Engineering of vitamin and cofactor synthesis in yeasts

Perli, T.

DOI
10.4233/uuid:b9241a98-34b9-4295-93e3-af8d4e796628

Publication date
2021

Document Version
Final published version

Citation (APA)
Perli, T. (2021). *Engineering of vitamin and cofactor synthesis in yeasts*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:b9241a98-34b9-4295-93e3-af8d4e796628

Engineering of vitamin and cofactor synthesis in yeasts

Thomas Perli

Engineering of vitamin and cofactor synthesis in yeasts

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 8 July 2021 at 10:00 o'clock

by

Thomas Perli

Master of Science in Cellular and Molecular Biotechnology
University of Trento, Italy
born in Trento, Italy

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
Dr.ir. J-M.G. Daran	Delft University of Technology, promotor

Independent members:

Prof.dr. F. Hollmann	Delft University of Technology
Dr. M. Krogh Jensen	Technical University of Denmark
Prof.dr. P. Branduardi	University Milano Bicocca
Prof.dr. D. Claessen	Leiden University
Dr. D. G. Weissbrodt	Delft University of Technology

Reserve member:

Prof.dr.ir. H. J. Noorman	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. The project was financed by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 722287.

The contents of this publication are the sole responsibility of the author and do not necessarily reflect the opinion of the European Union.

Cover	Elisa Godino
Layout	Thomas Perli
Printed by	ProefschriftMaken || www.proefschriftmaken.nl
ISBN	978-94-6423-329-2

Table of contents

Samenvatting

De toenemende wereldbevolking en het gebruik van niet-duurzame fossiele brandstoffen in de wereldeconomie zijn samen de belangrijkste factoren in de stijging van de gemiddelde temperatuur op aarde na het pre-industriële tijdperk. Overschakelen naar een duurzamere, circulaire en op biologische productie gebaseerde economie is een van de belangrijkste pijlers van het akkoord van Parijs om de uitstoot van broeikasgassen in de komende decennia te verminderen. Biotechnologie kan in deze context een cruciale rol spelen door het ontwikkelen van nieuwe, duurzamere processen voor de productie van voedsel, medicijnen, brandstoffen en chemicaliën.

Door de exponentiële daling van de kosten voor het lezen, bewerken en schrijven van DNA- en RNA-codes hebben vele wetenschappers en filantropen de 21ste eeuw alvast beschreven als de eeuw van de biologie. Deze nieuwe ontwikkelingen bieden uitgebreide mogelijkheden om microbiële gastheren aan te passen en te ontwerpen voor, bijvoorbeeld, het efficiënt recyclen van afvalstromen en de duurzame productie van chemische bouwstenen. Miljoenen jaren van evolutie hebben een enorm aantal micro-organismen (levensvormen) op aarde gevormd en geselecteerd, en als resultaat hiervan is hun stofwisseling specifiek geoptimaliseerd om te overleven in verschillende niche-omgevingen. Met het gebruik van recombinant-DNA-technologie kunnen wetenschappers profiteren van de aanwezige natuurlijke diversiteit van microbiële stofwisseling en kunnen ze verschillende gewenste eigenschappen combineren in een enkel organisme dat als productieplatform dienst doet. Bakkersgist (*Saccharomyces cerevisiae*) is niet alleen in de voedingsindustrie maar ook in de industriële biotechnologie een zeer populair micro-organisme en productieplatform. In de industriële biotechnologie wordt *S. cerevisiae* onder andere toegepast voor de efficiënte productie van verschillende biochemische producten, variërend van eenvoudige bulkchemicaliën zoals ethanol en melkzuur tot complexe hoogwaardige moleculen zoals opioïden en het antimalariamiddel artemisinine. Tot de belangrijkste kenmerken die *S. cerevisiae* tot een aantrekkelijke microbiële gastheer maken behoren toegankelijkheid voor genetische manipulatie, het vermogen om ook in de afwezigheid van zuurstof met hoge snelheden te groeien, en robuustheid onder stressvolle procesomstandigheden.

Ondanks deze voordelen zijn de eerdergenoemde productieprocessen van verschillende biochemische producten vaak nog verre van optimaal en is er nog een aantal beperkingen dat overwonnen moet worden. Een van de belangrijkste uitdagingen bij het economisch aantrekkelijk maken van bioprocesalternatieven is het optimaliseren van de productopbrengst en het elimineren van bijproductvorming. Om de productopbrengst te maximaliseren is het nodig om de intracellulaire beschikbaarheid van de biochemische

intermediairen te verhogen. Daarnaast moet de microbiele gastheer bestand zijn tegen stressomstandigheden en verontreinigingen die aanwezig zijn in grondstoffen voor tweede- of derde-generatie industriële processen. Het succesvol tot expressie brengen van nieuwe biochemische routes voor specifiek producten kan vereisen dat ook genetische informatie voor synthese of opname van nieuwe cofactoren wordt geïntroduceerd. In deze context kan het optimaliseren van biosyntheseroutes voor natieve vitamin- en cofactorbiosyntheseroutes, alsmede het introduceren van heterologe routes voor *de novo* cofactorbiosynthese, een belangrijke rol spelen bij het realiseren van reductie van proceskosten en een hogere robuustheid van processen.

In **Hoofdstuk 1** wordt de rol van vitamines uit de B-groep, hun biosynthese en de regulatie hiervan in bakkersgist gedocumenteerd. Ook analyseert dit hoofdstuk de aanwezigheid van genen die betrokken zijn bij de biosynthese van deze vitamines in verschillende *Saccharomyces* gistsoorten, en bespreekt het eerder onderzoek aan, en nieuwe mogelijkheden voor, de uitbreiding van de nu beschikbare set van cofactoren die van nature door gisten wordt geproduceerd.

In **Hoofdstuk 2** werden de specifieke vitaminebehoeften van de populaire *S. cerevisiae* laboratoriumstam CEN.PK113-7D gekarakteriseerd door groeisnelheden te bestuderen in synthetische media waarin steeds één vitamine werd weggelaten. De resultaten van dit onderzoek toonden aan dat het verwijderen van de vitaminen inositol en nicotinezuur uit het media de groeisnelheid van de *S. cerevisiae*-stam niet beïnvloedde. Laboratoriumevolutie werd met succes toegepast om nieuwe gistisolaten te verkrijgen die zonder de vitaminen para-aminobenzoëzuur, pantotheenzuur, thiamine of pyridoxine in het media even snel groeiden als de originele stam in aanwezigheid van alle vitaminen. Analyse van de DNA-volgorde van de genomen van de geëvolueerde isolaten leidde tot nieuwe inzichten in de regulatie van de biosynthese van vitamines. Het introduceren van de mutaties die in geëvolueerde stammen werden aangetroffen in de originele, niet-geëvolueerde stam, leidde tot een substantiële toename van de specifieke groeisnelheid in niet-gesupplementeerde media.

De novo biosynthese van de vitamines thiamine, nicotinamide dinucleotide (NAD^+) en pantotheenzuur in *S. cerevisiae* vereist de aanwezigheid van zuurstof, waardoor toevoeging van deze vitamines een strikt vereiste is voor anaërobe culturen van gist. Daarom werd in **Hoofdstuk 3** onderzocht of de zuurstofonafhankelijke synthese van zowel pantothenaat en NAD^+, zoals die voorkomt in obligaat anaërobe pensschimmels uit het fylum Neocallimastigomycota, kon worden geïntroduceerd *S. cerevisiae*. Het L-aspartaatdecarboxylase (adc) die betrokken is bij de zuurstofonafhankelijke panothenaatsynthese in deze organismen werd geidentificeerd met behulp van genoom- en fylogenetische analyse. Van twee eerder geïdentificeerde genen betrokken bij zuurstofonafhankelijke synthese van NAD^+ in Neocallimastigomycota is beschreven dat ze

zijn verworven door horizontale genoverdracht. Deze beide genen werden samen met het geïdentificeerde adc-enzym geïntroduceerd in *S. cerevisiae* knock-outstammen waarin de natuurlijke cofactor-syntheseroutes waren uitgeschakeld. Deze genetisch gemodificeerde stammen waren in staat om snel te groeien in de afwezigheid van zuurstof in media zonder nicotine of pantothenaat. Deze waarneming toonde aan dat expressie van de heterologe enzymen de zuurstofvereisten van de oorspronkelijke syntheseroutes van deze cofactoren omzeilden. Bovendien resulteerde dit werk, waarin gist als modelsysteem werd gebruikt, tot een beter inzicht in de evolutionaire aanpassingen aan anaërobe omgevingen en de selectieve voordelen van gepostuleerde horizontale genoverdrachtgebeurtenissen in de evolutionaire geschiedenis van anaërobe schimmels.

Het doel van **Hoofdstuk 4** was het functioneel tot expressie brengen van de Molybdeen-cofactor (Moco) biosyntheseroute en de Moco-afhankelijke nitraatassimilatieroute in *S. cerevisiae*. Het succesvol tot expressie brengen van de Moco-biosynthese opent de mogelijkheid om een hele nieuwe klasse van Moco-afhankelijke enzymen in *S. cerevisiae* te gebruiken. In dit hoofdstuk werd eerst de Moco-biosyntheseroute van de nitraat-assimilerende gist *Ogataea parapolymorpha* functioneel gekarakteriseerd door met CRISPR-Cas9 deletiemutaties aan te brengen in vermoedelijk betrokken genen. De genen die noodzakelijk waren voor Moco-biosynthese, molybdaattransport en nitraatassimilatie werden daarna tot expressie gebracht in *S. cerevisiae*. Laboratoriumevolutie van de aldus gemodificeerde giststammen voor groei op nitraatbevattende media resulteerde in stammen met hogere Moco-afhankelijke nitraatreductase-activiteit. Deze verandering kon worden toegewezen aan een verhoogd expressieniveau van de heterologe genen. Het introduceren van nitraatassimilatie in *S. cerevisiae* leidde tot een hogere robuustheid tegen contaminaties met de gist *Brettanomyces bruxellensis*, die vaak als contaminatie gevonden wordt in industriële processen.

Naast *S. cerevisiae* is ook *Yarriowia lipolytica* een industrieel relevante gist, die bovendien als model wordt gebruikt voor het bestuderen van vetstofwisseling. In **Hoofdstuk 5** werd aangetoond dat dezelfde genen uit *O. parapolymorpha* die in Hoofdstuk 4 Moco-afhankelijke nitraatassimilatie mogelijk maakten in *S. cerevisiae*, ook in *Y. lipolytica* werkten. Deze vinding suggereert dat het mogelijk zou kunnen zijn om Moco-afhankelijke enzymen, zoals de nitraatreductase, in elke andere gistsoort tot expressie te brengen. De gemodificeerde *Y. lipolytica*-stam kon door gebruik van adaptieve laboratoriumevolutie op nitraatbevattende media worden verbeterd. Het bepalen van de genoomsequentie van de hieruit voortkomende geëvolueerde isolaten leverde nieuwe inzichten in mogelijke limitaties in de biochemische route en duidde op een mitochondriële lokalisatie van de eerste stap van de Moco-biosynthese in *Y. lipolytica*.

Summary

The increase in world population together with an unsustainable and fossil fuel-based economy are at the root of the significant increase in the average global temperature since the pre-industrial era. Switching to a more sustainable and circular bio-based economy is one of the main pillars included in the Paris agreement aimed at reducing green-house gasses emissions in the next decades. Biotechnology promises to play a crucial role in this context by providing novel, more sustainable processes for the production of food, drugs, fuels and chemicals.

Many scientists and philanthropists have already described the 21st century as the century of biology. With the exponential decrease in costs of reading, editing and rewriting the code of life (DNA/RNA), biotechnologists are now equipped with a range of tools that enables them to engineer microbial hosts for the efficient recycling of waste streams and/or for the sustainable production of chemical building blocks. Millions of years of evolution have shaped and selected a massive number of life forms on Earth and, as result, their metabolism has been specifically optimized to survive in different niche environments. With the use of recombinant DNA technology, scientists can take advantage of such metabolic diversity and transfer it to other organisms of choice to combine advantageous characteristics in a single production host.

Saccharomyces cerevisiae is a highly popular microorganism, not only in the food fermentation industry but also in industrial biotechnology, where it is applied for the efficient production of several products ranging from simple bulk chemicals such as ethanol and lactic acid to complex, high-value molecules such as opioids and the anti-malarian drug artemisinin. Among the main characteristics that make *S. cerevisiae* an attractive microbial host are its genetic accessibility, its ability to easily ferment at high rates, in large volumes, its ability to grow in the absence of oxygen, and its robustness against stressful process conditions. However, as good as these advantage sounds, the reality is that, in many cases, individual yeast species are still far from being an optimal production hosts and there are a number of limitations that still need to be overcome. One of the main challenges in making bioprocess alternatives economically attractive concerns the system-level optimization and resource channelling toward the desired product in order to maximize product yield and to minimize the formation of by-products. For example, the intracellular availability of precursors needs to be maximized to boost activity of desired product pathway. Moreover, strain robustness against stress conditions and process contamination need to further improve in order to take full advantage of the less pure second or third generation industrial media. Finally, the introduction of new product

pathways may require the parallel engineering biosynthesis of a non-native cofactor or its transport in order to achieve the functional expression of heterologous enzymes. In this context, the optimization of native vitamin and cofactor biosynthesis pathways and the engineering of heterologous *de novo* cofactor biosynthesis pathways might play a crucial role for enabling process cost reduction, increased process robustness and introduction of new metabolic routes in engineered microbial hosts.

Chapter 1 provides an introduction to this research field by reviewing the roles of B-group vitamins in baker's yeast, their biosynthesis and regulation. In addition, this chapter analyses the occurrence of genes involved in biosynthesis of these vitamins in members of the *Saccharomyces* genus, and discusses previous work and possible research opportunities regarding the expanding the set of cofactors that is naturally produced by yeasts.

In **Chapter 2**, the specific vitamin requirements of the popular *S. cerevisiae* laboratory strain CEN.PK113-7D were assessed by studying growth performance in synthetic media variants from which a single vitamin at a time had been omitted. Results showed that removal of inositol and nicotinic acid from the media did not affect the growth rate of this *S. cerevisiae* strain. Evolutionary engineering was successfully applied to obtain isolates that were able to grow as fast in media lacking either *para*-aminobenzoic acid, pantothenic acid, thiamine, or pyridoxine as the parental strain in complete synthetic medium. Whole-genome sequencing of the evolved isolates revealed new insights in the regulation of vitamin biosynthesis. Reverse engineering of a few selected mutations in the unevolved background strain was sufficient to achieve a substantial increase in its specific growth rate in non-supplemented media.

De novo biosynthesis of thiamine, nicotinamide dinucleotide (NAD^+), and pantothenate in *S. cerevisiae* requires the presence of oxygen, thus imposing strict vitamin requirements when growing yeast under anaerobic conditions. In **Chapter 3**, alternative pathways for the oxygen-independent synthesis of pantothenate and NAD^+, sourced from obligate anaerobic gut fungi belonging to the *Neocallimastigomycota* phylum, were introduced in *S. cerevisiae*. Genomic analysis followed by phylogenetic analysis helped to identify the l-aspartate decarboxylase (adc) gene responsible for oxygen-independent pantothenate synthesis in these organisms and shed new light on its evolutionary origin. Previously identified genes for oxygen-independent synthesis of NAD^+ reported to be acquired by a horizontal gene transfer event, together with the identified adc enzyme, were introduced in *S. cerevisiae* knock-out strains devoid of essential steps in the native cofactor-synthesis pathways. The resulting engineered strains were able to grow fast in the absence of oxygen in media lacking either nicotinate or pantothenate, thus demonstrating that the heterologously expressed enzymes by-passed oxygen requirements of the native synthesis pathways for these cofactors. Moreover, this work demonstrated how heterologous expression studies in yeast can provide more insights into evolutionary adaptation to

anaerobic environments and into selective advantages conferred by horizontal gene transfer events in difficult-to-cultivate anaerobic fungi.

In **Chapter 4**, the functional expression of a nitrate assimilation pathway was achieved in *S. cerevisiae* by the parallel expression of a heterologous *de novo* biosynthesis pathway for a Molybdenum cofactor (Moco). This result opened up the possibility to use of an entire new enzyme family in baker's yeast. In this work, Moco biosynthesis was first functionally characterized in the nitrate-assimilating yeast *Ogataea parapolymorpha* by introducing single-knockout mutations in putative genes involved in Moco biosynthesis with CRISPR/Cas9 gene editing, followed by functional analysis of the resulting mutant strains. Then, a set of 11 selected genes, including 7 Moco biosynthesis genes, 1 high-affinity molybdate-transporter and 3 genes encoding nitrate-assimilation enzymes were expressed in *S. cerevisiae*. Evolution of the engineered strains on nitrate-containing medium resulted in higher Moco-dependent nitrate-reductase activity, which could be attributed to an increased gene dosage of heterologously expressed *O. parapolymorpha* genes and an increased level of their protein products. The industrial relevance of a nitrate-assimilating *S. cerevisiae* strain was confirmed by showing how the use of a nitrate-assimilating *S. cerevisiae* strain led to increased robustness against contaminations by the spoilage yeast *Brettannomyces bruxellensis*.

In **Chapter 5**, we demonstrated that the same set of *O. parapolymorpha* genes as studied in Chapter 4 could confer Moco-dependent nitrate assimilation in the evolutionarily distant, industrially relevant and lipid accumulating yeast *Yarrowia lipolytica*. The engineered strain was also shown to express an active Moco-dependent nitrate reductase. This observation, by extrapolation, suggested that it may be possible to express Moco-dependent enzymes in any yeast species. By performing adaptive laboratory evolution on nitrate-containing media, followed by genome sequencing of evolved isolates, further insights into possible pathway limitations were found, with specific focus to the mitochondrial localization of the first step in Moco biosynthesis.

Chapter 1: Introduction

Adapted from the publication entitled
"Vitamin requirements and biosynthesis in *S. cerevisiae*",
by Thomas Perli*, Anna K. Wronska*, Raúl A. Ortiz-Merino,
Jack T. Pronk, Jean-Marc Daran
*These authors contributed equally to this work.
Yeast 2020; 37: 283-304.
https://doi.org/10.1002/yea.3461

Introduction

"No animal can live on only pure protein, fat, and carbohydrates, but other dietary factors are required for life" [1]. This observation eventually led to the vitamine (later changed to vitamin) theory established by Casimir Funk [2]. An organic compound is defined as a vitamin if it is essential, cannot be synthesized by the organism itself, and therefore needs to be taken up from the environment [3]. Whether a compound is a vitamin therefore depends on the organism studied and, potentially, on growth conditions.

Chemically defined media for cultivation of yeasts (CDMY) are essential for fundamental as well as applied research. In contrast to complex media, which contain non defined components such as yeast extract and/or peptone, defined media enable the generation of highly reproducible data, independent variation of the concentrations of individual nutrients and, in applied settings, design of balanced media for high-biomass-density cultivation and application of defined nutrient limitation regimes. The use of CMDY prevents thus unwanted variations. Lot to lot variation of the complex raw materials as yeast extract may lead to up to 50% difference in growth rate and biomass levels that [4, 5]. The control of process variability is not only crucial to be in line with FDA regulations but also for maintaining high productivity and maximize process economics [6]. The CDMY that are now used in yeast research laboratories around the world are based on an early investigation of the requirements for riboflavin (B_2), biotin (B_7), thiamine (B_1), pyridoxine (B_6), inositol (B_8), nicotinic acid (B_3) and pantothenate (B_5) of over a hundred yeast species [7]. With the exception of riboflavin, which could be universally omitted, yeast species exhibited diverse auxotrophy patterns for the remaining six compounds, which were therefore included in the first CDMY. *para*-Aminobenzoic acid (*p*ABA, formerly referred to as vitamin B_{10}) was later added as it was found to stimulate growth of brewing yeasts [8]. These seven compounds with riboflavin (vitamin B_2) and folate (vitamin B_9) are still included in the widely used CDMY known as Yeast Nitrogen Base [9, 10] (YNB; Table 1). The concentration of the vitamins contained in YNB have been empirically defined but without quantitative assessing the exact yeast requirement [11]. In another popular CDMY, often referred to as Verduyn medium (Table 1), concentrations of media components were adjusted to support yeast biomass concentrations up to 10 g L^{-1} in aerobic, glucose-limited cultures that exhibit a fully respiratory metabolism [12, 13].

Athough meant to suit all *S. cerevisiae* strains, it may happen that in specific growth conditions or for specific strains, these recipes have to be adjusted. Strains of the popular *S. cerevisiae* BY lineage [15] require additional inositol to support fast growth until glucose exhaustion in YNB medium [16]. Inositol concentration represents one of main difference between the YNB and the Verduyn media, the latest containing an inositol (B_8) concentration 12.5-fold higher to prevent occurrence of undesired growth limitation. (Table 1). Information of yeast biomass vitamin content (per $gram_{DW}$) would allow to

prepare tailor-made media based on exact nutritional requirements. However, data of intracellular vitamin concentrations remain scarse and difficult to compare. As an example the range of measured intracellular biotin concentration in *S. cerevisiae* varies by order of magnitude likely influenced by the used detection method that oscilates between bioassay based on growth of an auxotroph organism (1.4-1.5 µg/g) [17], immunodetection (0.053 0.004 ng/g) [18] or liquid chromatography. It is obvious that more complete and accurate quantitative information regarding intracellular vitamin concentration is needed. This knowledge will be key to further understand the physiological role of class B vitamin in yeast metabolism.

Based on their essentiality in the human diet, the molecules precedently mentioned can all be classified as B vitamins, which are water-soluble compounds involved in cell metabolism. However, as will be discussed below, they have widely different chemical structures and roles in cellular metabolism [3]. Early studies already demonstrated that growth of some yeasts, including *Saccharomyces* species, was not strictly dependent on addition of all of these compounds, although omission of individual compounds might result in sub-optimal growth [19-21]. These observations suggested that these yeast strains could *de novo* synthesize some of these compounds, in which case they should formally not be referred to as vitamins but, if their addition leads to improved growth, as growth factors.

It is well established that vitamin and/or growth-factor requirements of yeasts are not only species dependent, but can also strongly vary with growth conditions. In particular, ergosterol and unsaturated fatty acids, whose synthesis by *S. cerevisiae* requires molecular oxygen, are routinely included in CDMY for anaerobic yeast cultivation [22, 23]. These anaerobic nutritional requirements of yeasts are addressed in several dedicated reviews [24, 25] and will not be discussed here. For information on the applications and physiological impacts of artificially introduced auxotrophic requirements in *S. cerevisiae*, readers are referred to a previous minireview [26].

The present paper aims to review current knowledge on the capability of *S. cerevisiae* for *de novo* synthesis of the seven 'vitamins' that are commonly added to CDMY and on the pathways and genes involved in their biosynthesis. Riboflavin (B_2) and folic acid (B_9) that are only present in YNB will not be discussed further. *S. cerevisiae* and more generally Saccharomycotina yeasts are B_2 prototroph under both aerobic and anaerobic conditions [7]. Folic acid (B_9) synthesis depends on *p*ABA as a rate limiting precursor, whose *de novo* synthesis and metabolic implication are reviewed below. In addition, based on the existing knowledge on *S. cerevisiae* and a comparative analysis of the genomes of *Saccharomyces* species, we present a brief assessment of the distribution of these metabolic pathways across *Saccharomyces* species.

Table 1: Composition of Yeast Nitrogen Base (YNB) [14] and Synthetic Media (SM) [13] for aerobic growth. Values are for 1 L of media.

		Yeast Nitrogen Base w/o amino acids	Verduyn Synthetic Media
Nitrogen source	Ammonium sulfate ($(NH_4)_2SO_4$)	5 g	5 g
Salts	Potassium phosphate monobasic (KH_2PO_4)	850 mg	3 g
	Potassium phosphate dibasic (K_2HPO_4)	150 mg	
	Sodium chloride (NaCl)	100 mg	
	Calcium chloride ($CaCl_2$)	100 mg	3.39 mg
	Boric acid (H_3BO_3)	0.5 mg	1 mg
	Copper sulfate ($CuSO_4$)	0.04 mg	0.19 mg
	Cobalt chloride ($CoCl_2$)		0.16 mg
	Potassium iodide (KI)	0.1 mg	0.1 mg
	Ferric chloride ($FeCl_3$)	0.2 mg	
	Iron sulfate heptahydrate ($FeSO_4 \cdot 7\ H_2O$)		3 mg
	Magnesium sulfate ($MgSO_4$)	0.5 g	0.244 g
	Manganese chloride ($MnCl_2$)		0.64 mg
	Manganese sulfate ($MnSO_4$)	0.4 mg	
	Sodium molybdate (Na_2MoO_4)	0.2 mg	0.34 mg
	Zinc sulfate ($ZnSO_4$)	0.4 mg	2.53 mg
	EDTA		15 mg
Growth factors	Biotin	0.002 mg	0.05 mg
	Calcium pantothenate	0.4 mg	1 mg
	Folic acid	0.002 mg	
	Inositol	2 mg	25 mg
	Nicotinic acid	0.4 mg	1 mg
	para-Aminobenzoic acid	0.2 mg	0.2 mg
	Pyridoxine	0.4 mg	0.82 mg
	Riboflavin	0.2 mg	
	Thiamine	0.32 mg	0.79 mg

Vitamins that act as enzyme cofactors

Pyridoxine (B_6)

Pyridoxine (PN), pyridoxal (PL), pyridoxamine (PM) and their phosphorylated derivatives pyridoxine 5'-phosphate (PNP) and pyridoxamine 5'-phosphate (PMP) can be interconverted intracellularly and together form the B6 vitamers. A vitamer is defined as a molecule having a similar structure and the same nutritional impact as the biologically active form of the vitamin. Pyridoxine was isolated and synthesized after its identification as a substance preventing dermatitis in rats [27-29]. Its chemical structure is characterized by a tetra-substituted pyrimidine ring with one methyl, one hydroxyl and two methyl-hydroxyl groups (Figure 1). Pyridoxine was first reported to stimulate yeast growth in 1939 [30]. Although mainly supplied to CDMY as the vitamer pyridoxine, pyridoxal 5'-phosphate (PLP) is the active form. PLP serves as coenzyme and/or substrate for at least 50 *S. cerevisiae* enzymes involved in amino-acid, glucose and lipid metabolism, as well as in thiamine biosynthesis and regulation (Table 2).

PLP formation from PM, PN or PL involves a salvage pathway that is widespread in nature [89]. These three vitamers can be imported in *S. cerevisiae* by the high-affinity proton symporter Tpn1 [90]. In the yeast cytosol, PN, PM and PL are phosphorylated to form PNP, PMP and PLP, respectively, most probably by the putative pyridoxine kinase Bud16. PNP and PMP are subsequently oxidized to PLP by the pyridoxine oxidase Pdx3 [91].

De novo synthesis of PLP by *S. cerevisiae* [92] involves a single reaction catalyzed by PLP synthase, which is a heterodimeric enzyme [93] (Figure 1). Its glutamine-hydrolase subunit (Sno) catalyzes the hydrolysis of L-glutamine, producing L-glutamate and ammonia [94]. Ammonia generated in this reaction is not released from the enzyme, but channeled to the active site of the synthase subunit (Snz), which condenses it with D-ribulose 5-phosphate and D-glyceraldehyde 3-phosphate to yield PLP [95]. The Snz protein not only catalyses PLP formation but also isomerizes dihydroxyacetone-phosphate and ribose-5-phosphate to glyceraldehyde-phosphate and ribulose-5-phosphate, respectively, with the latter being the favoured substrate for PLP formation [96].

The *S. cerevisiae* genome carries three members of the *SNO* and *SNZ* genes familes (*SNO1,2 and 3, SNZ1,2 and 3*). These *SNO* and *SNZ* genes form colocalized gene pairs, each expressed from a single bi-directional promoter. The *SNZ1/SNO1* pair has been shown to be involved in *de novo* PLP biosynthesis and its transcription is activated in late stationary phase [97]. Transcriptional activation of *SNZ1/SNO1* under amino acid starvation, mediated by the Gcn4 master regulator, is consistent with the PLP requirement of aminotransferases [98]. The *SNZ1/SNO1* gene pair is coregulated by the adenine and histidine biosynthesis transcription factor Bas1 [99-101] in the presence of glycine [102]. In contrast to the *SNZ1/SNO1* gene pair, which is located in the middle of the right arm of CHRXIII, *SNZ2*/*SNO2* and *SNZ3*/*SNO3* are found in sub-telomeric regions of CHRXIV

Table 2: *S. cerevisiae* S288C proteins requiring pyridoxal-5-phosphate, thiamine diphosphate and biotin as cofactor or as substrate.

Cofactor	Protein	Protein name
Pyridoxal-5-phosphate	Uga1	4-aminobutyrate aminotransferase [31]
	Hem1	5-aminolevulinate synthase [32]
	Arg8	Acetylornithine aminotransferase [33]
	Bio3	Adenosylmethionine-8-amino-7-oxononanoate aminotransferase [34]
	Agx1	Alanine-glyoxylate aminotransferase 1 [35]
	Abz2	Aminodeoxychorismate lyase [36]
	Aro9	Aromatic amino acid aminotransferase 2 [37]
	Aro8	Aromatic/aminoadipate aminotransferase [38]
	Aat2	Aspartate aminotransferase 2 [39]
	Aat1	Aspartate aminotransferase 1 [40]
	Bat2	Branched-chain-amino-acid aminotransferase 2 [41]
	Bat1	Branched-chain-amino-acid aminotransferase 1 [41]
	Cha1	Catabolic L-serine/threonine dehydratase [42]
	Str3	Cystathionine beta-lyase [43]
	Cys4	Cystathionine beta-synthase [44]
	Cys3	Cystathionine gamma-lyase [45]
	Str2	Cystathionine gamma-synthase [46]
	Nfs1	Cysteine desulfurase [47]
	Dsd1	D-serine dehydratase [48]
	Gad1	Glutamate decarboxylase [49]
	Gcv2	Glycine dehydrogenase [50]
	Gph1	Glycogen phosphorylase [51]
	His5	Histidinol-phosphate aminotransferase [52]
	Met17	Homocysteine/cysteine synthase [53]
	Bna5	Kynureninase [54]
	Sry1	L-threo-3-hydroxyaspartate ammonia-lyase [55]
	Gly1	Low specificity L-threonine aldolase [56]
	Car2	Ornithine aminotransferase [57]
	Spe1	Ornithine decarboxylase [58]
	Ser1	Phosphoserine aminotransferase [59]
	Alt2	Probable alanine aminotransferase 2 [60]
	Alt1	Probable alanine aminotransferase 1 [61]
	Bna3	Probable kynurenine-oxoglutarate transaminase [62]

	Irc7	Putative cystathionine beta-lyase [62]
	Yll058w	Putative cystathionine gamma-synthase [63]
	Yml082w	Putative cystathionine gamma-synthase [46]
	MCY1	Putative cysteine synthase [64]
	Shm2	Serine hydroxymethyltransferase 2 [65]
	Shm1	Serine hydroxymethyltransferase 1 [65]
	Lcb1	Serine palmitoyltransferase 1 [66]
	Lcb2	Serine palmitoyltransferase 2 [66]
	Dpl1	Sphingosine-1-phosphate lyase [67]
	Ilv1	Threonine dehydratase [68]
	Thr4	Threonine synthase [69]
	Trp5	Tryptophan synthase [70]
	Yhr112c	Uncharacterized trans-sulfuration enzyme [71]
	Thi5	4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) synthase [72]
	Thi11	4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) synthase [72]
	Thi12	4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) synthase [72]
	Thi13	4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) synthase [72]
Thiamine diphosphate	Kgd1	2-oxoglutarate dehydrogenase [73]
	Ilv2	Acetolactate synthase catalytic subunit [74]
	Pxp1	Putative 2-hydroxyacyl-CoA lyase [75]
	Pdc1	Pyruvate decarboxylase isozyme 1 [76]
	Pdc5	Pyruvate decarboxylase isozyme 2 [77]
	Pdc6	Pyruvate decarboxylase isozyme 3 [78]
	Pda1	Pyruvate dehydrogenase E1 component subunit alpha [79]
	Pdb1	Pyruvate dehydrogenase E1 component subunit beta [80]
	Thi3	Thiamine metabolism regulatory protein THI3 [81]
	Aro10	Transaminated amino acid decarboxylase [82]
	Tkl1	Transketolase 1 [83]
	Tkl2	Transketolase 2 [84]
Biotin	Acc1	Acetyl-CoA carboxylase [37]
	Hfa1	Acetyl-CoA carboxylase, mitochondrial [85]
	Bpl1	Biotin protein ligase [86]
	Pyc1	Pyruvate carboxylase 1 [87]
	Pyc2	Pyruvate carboxylase 2 [87]
	Dur1,2	Urea amidolyase [88]

and VI, respectively and are flanked by the thiamine biosynthetic genes *THI12* and *THI5*, respectively. Their increased expression upon

Thiamine (B_1)

Thiamine, also known as vitamin B1, was first isolated by Jansen and Donath [105] and later obtained in sufficient amounts for extended chemical analysis [106]. In animals, which cannot synthesize thiamine, a lack of dietary supply causes beriberi, a disease affecting the nervous system [107]. Thiamine is essential for cellular energy metabolism and its major biologically active derivative thiamine diphosphate (TDP) serves as cofactor for a variety of enzymes, including pyruvate and oxoglutarate dehydrogenases, transketolases, 2-hydroxy-3-oxoadipate synthase, acetolactate synthase and 2-oxo acid decarboxylases (Table 2). As an electrophilic cofactor, TDP forms covalent intermediates with enzyme substrates. Thiamine can also perform intramolecular proton transfers, which is a rare function among cofactors [108]. It has been proposed that a general stress-protective role of thiamine in *S. cerevisiae* is partially unrelated to its role as a cofactor [109]. Thiamine is synthesized *de novo* by plants and many microorganisms including yeast species.

Thiamine consists of two substituted aromatic moieties, 4-amino-2-methyl-5-pyrimidyl (HMP) and 5-(2-hydroxylethyl)-4-methylthiazolium (HET), which are connected by a methylene bridge (Figure 1). In addition to free thiamine and the biologically active form TDP, thiamine monophosphate (TMP), and thiamine triphosphate are also found intracellularly. All thiamine-prototrophic organisms synthesize TDP via condensation of the precursors 5-(2-hydroxylethyl)-4-methyl thiazole phosphate and 4-amino-2-methyl-5-pyrimidine diphosphate (HMP-PP) to TMP by TMP diphosphorylase (Thi6 in *S. cerevisiae*) (Figure 1). Although bacteria can synthesize TDP from TMP in a single reaction, catalysed by a TMP kinase, eukaryotes utilize a pathway in which TMP is first dephosphorylated to thiamine, which is then pyrophosphorylated to TDP by a thiamine pyrophosphokinase (Thi80 in *S. cerevisiae*) [110]. Two transporters involved in the acquisition of exogenous thiamine have been identified in *S. cerevisiae*: a high-affinity transporter encoded by *THI10* [111] and a periplasmic acid phosphatase encoded by *PHO3* that releases thiamine from thiamine phosphates [112].

In *S. cerevisiae,* the thiamin precursor HMP-PP is synthesized in two steps. First, 4-amino-2-methyl-5-pyrimidine phosphate (HMP-P) is formed from pyridoxal-5-phosphate and histidine. The histidine used for HMP-P synthesis is provided from the active site of HMP-P synthase in a suicide reaction [72, 113]. HMP-P synthase is encoded by four highly similar *S. cerevisiae* genes (*THI5*, *THI11*, *THI12,* and *THI13*). These genes are located in sub-telomeric regions of different chromosomes, suggesting that an increased copy number conferred a selective advantage in thiamine-poor environments [114]. In a second step, HMP-P is phosphorylated to HMP-PP by HMP-P kinase in an adenosine

triphosphate-dependent reaction [115]. The *S. cerevisiae* genome harbors two paralogous genes encoding HMP-P kinase, *THI20,* and *THI21*, of which the former encodes the major isoform [114]. Thi20 is a trifunctional protein that displays thiamine biosynthesis and thiamine degradation activities in a single protein. Its N-terminal domain is active as HMP and HMP-P kinase, while its C-terminal domain has thiaminase II activity [116]. Although molecular oxygen is not directly required for HMP biosynthesis, activity of this branch of the thiamine biosynthetic pathway was shown to be oxygen dependent [114]. However, based on gene deletion studies it has been proposed that *S. cerevisiae* can still synthesize the pyrimidyl moiety under anaerobic conditions via an alternative, as yet unidentified, pathway [117].

For the synthesis of the thiazole moiety, eukaryotic cells use a single enzyme to form 5-(2-hydroxylethyl)-4-methyl thiazole phosphate from glycine and nicotinamide adenine dinucleotide (NAD^+), encoded by *THI4* in *S. cerevisiae* [118]. Thi4 acts as a substrate in the reaction by providing the sulfur atom needed for thiazole formation in an iron-dependent sulfide transfer from a conserved cysteine. Therefore, similar to Thi5, Thi4 acts as a suicide enzyme undergoing only a single catalytic turnover [119-121]. Under thiamine-depleted conditions, Thi5 and Thi4 are among the most abundant proteins in *S. cerevisiae* [122]. Strains harboring a *THI4* deletion have an increased sensitivity to DNA damaging agents such as UV light and methyl methanesulfonate. The mechanism of this protection is not fully understood [114, 123].

Involvement of two suicide enzymes makes *de novo* thiamine biosynthesis in yeast an energetically very expensive process: for each molecule of thiamine produced, two complete proteins (Thi4 and Thi5/11/12/13) have to be synthesized and degraded. Tight regulation of thiamine synthesis occurs mainly at the transcriptional level [124, 125]. As a result, the *THI* regulon is repressed in the presence of high intracellular levels of TDP. A strain carrying a partially inactivated form of Thi80 was shown to constitutively express the *THI* genes, suggesting that TDP is the molecule acting in this negative feedback regulation loop [126]. Three positive regulators for thiamine biosynthesis have been identified to date: Thi2, Thi3, and Pdc2 [81, 127, 128]. Elimination of any of these three proteins abolishes *THI* genes expression. The expression of *THI2* and *THI3*, but not *PDC2*, strongly increased under thiamine-depleted conditions [129]. Deletion of *THI2* results in repression of all *THI* genes except for *THI10*, whereas deletion of *THI3* causes repression of all *THI* genes. Thi3, which binds TDP, was originally proposed to also act as a 2-oxo acid decarboxylase involved in the Ehrlich for fusel alcohol biosynthesis [130] but this conclusion was later refuted [131-133]. A strain that only carried a *thi3* allele encoding a protein unable to bind TDP showed constitutive expression of *THI* genes in thiamine-containing medium, suggesting that Thi3 acts as a TDP sensor. However, Thi3 lacks a clear DNA-binding motif and is likely to act through interaction with other proteins, such as

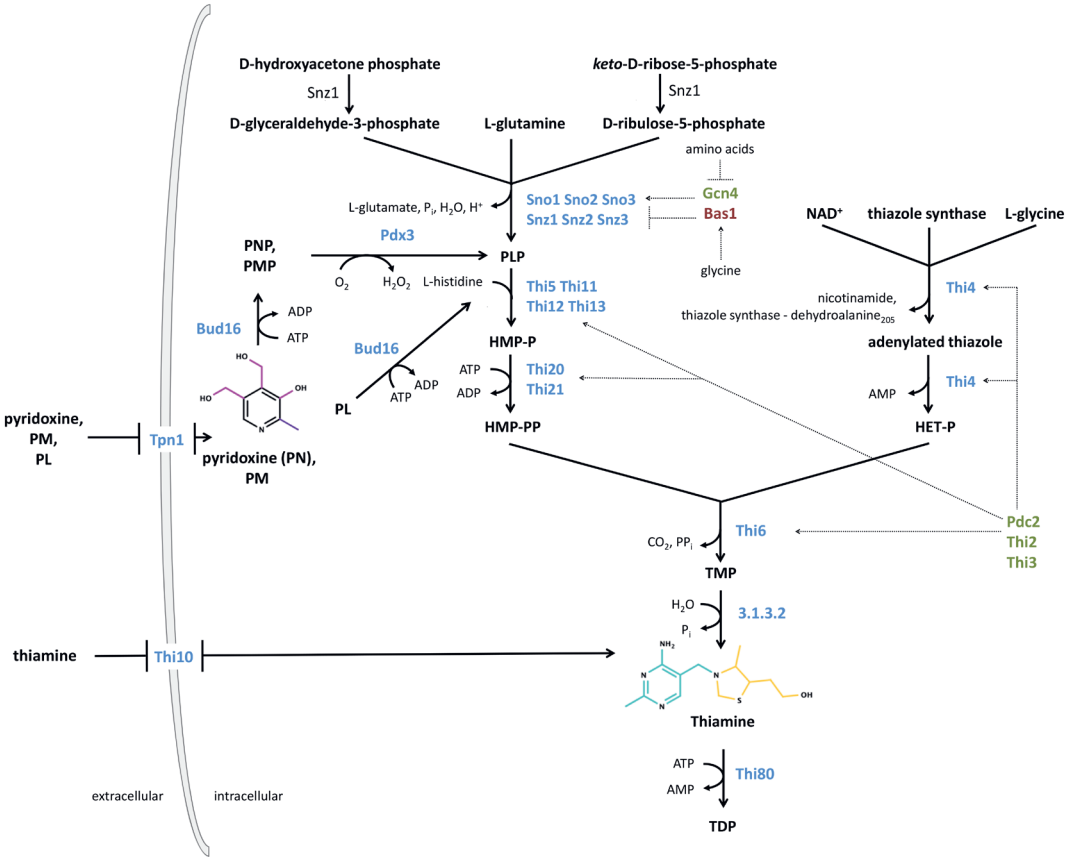


Figure 1: PLP and TDP *de novo* synthesis pathway in *S. cerevisiae*. D-glyceraldehyde 3-phosphate, L-glutamine and *keto*-D-ribose 5-phosphate are converted to PLP by the catalytic activity of the gene products of *SNO1,2,3* and *SNZ1,2,3*. Gcn4 acts as positive regulator of *de novo* PLP biosynthesis whereas Bas1 acts as an inhibitor. Gcn4 is inhibited by amino acids and activated under amino-acid starvation. Bas1 instead is upregulated in the presence of glycine. PN, PM and PL are imported by Tpn1. PN is converted at the expense of ATP to PNP by Bud16 whereupon Pdx3, produces PLP and hydrogen peroxide in an oxygen-dependent reaction. Similarly, PLP can be formed starting from PM in two steps by action of Bud16 and Pdx3, with PMP as intermediate. Moreover, PL can also be converted to PLP by action of Bud16. PLP is used as cofactor or converted HMP-P by one of the four homologous enzymes Thi5, Thi11, Thi12 and Thi13, under consumption of L-histidine. HMP-P is the intermediate for the formation of the pyrimidyl moiety of thiamine (shown in cyan). Thi20 and Thi21 further phosphorylate HMP-P to 4-amino-2-methyl-5-pyrimidine diphosphate (HMP-PP). The thiazole moiety (shown in yellow) is synthesized by activity of Thi4 in a suicide mechanism, leading to HET-P. HMP-PP and HET-P are merged by the gene product of *THI6* to TMP. The following reaction catalysed by an acid phosphatase (EC number 3.1.3.2) yields thiamine. Thiamine can be taken up with the aid of the transporter Thi10. Finally, thiamine is converted to its biologically active form thiamine diphosphate (TDP) under consumption of ATP by Thi80. Pdc2, Thi2 and Thi3 are responsible for the upregulation of transcription of Thi5/11/12/13, Thi20/21, Thi6 and Thi4. Alcohol and methyl substitutions on the pyridoxine pyrimide ring are shown in magenta

and purple respectively. Metabolites, proteins and positive regulators are shown in bold, blue and green respectively. ATP, adenosine triphosphate; HET-P, 5-(2-hyroxylethyl)-4-methylthiazole phosphate; HMP-P, 4-amino-2-methyl-5-pyrimidine phosphate; HMP-PP, 4-amino-2-methyl-5-pyrimidine diphosphate; PL, pyridoxal; PLP, pyridoxal-5'-phosphate; PM, pyridoxamine; PMP, pyridoxamine-5'-phosphate; PN, pyridoxine; PNP, pyridoxine-5'-phosphate; TDP, thiamine diphosphate; TMP, thiamine phosphate.

Thi2 and Pdc2. Pdc2 is a transcriptional regulator that activates both *THI* genes and *PDC* genes encoding pyruvate decarboxylases [128, 134]. These regulatory proteins therefore provide an interesting link between the biosynthesis of pyruvate decarboxylase, the most highly expressed TDP-dependent enzyme in *S. cerevisiae*, and its cofactor. A regulatory link between the biosynthesis of thiamine and that of nicotinic acid, another member of the B-complex vitamins, was demonstrated when the NAD^+-dependent histone deacetylase Hst1 was found to act as a repressor of basal *THI*-gene expression [135].

Biotin (B_7)

During the first half of the 20th century, biotin was discovered as an essential growth factor for various organisms [136, 137]. Biotin plays an important role as coenzyme in carboxylases involved in fatty acid synthesis, sugar, and amino acid metabolism [138]. The cytosolic (Acc1; [139] and mitochondrial (Hfa1) acetyl-CoA carboxylases [85], pyruvate carboxylase (Pyc1,2; [140]), urea carboxylase (Dur1,2; [141]) and a tRNA-aminoacylation cofactor (Arc1; [142]) are the only biotin-dependent enzymatic activities in *S. cerevisiae* (Table 2). Covalent linkage of the carboxyl group of biotin to an ε-lysine residue of apo-Acc1 and apo-Pyc1 and 2 is catalyzed by the biotin protein ligase, Bpl1 [86, 143]. Although not characterised, a similar mechanism is likely to occur for the mitochondrial acetyl-CoA carboxylase [144, 145]. Biotin can be taken up via the proton symporter Vht1 [146]. Alternatively, the biotin intermediates 8-amino-7-oxonanote (KAPA) and 7,8-diaminopelargonate (DAPA) can be transported into yeast via the Bio5 membrane protein [34].

The molecular structure of biotin is characterized by an imidazole, or ureido ring, fused with a sulfur-containing tetrahydrothiophene ring, substituted with a valeric acid chain (Figure 2). The reactions involved in the formation of the ring structures of biotin from KAPA are highly conserved among yeast and bacteria and require three steps starting with the conversion of KAPA to DAPA. This reaction is catalysed by Bio3, a DAPA aminotransferase that requires S-adenosyl-methionine (SAM) and PLP as cofactors. The following step, catalysed by the dethiobiotin synthetase Bio4, converts DAPA to dethiobiotin at the expense of adenosine triphosphate [34]. In the final step, the biotin synthase Bio2, a mitochondrial iron-sulfur-cluster protein, converts dethiobiotin to biotin by incorporating a sulfur atom [147], presumably acting as a suicide enzyme [148].

The pathway for synthesis of the valeric side chain of biotin remains elusive and probably

involves Bio1 and Bio6, both of which are required for biotin-independent growth of *S. cerevisiae* [149]. Presence of Bio1 and Bio6 is strain dependent. For example, the reference strain S288C [150] lacks these two genes and is unable to grow on CDMY lacking biotin [151]. In contrast, sake strains of *S. cerevisiae* [152], *S. cerevisiae* strains isolated from cachaça fermentations [153], and the laboratory strains A364a [149] and CEN.PK113-7 [151, 154] do carry these two genes and exhibit growth, albeit very slowly, on CDMY without biotin. *BIO6* has been proposed to have evolved from a duplication and neo-functionalization of *BIO3*, after *BIO3* and *BIO4* had been simultaneously acquired by horizontal gene transfer, with *BIO1* similarly having evolved from duplication and neo-functionalization of the uncharacterized ORF *YJR154W* [149].

In view of its 55% amino-acid sequence similarity with *Escherichia coli* BioA, *BIO6* probably encodes an adenosylmethionine-8-amino-7-oxononanoate transaminase [152] proposed that *BIO1* encodes a coenzyme A (CoA) ligase that activates pimelic acid, a C7 dicarboxylic acid, to pimeloyl-CoA. Although such a CoA ligase (BioW) was identified in the gram-positive bacterium *Bacillus subtilis* [155], *S. cerevisiae* Bio1 protein does not show similarity to that enzyme. Additionally, biosynthesis of pimelic acid by *S. cerevisiae* has not been reported and pimelic acid feeding to a strain carrying the full biotin biosynthesis pathway was not able to stimulate growth on medium lacking biotin [156].

Laboratory evolution studies highlighted the role of the enigmatic Bio1 protein in biotin prototrophy of *S. cerevisiae*. Prolonged cultivation of the laboratory strain CEN.PK113-7D in biotin-free accelerostats yielded an evolved strain that showed the same high specific growth rate (0.36 h^{-1}) in the absence and presence of biotin. Whole-genome re-sequencing of evolved isolates revealed a massive 20-to 40-fold amplification of the physically linked *BIO1* and *BIO6* gene copies [154]. Overexpression of *BIO1*, but not *BIO6*, from a multi-copy plasmid sufficed to increase specific growth rates of the non-evolved strain on biotin-free CMDY without biotin from 0.01 h^{-1} to 0.15 h^{-1}. Despite its unknown function, these results show that *BIO1* is a key bottleneck of in *de novo* biotin synthesis in *S. cerevisiae* [154]. Strategies to generate biotin-prototrophic *S. cerevisiae* strains are likely to benefit from elucidation of the reaction catalysed by Bio1. In fact, more recently, heterologous expression of the *BIO1* gene from the biotin prototrophic yeast *Ciberlidnera fabianii* in *S. cerevisiae* was proven to be sufficient to increase the engineered strain's grow at a rate up to 0.40 h^{-1} in biotin-free CDMY [157].

The biotin biosynthetic genes *BIO5, BIO2, BIO4, BIO3, BIO6, VHT1* and *BPL1* showed a concerted upregulation during biotin starvation [18, 152]. The promoter regions of *BIO5, VHT1, BIO2* and *BPL1* contains an upstream activating element that, in the absence of biotin, is bound by the transcription factor Vhr1 which upregulates transcription. This activation ensures expression of biotin and DAPA transporters, *de novo* biotin synthesis and enzyme biotinylation [158]. The transcriptional regulation of the biotin permease

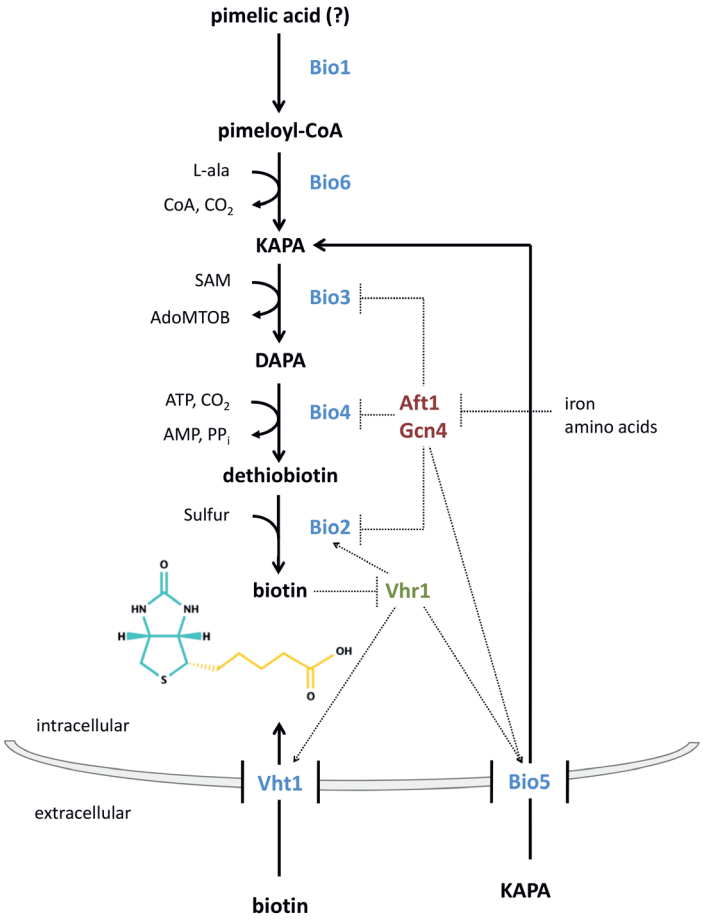


Figure 2: Biotin *de novo* biosynthesis pathway in *S. cerevisiae*. Biotin is composed of an ureido and a tetrahydrothiophene ring (shown in cyan) fused to a valeric acid chain (shown in yellow). The five final steps of *de novo* biotin synthesis are carried out by Bio1, Bio6, Bio3, Bio4 and Bio2. Origin of pimelic acid remains elusive in *S. cerevisiae* (indicated by question mark(?)). Pimeloyl-CoA formed by Bio1 is converted via 8-amino-7-oxonanoate (KAPA) to 7,8-diaminopelargonate (DAPA) by Bio6 and Bio3. DAPA is subsequently converted by Bio4 to dethiobiotin and finally to biotin by Bio2. The intermediate KAPA and biotin can be imported via the membrane transporters Bio5 and Vht1, respectively. In the absence of biotin, the regulator Vhr1 upregulates expression of genes encoding the transporters Vht1 and Bio5 as well as Bio2. In iron and amino acid rich conditions the transcriptional regulator genes *AFT1* and *GCN4* are transcriptionally repressed, which under iron and amino acid scarce conditions would not activate transcription of *BIO3*, *BIO4* and *BIO2* and relieve *BIO5* expression. Metabolites, proteins, positive and negative regulators are shown in bold, blue, green and red respectively.

gene *VHT1* is additionally controlled by the transcription factors Aft1 and Gcn4, which are involved in iron homeostasis and global control of nitrogen metabolism, respectively. Downregulation of biotin biosynthesis genes and a parallel upregulation of biotin transport upon low availability of iron and/or nitrogen [159] has been interpreted as a strategy to decrease the metabolic burden of *de novo* biotin synthesis under these conditions [18].

Vitamins that act as metabolic precursors for cofactor biosynthesis

Pantothenic acid (B_5)

Vitamin B_5 was discovered by 1933 [160] and, based on its presence in all animal tissues, named pantothenate (παντοθεν, from everywhere). Pantothenate is not a cofactor, but a key precursor for synthesis of CoA and acyl carrier protein, which play key roles in metabolism. When supplied to media, pantothenate is imported into *S. cerevisiae* by the plasma-membrane pantothenate-proton symporter Fen2 [90]. Only plants and microorganisms, including fungi, can perform *de novo* pantothenate biosynthesis. However, most sake strains of *S. cerevisiae* strains are entirely auxotrophic for pantothenate when grown in media that exclusively contain organic nitrogen sources and, in some cases, also when an inorganic nitrogen source is provided [161]. Many *S. cerevisiae* strains can synthesize pantothenic acid. In such strains, removal of the molecule from the medium typically results in impaired growth on glucose, but not on glycerol or acetate [162].

Panothenate is formed by fusion of pantoate and β-alanine, in a reaction catalysed by pantoate-beta-alanine ligase (Pan6 in *S. cerevisiae*, Figure 3). In *S. cerevisiae*, β-alanine is produced from spermine in two steps [162]. The first step is catalyzed by the polyamine oxidase Fms1, which produces 3-aminopropanal from spermine. 3-Aminopropanal is then oxidized to βalanine by the cytosolic aldehyde dehydrogenases Ald2 and Ald3. The reaction catalyzed by Fms1 has been reported to be rate limiting for pantothenate biosynthesis and Fms1 overexpression results in the secretion of pantothenic acid [163]. Pantoate is synthesized in *S. cerevisiae* from 2-*keto*-isovalerate, an intermediate of the valine biosynthesis. After conversion of 2-keto-isovalerate into 2dehydropantoate [164] by keto-pantoate hydroxymethyltransferase (Ecm31), 2dehydropantoate is transformed into pantoate by 2dehydropantoate 2-reductase (Pan5) in an NADPH-dependent reduction [165].

In comparison with the regulation of the other biosynthetic pathways discussed in this review, regulation of pantothenate acid biosynthesis in *S. cerevisiae* has not been intensively studied and, therefore, is still incompletely understood. Expression of *ECM31* and *PAN6* was shown to be low, constititutive and unaffected by extracellular panthotenate concentrations [168], while transcript levels of the pantothenic acid biosynthetic genes (*ECM31, PAN5, FMS1, ALD2, ALD3* and *PAN6*) across 55 different culture conditions

[166, 167] (Figure 4) did not reveal indications for coregulation.

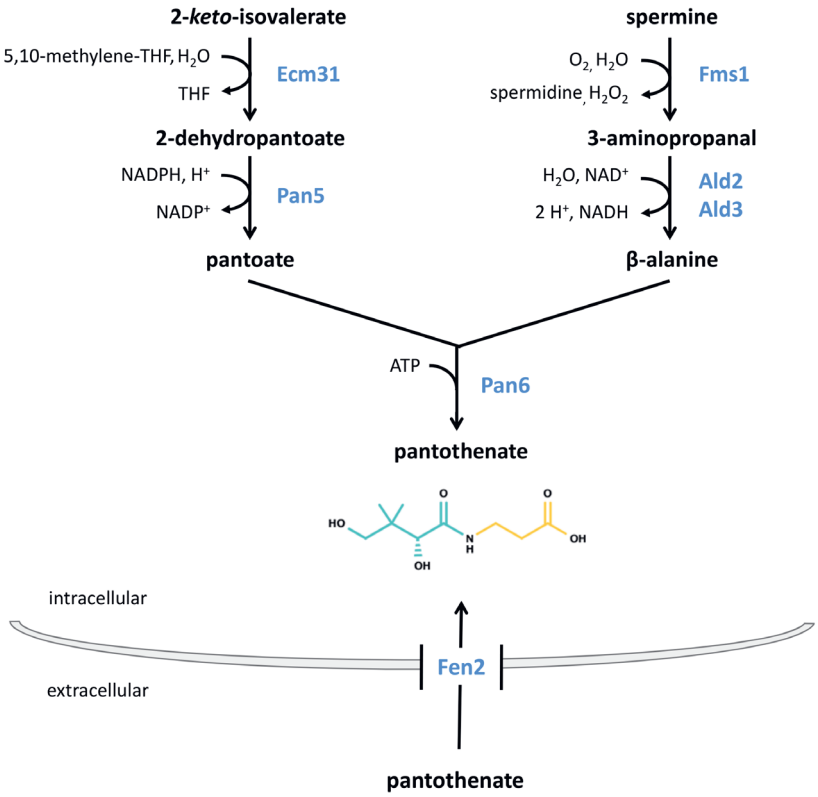


Figure 3: Pantothenate *de novo* synthesis pathway in *S. cerevisiae*. Pantothenate can be imported by the proton symporter Fen2 or synthesized *de novo* by condensation of pantoate (shown in cyan) and β-alanine (shown in yellow) in an ATP-dependent reaction catalysed by Pan6. Pantoate is formed in a two-step pathway from 2-*keto*-isovalerate catalysed by Ecm31 and Pan5 with 2-dehydropantoate as intermediate. β-alanine is formed starting from spermine by the enzymes Fms1 and Ald2-3 via 3-aminopropanal. ATP, adenosine triphosphate; $NADP^+$, nicotinamide adenine dinucleotide phosphate.

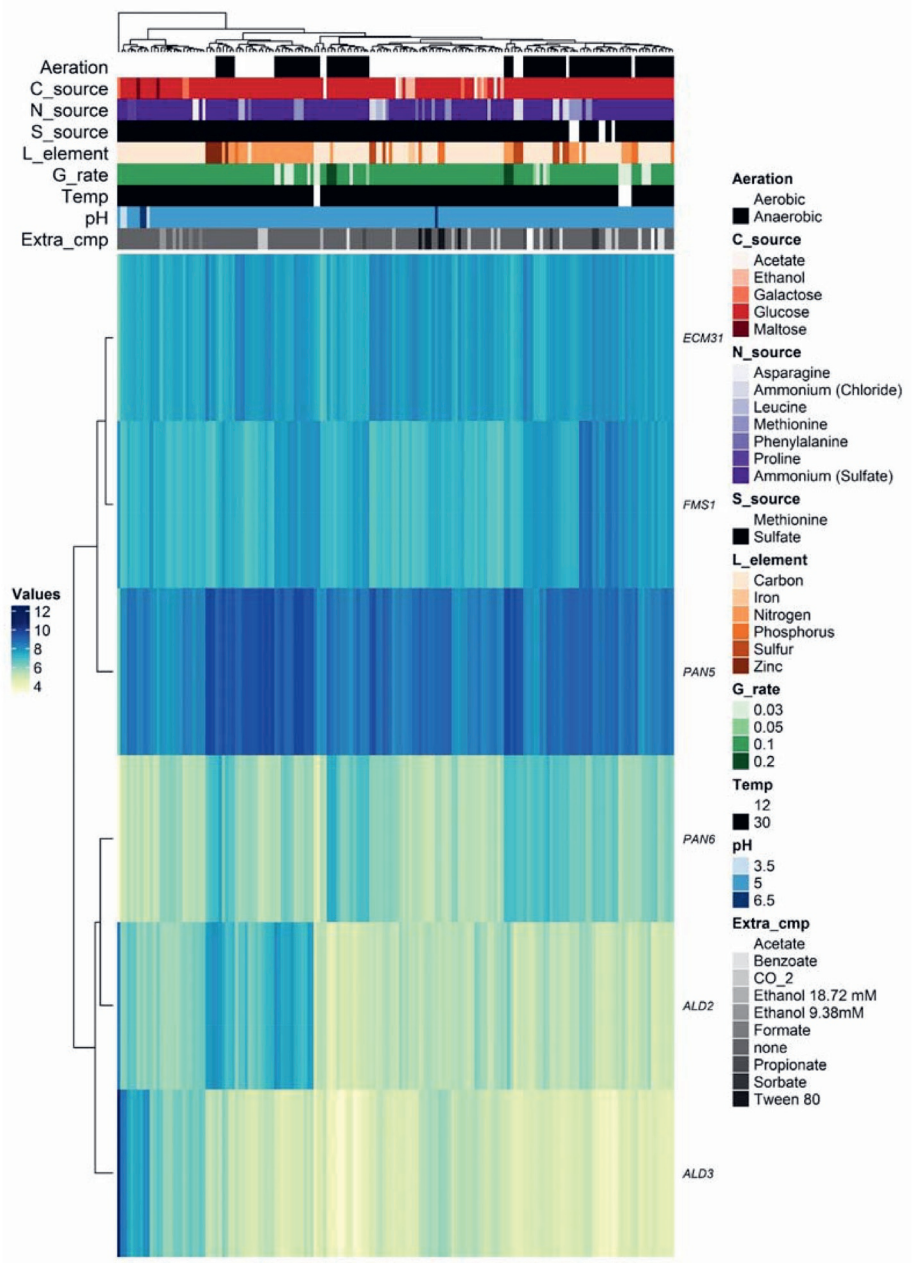


Figure 4: Heatmap showing mRNA levels for pantothenate biosynthetic genes measured under 70 different conditions in chemostat cultures. Each row shows a gene involved in *de novo* pantothenate biosynthesis while each column represents one condition. Data are derived from [166, 167] and code for generating this plot is available at https://gitlab.tudelft.nl/rortizmerino/sacch_vitamins.

*p*ABA (B_{10})

*p*ABA, also known as vitamin B_{10}, is a water-soluble B complex vitamin. It was discovered in 1920s [169, 170]. A temporary sunscreen application after the WW2 was soon withdrawn as it caused dermatitis and auto-immune responses [171, 172]. *p*ABA is an important intermediate in the biosynthesis of folates, a class of cofactors involved in transfer of C1-units in nucleic acid and amino-acid metabolisms, as well as in ubiquinone biosynthesis [173, 174]. Folates result from the association of three precursors *p*ABA (B_{10}), GTP, and glutamate, out of which *p*ABA is the less abundant intracellularly and limit folic acid (B_9) synthesis. Additionally, growth deficiency in the presence of *p*ABA and absence of folic acid has not been reported before, making this vitamin dispensable for CDMY.

In *S. cerevisiae*, *p*ABA biosynthesis starts from chorismate which, as indicated by its name (χωρίζω; to separate) is located at the intersection of the biosynthesis of (a) tyrosine and phenylalanine, (b) tryptophan, and (c) *p*ABA and folates. Conversion of chorismate into *p*ABA involves two enzyme reactions (Figure 5A). First, amino-deoxy-chorismate synthase (Abz1) uses glutamine as amino donor to produce 4amino-4-deoxy-chorismate. Subsequently, amino-deoxy-chorismate lyase (Abz2) removes the pyruvate moiety of chorismate, resulting in *p*ABA [175]. Chorismate is a key intermediate of the shikimate pathway for aromatic amino-acid biosynthesis. The shikimate pathway is tightly regulated, not only transcriptionally but also by allosteric feedback regulation of its first committed enzyme, 3-Deoxy-D-arabinoheptulosonate 7-phosphate (DAHP) synthase. *S. cerevisiae* contains two isoenzymes of DAHP, Aro3 and Aro4, which are feedback inhibited by phenylalanine and tyrosine, respectively [176, 177]. This regulation ensures that intracellular chorismate availability is strongly influenced by aromatic amino-acid concentrations. *ABZ1* and *ABZ2*, which encode the key enzymes of the *p*ABA pathway, are transcribed constitutively [178], suggesting that any regulation of *p*ABA biosynthesis occurs is post-transcriptional.

Rates of fermentation and nitrogen assimilation of *S. cerevisiae* strains have been correlated with specific alleles of *ABZ1*, thereby linking *p*ABA synthesis to overall strain performance [178, 179]. This genetic heterogeneity has been exploited to engineer *S. cerevisiae* for *p*ABA production by overexpressing *ABZ1-2* alleles from wine strains that encode highly active enzymes [180].

Nicotinic acid (B_3)

Nicotinic acid, also known as niacin, was first isolated from liver in 1937 and was identified as "pellagra-preventing factor" and "anti-black tongue factor" [181]. Together with nicotinamide, it makes up the vitamin B_3 complex. Nicotinic acid is an important precursor for the essential redox cofactors NAD^+ and nicotinamide adenine dinucleotide phosphate ($NADP^+$).

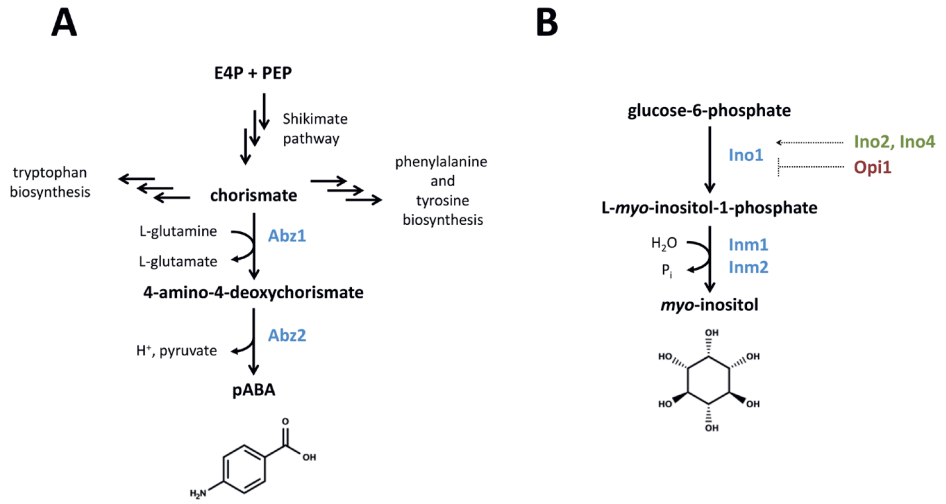


Figure 5: *p*ABA and *myo*-inositol *de novo* synthesis pathways in *S. cerevisiae*. (A) The genes *ABZ1* and *ABZ2* code for a two-step pathway producing *p*ABA from chorismate via the intermediate 4-amino-4-deoxychorismate. Chorismate is synthesized from erythrose-4-phosphate and phosphoenolpyruvate via the shikimate pathway. In addition to being precursor for *p*ABA biosynthesis, chorismate also serves as precursor for tryptophan, phenylalanine and tyrosine biosynthesis. (B) *myo*-Inositol is formed from glucose-6-phosphate via Ino1 yielding L-*myo*-inositol-1-phosphate, which is in a second step converted to *myo*-inositol by Inm1 or Inm2. *INO2, INO4* genes encode *INO1* transcriptional activators while *OPI1* encodes the antagonist reglator of the gene encoding the initial step of inositol synthesis (*INO1*). Metabolites, proteins, positive regulators and positive regulators are shown in bold, blue, green and red respectively. *p*ABA, *para*-Aminobenzoic acid.

S. cerevisiae can either obtain NAD^+ from *de novo* biosynthesis or from salvage routes that regenerate NAD^+ from its nicotinamide degradation products [182, 183] (Figure 4). These pathways converge at the level of nicotinic acid mononucleotide (NaMN) and share the last two reactions towards NAD^+ formation.

In the *de novo* biosynthesis pathway, NaMN is synthesized from L-tryptophan in a series of six enzymatic reactions (catalysed by Bna1-2 and Bna4-7) and one spontaneous reaction (Figure 5). Three of the enzymes involved in the *de novo* biosynthesis pathway, indoleamine 2,3 dioxygenase (Bna2), kynurenine 3-monooxygenase (Bna4), and 3-hydroxyanthranilate 3,4-dioxygenase (Bna1), require molecular oxygen as a substrate, thereby explaining the strict requirement of anaerobic *S. cerevisiae* cultures for nicotinic acid supplementation [54]. In the salvage pathway, nicotinamide and nicotinic acid are converted to NaMN via the so-called Preiss-Handler pathway I [184, 185] which involves Pnc1 and Npti1 as key enzymes. Extracellular nicotinic acid can be imported into yeast cells by the plasma-membrane transporter Tna1 and then used to form NAD^+ through the

salvage pathway [186, 187].

In yeast, there are other four additional pathways for NAD^+ biosynthesis: two salvage pathways from nicotinamide riboside (NR) and two salvage pathways from nicotinic acid riboside [188-190]. Three of these salvage pathways converge first with the Preiss-Handler NAD^+ salvage pathway and then with the *de novo* NAD pathway (Figure 4). In the NR salvage pathway I, which is not connected to the other pathways, NR is first phosphorylated to nicotinamide nucleotide by the Nrk1 kinase and then adenylated to NAD^+ by Nma1 or Nma2 (Figure 6).

NAD^+ and $NADP^+$ are essential redox cofactors for many oxido-reductases [191]. In addition to its role as a redox cofactor, NAD^+ is a substrate for several enzymes in yeast including sirtuin protein deacetylases (Sir2, Hst1-4) and cyclic ADP-ribose (cADPR) synthases (Tpt1; [192, 193]. These enzymes have important roles in the maintenance and regulation of chromatin structure, calcium signaling, life-span and DNA repair [183, 194-197]. NAD^+ is also a precursor for $NADP^+$ which, like NAD^+ is involved in many cellular redox reactions [198]. Intracellular NAD^+ levels are controlled by a complex regulation network. Hst1 (Homologue of Sir2) acts as a NAD^+ sensor that represses *BNA* genes when NAD^+ is abundant. Hst1 does not bind the DNA directly but interacts with Rfm1 and Sum1 to form a repressor complex. Mac1, which was previously characterized as a copper-sensing transcription factor, has been shown to also be involved in regulation of *BNA* genes, together with Hst1 [182, 199-201]. When NA is abundantly available, NA salvage metabolism is preferred over use of the *de novo* biosynthetic pathway, which is repressed by Hst1 [182, 201]. In *S. cerevisiae*, NAD^+ metabolism is regulated together with phosphate and purine nucleotide metabolism, although the exact mechanisms remain uncharacterised [202, 203]. NR can be produced and stored in vacuoles and then released into the cytosol by the Fun26 transporter, thereby enabling cells to feed NR stores into NAD^+ synthesis [204, 205].

Inositol (B_8)

Of the seven organic supplements that are added to commonly used CDMY, only inositol (Table 1) is not a cofactor or cofactor precursor. First isolated in 1928 [206], inositol is a polyol (cyclohexane-1,2,3,4,5,6-hexol) that serves as precursor for phosphatidylinositol, a main constituent of phospholipid membranes [207]. Upon its cleavage into inositol phosphate and diacylglycerol by phospholipase C, phosphatidylinositol also plays a central role in inositol-phosphate signalling [208]. Moreover, inositol is a precursor for the synthesis of glycosylphosphatidylinositol anchor proteins [209].

Myo-inositol is physiologically the most common stereoisomer among the eight possible inositol enantiomers. In organisms capable of synthesizing *myo*-inositol, it is formed from glucose-6-phosphate via two enzyme-catalysed reactions.

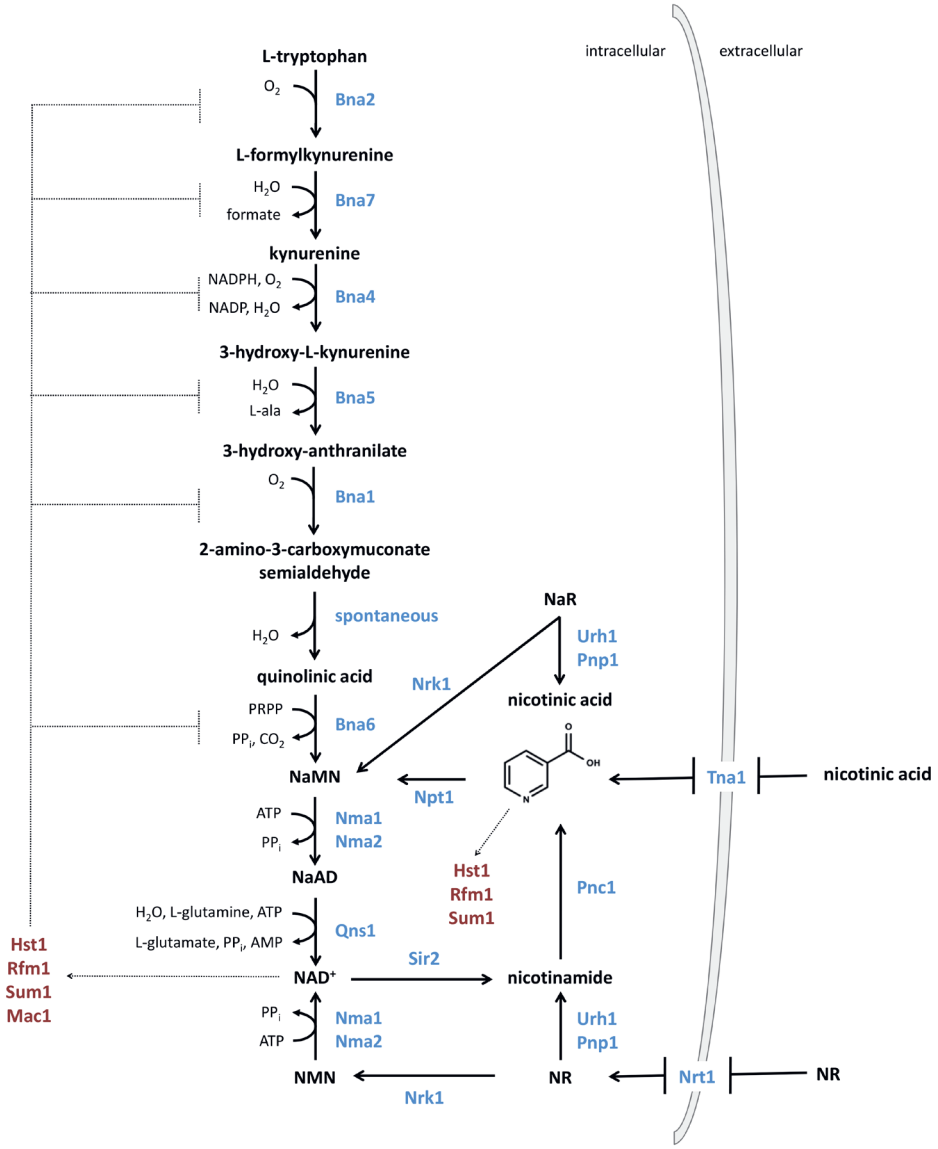


Figure 6: Nicotinic acid *de novo* synthesis and salvage pathway in *S. cerevisiae*. NAD^+ is *de novo* synthesized from L-tryptophan in nine catalytic steps involving the Bna enzyme family and enzymes Nma1, Nma2 and Qns1. Nicotinic acid can be imported into the cell via Tna1 and enters the NAD synthesis pathway as NaMN by catalytic activity of Npt1. Similarly, NaR can be salvaged by catalytic activity of Nrk1 to form NaMN. NaR can be also converted to nicotinic acid by Urh1 and Pnp1. Nrk1 also converts NR into NMN subsequently converted to NAD^+ by Nma1 and Nma2. NR is imported by activity of Nrt1 transporter and might be used by Pnp1 or Urh1 to form nicotinamide. Alternatively, nicotinamide can be synthesized via Sir2 from NAD^+. Pnc1 uses nicotinamide to form nicotinic acid. The regulators Hst1 (with aid of Rfm1 and Sum1) and Mac1 repress the

expression of genes encoding Bna enzymes upon binding to NAD^+ and nicotinic acid. Metabolites, proteins and negative regulators are shown in bold, blue and red respectively. NAD^+, nicotinamide adenine dinucleotide; NaMN, nicotinic acid mononucleotide; NAR, nicotinamide riboside; NMN, nicotinamide nucleotide; NR, nicotinamide riboside.

The genes involved in the *S. cerevisiae* inositol biosynthesis pathway were discovered by complementation of inositol-requiring mutants [210] (Figure 5B). First, L-*myo*-inositol 1-phosphate is generated from glucose-6-phosphate by L-*myo*-inositol 1phosphate synthase (Ino1) [211]. Subsequently, *myo*-inositol is generated by dephosphorylation of L-*myo*-inositol 1-phosphate by the heterodimeric enzyme inositol 3-phosphate monophosphatase (Inm1/Inm2) [212].

Lipid metabolism in eukaryotic cells, including yeasts, is rigorously regulated. Yeast cells continuously monitor lipid status and quickly respond to alterations by a dual regulatory control. Many insights into how the yeast cells regulate their phospholipid metabolism derive from research on regulatory responses to variations in the inositol content of growth media [213]. Analysis of inositol-auxotrophic *S. cerevisiae* strains enabled discovery of *INO2* and *INO4*, which encode positive transcriptional regulators for *INO1* and a large number of other genes involved in phospholipid synthesis [214]. A key negative feedback mechanism for transcriptional regulation was discovered by the characterization of mutants able to secrete inositol, a phenotype also referred to as the Opi^- phenotype [215]. The transcriptional factor Opi1 was shown to act as a negative regulator in the presence of inositol, with some and mutations in *OPI1* resulting in constitutive *INO1* expression. The Opi^- phenotype has also been in identifying other *S. cerevisiae* genes involved in phospholipid biosynthesis, transcription, protein processing, and trafficking [216].

Systematic search for components of the class B vitamin biosynthesis pathways in *Saccharomyces* species.

Although strain-to-strain differences occur, the *S. cerevisiae* pangenome harbours all necessary genetic information to synthesize inositol, biotin, thiamine, nicotinic acid, pantothenate, pyridoxine and *p*ABA. Since the work of Burkholder, McVeigh and Moyer in 1944, no systematic analysis has been performed to assess growth factor requirements of different species within the *Saccharomyces* genus. To explore this issue, we screened the genomes of the type strains of *Saccharomyces* species for annotated sequences homologous to the structural genes encoding enzymes involved in biosynthesis of class B vitamins in *S. cerevisiae* [217]. Based on this screen, the genomes of most *Saccharomyces* type species encode complete biosynthetic pathways for these compounds (Figure 7). Two notable exceptions are *S. arboricola*, which misses key genes required for biosynthesis of pyridoxine, thiamine, and biotin (*SNO2/3*, *SNZ2/3*, *THI5-13*, *BIO1*) and *S. kudriavzevii* which lacks genes involved in biosynthesis of pyridoxine, pantothenate, *p*ABA and

inositol (*SNO1, FMS1, PAN6, ABZ1/2, INO1*). Absence of *SNO2/3* in *S. paradoxus* should not compromise its pyridoxine prototrophy as its genome does harbour the main paralog *SNO1*.

Some *Saccharomyces* species show higher copy numbers for individual vitamin biosynthesis genes than *S. cerevisiae*. In particular, *S. jurei* harbours additional copies of *SNO2/3, SNZ2/3, THI5* and *THI11-13*, while *S. paradoxus* carries two copies of *BIO1* and *BIO6*. These genes are all located in subtelomeric regions in *S. cerevisiae*. Subtelomeric regions are known hotspots for genetic plasticity that contain many gene families involved in interaction between the cell and its environment [219]. Assuming conserved syntheny within the *Saccharomyces* genus, these gene amplifications may therefore reflect evolutionary adaptations to the environmental conditions these different species were exposed to.

Figure 7: Occurrence of vitamin biosynthesis genes in *Saccharomyces* species. A homology search was conducted using HMMER v3 [218] with *S. cerevisiae* S288C proteins as queries (left side row names) against a database of annotated proteins from the *Saccharomyces* species listed in the column headers. For Bio1 and Bio6, *S. cerevisiae* K7 proteins were used as queries (indicated with *) because S288C is known to lack such proteins. Available genome annotations from species in the monophyletic *Saccharomyces* clade (formely known as sensu stricto; Table 3) were used to build a protein sequence database. Besides *S. cerevisiae* S288C and CEN.PK113-7D, sequences in the database belong to type strains. This database was then searched for sequence homologs using the queries listed on the left-hand side. Queries are grouped and labeled on the right-hand side and depending on the biosynthetic pathway they are involved in. Boxes are colored depending on the number of hits (e-value > 1e-5, percentage of alignment > 75%) obtained by each query on each strain. The color code is shown at the bottom. Hits from queries belonging to the same biosynthetic pathway were ranked according to lowest e-value, then highest percentage of alignment and best hits were uniquely assigned to each query (i.e., a sequence considered as best hit is never used more than once and best hits with a count > 1 are all identical). This last step accounts for the presence of paralogs and the high level of similarity between proteins in the same pathway, especially in the pyridoxine and thiamine pathways (see Thi5 and Thi20 for instance). Code for this search is available in https://gitlab.tudelft.nl/rortizmerino/sacch_vitamins and sequences are deposited under BioProject accession PRJNA578688 as indicated in Table 3.

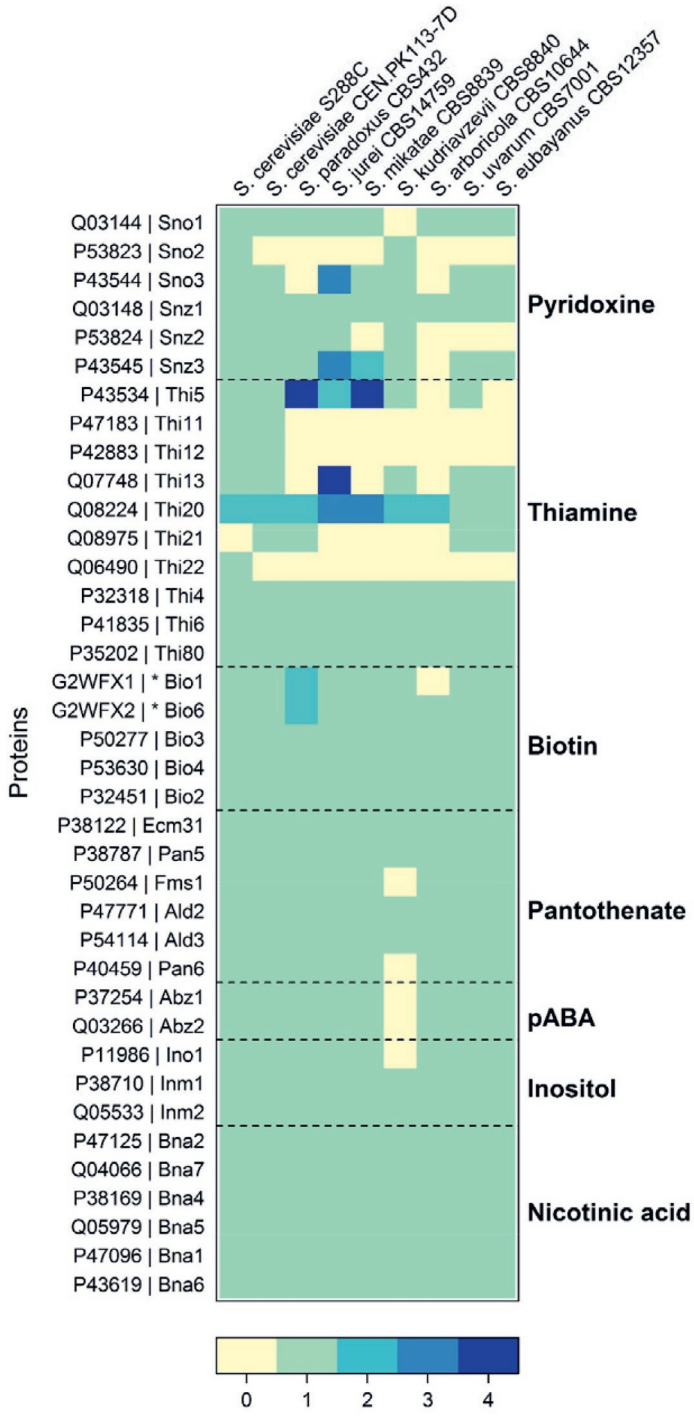

S. cerevisiae S288C
S. cerevisiae CEN.PK113-7D
S. paradoxus CBS432
S. jurei CBS14759
S. mikatae CBS8839
S. kudriavzevii CBS8840
S. arboricola CBS10644
S. uvarum CBS7001
S. eubayanus CBS12357
Proteins
Q03144 | Sno1
P53823 | Sno2
P43544 | Sno3
Q03148 | Snz1
P53824 | Snz2
P43545 | Snz3
P43534 | Thi5
P47183 | Thi11
P42883 | Thi12
Q07748 | Thi13
Q08224 | Thi20
Q08975 | Thi21
Q06490 | Thi22
P32318 | Thi4
P41835 | Thi6
P35202 | Thi80
G2WFX1 | * Bio1
G2WFX2 | * Bio6
P50277 | Bio3
P53630 | Bio4
P32451 | Bio2
P38122 | Ecm31
P38787 | Pan5
P50264 | Fms1
P47771 | Ald2
P54114 | Ald3
P40459 | Pan6
P37254 | Abz1
Q03266 | Abz2
P11986 | Ino1
P38710 | Inm1
Q05533 | Inm2
P47125 | Bna2
Q04066 | Bna7
P38169 | Bna4
Q05979 | Bna5
P47096 | Bna1
P43619 | Bna6
Pyridoxine
Thiamine
Biotin
Pantothenate
pABA
Inositol
Nicotinic acid
0
1
2
3
4

Table 3: *Saccharomyces* species involved in a comparative analysis of the presence of annotated vitamin biosynthesis genes.

TaxID	Species	Strain	Other Identifiers	Reference	Biosample
1080349	*S. eubayanus*	CBS12357^T	NBRC111513^T	[220]	NA
226127	*S. uvarum*	CBS7001^T	MCYC623^T	[221]	SAMN13069661
1160507	*S. arboricola*	CBS10644^T	H-6^T	[222]	SAMN13069660
226230	*S. kudriavzevii*	CBS8840^T	IFO1802^T	[221]	NA
226126	*S. mikatae*	CBS8839^T	IFO1815^T	[221]	SAMN13069662
1987369	*S. jurei*	CBS14759^T	NCYC3947^T	[223]	SAMN13069663
226125	*S. paradoxus*	CBS432^T	NRRLY-17217^T	[224]	SAMN13069659
559292	*S. cerevisiae*	S288C	CBS8803	[150]	NA
889517	*S. cerevisiae*	CEN.PK113-7D	CBS8340	[25]	SAMN13069664

Note. Biosamples can be accessed under bioproject accession PRJNA578688 (https://www.ncbi.nlm.nih.gov/bioproject)

Engineering of non-native cofactors in yeast

Cofactors are defined as non-protein organic molecules or inorganic compounds (mainly in the form of metal ions) that are essential for catalytic activity of an enzyme [226]. As many enzymes are dependent on the presence of one or more essential cofactors in order to be active, the introduction of a new pathway in a microbial host through metabolic engineering can impose additional nutritional or genetic requirements for cofactor supply. In particular, the successful implementation of new heterologous pathways may require the parallel broadening of the cofactor repertoire of the host cell, either by engineering *de novo* cofactor biosynthesis or by enabling import of cofactors or their precursors from the growth medium.

In nature, two types of cofactor are found: inorganic cofactors and organic cofactors (Table 4). Inorganic cofactors include metal ions such as Fe^{2+}/Fe^{3+}, Mg^{2+} and Mn^{2+} and metal-containing structures such as iron-sulfur clusters. The range of transition metal ions required for growth of a microorganism under defined growth conditions reflects their involvement in key metabolic enzymes. Expression of a new biocatalyst whose activity requires a metal ion that is not used in the native proteome often needs to be accompanied by an adaptation of the metal ion transport systems. Although native ion transport systems may exhibit promiscuous activities and enable intake of novel ion species, they often exhibit low transport affinity. The resulting requirement for supplementation of media with high concentrations of a metal ion may then lead to toxicity.

Several examples have shown that toxicity can be prevented by the expression of a high-

Table 4: Overview of the main inorganic and organic (vitamin derived and non-vitamin derived) cofactors present in nature, example of cofactor-dependent enzymes and information about the presence of a *de novo* biosynthesis pathway in *S. cerevisiae*. FeMoco, iron molybdenum cofactor; TDP, thiamine diphosphate; NAD^+, nicotinamide adenine dinucleotide; $NADP^+$, nicotinamide adenine dinucleotide phosphate; PLP, pyrodoxal-5-phosphate; CoA; THFA, tetrahydrofolic acid; FAD, flavin adenine dinucleotide; FMN, flavin mononucleotide; ATP, adenine triphosphate; SAM, S-adenosyl methionine; CTP, cytidine triphosphate; Mo-MPT, molybdenum cofactor; Mo-MCD, Mo-molybdopterin cytosine dinucleotide; Mo(bis-MGD), Mo-bis(molybdopterin guanine dinucleotide); PAPS, 3-phosphoadenosine-5-phosphosulfate; PQQ, pyrroloquinoline quinone; THB, tetrahydrobiopterin; THMPT, tetrahydromethanopterin.

Cofactor type	Cofactor	Example of dependent enzyme	*De novo* synthesized and/or used by *S. cerevisiae*	Reference
Inorganic	Cu^{2+}	Superoxide dismutase	Yes	[228]
	Fe^{2+}/Fe^{3+}	Thiamine thiazole synthase	Yes	[121]
	Mg^{2+}	Hexokinase	Yes	[229]
	Mn^{2+}	Arginase	Yes	[230]
	Ni^{2+}	Urease	Yes*	[227]
	Zn^{2+}	Alcohol dehydrogenase	Yes	[231]
	[Fe-S] [2Fe-2S] [4Fe-4S]	Biotin synthase	Yes	[232]
	FeMoco	Nitrogenase	No	[233]
Organic (vitamins and derivatives)	TDP	Pyruvate decarboxylase	Yes	[76]
	$NAD^+/NADP^+$	Glucose-6-phosphate dehydrogenase	Yes	[234]
	PLP	Cysteine desulfurase	Yes	[47]
	Cobalamine	Methionine synthase	No	[235]
	Biotin	Pyruvate carboxylase	Yes	[87]
	Coenzyme A	Fatty acid synthase	Yes	[236]
	THFA	Serine hydroxymethyltransferase	Yes	[237]
	Menaquinone	Osteocalcin	No	[238]
	Ascorbic acid	Prolyl 3-hydroxylase	Yes*	[239]
	FAD	Squalene monooxygenase	Yes	[240]
	FMN	Dihydroorotate dehydrogenase	Yes	[241]
	Coenzyme F_{420}	methenyl-H4-tetrahydromethanopterin reductase	No	[242]

	ATP	Hexokinase	Yes	[243]
	SAM	Cysteine methyltransferase	Yes	[244]
	Coenzyme B	Methyl-coenzyme M reductase	No	[245]
	Coenzyme M	Methyl-coenzyme M reductase	No	[245]
	Coenzyme Q	NADH:ubiquinone oxidoreductase	Yes	[246]
	CTP	Diacylglycerol kinase	Yes	[247]
	Glutathione	Glutaredoxin	Yes	[248]
	Heme	Catalase	Yes	[249]
Organic (non vitamins)	Lipoamide	α-ketoglutarate dehydrogenase	Yes	[250]
	Methanofuran	Formylmethanofuran dehydrogenase	No	[251]
	Moco/Mo-MPT	Nitrate reductase	Yes*	[252]
	Mo-MCD	Aldehyde oxidoreductase	No	[253]
	Mo-bis(MGD)	Formate dehydrogenase	No	[254]
	Nucleotide sugars	Glycosyltransferase	Yes	[255]
	PAPS	Sulfotransferase	Yes	[256]
	PQQ	Glucose dehydrogenase	No	[257]
	THB	Tyrosine hydroxylase	Yes*	[258]
	THMPT	Methylene-tetrahydromethanopterin Dehydrogenase	No	[259]

* engineered in *S. cerevisiae*

affinity transporter that can efficiently import a metal ion at low extracellular level. For instance, *S. cerevisiae* does not express native nickel-dependent enzymes. To functionally express a nickel-dependent, ATP-independent urease from *Schizosaccharomyces pombe*, the presence of a high-affinity Ni^{2+} transporter from *Sch. pombe* sufficed to enable growth on urea, even when no Ni^{2+} was added to growth media. This observation demonstrated the efficiency of the heterologously expressed transporter at extremely low Ni^{2+} concentrations that probably arose from leaching from glassware. In contrast, strains in which this *nic1* transporter was not expressed needed a 20 μM Ni^{2+} to support growth on urea [227]. A similar approach is followed in Chapter 5 of this thesis, in which the impact co-expression of a high-affinity molybdate (MoO_4^{2-}) transporter affects molybdenum cofactor-dependent nitrate assimilation pathway in an engineered *S. cerevisiae* strain.

Organic cofactors are structurally more complex and are often divided in compounds that in human nutrition are qualified as vitamins and other organic compounds that are essential for activity of specific (sets of) enzymes in an organism. Some of the vitamin-type cofactors have been discussed above and can natively be synthesized by *S. cerevisiae*. Ascorbic acid (vitamin C) is an example of a non-native vitamin in *S. cerevisiae* for which a heterologous vitamin synthesis pathway was engineered into baker's yeast. Although

it was not used as cofactor, intracellular synthesis of ascorbic acid yielded a strain with increased resistance to reactive oxygen species [239]. Biosynthesis of other vitamins such as cobalamine (B_{12}) and menaquinone (K_2) in *S. cerevisiae* has not yet been achieved and holds great potential for future research in food and pharmaceutical industry [260-262]. Moreover, the riboflavin-derived (B_2) coenzyme F_{420} (8-hydroxy-5-deazaflavin) that is typically found in methanogens, is considered an under-explored resource for asymmetric redox biocatalysis and engineering of an efficient production strain for this compound could be beneficial for the fields of biocatalysis and bioremediation [263, 264]. Of the non-vitamin organic cofactors, Molybdenum cofactor (Moco or Mo-MPT), which is composed by a molybdate oxyanion coordinated by a molybdopterin scaffold, is a particularly interesting target for engineering in *S. cerevisiae*. To date, over 35 different catalytic activities have been described to depend on Moco-derived cofactors and enabling their functional expression in the industrial workhorse *S. cerevisiae* could enable the use of industrially relevant enzymes such as the metal-dependent formate dehydrogenase, nitrate reductase, and molybdoprotein furoyl-CoA dehydrogenase [265].

Scope and outline of this thesis

As discussed above, the optimization of native vitamin and cofactor biosynthesis and the engineering of heterologous *non native* cofactor biosynthesis pathway are crucial steps required for increasing process robustness, reducing costs and expanding metabolic network of a microorganism. Inspired by previous research demonstrating how adaptive laboratory evolution and metabolic engineering could successfully eliminate vitamin B_7 dependency [154] and enable biosynthesis of tetrahydrobiopterin to functionally expressed opioids in *S. cerevisiae* [258], the goals of the present study were two-fold: i) investigating whether vitamin prototrophy of *S. cerevisiae* for all seven class-B vitamins could be achieved, and ii) whether *S. cerevisiae* could be engineered for synthesis of Molybdenum cofactor, a coenzyme new to *S. cerevisiae*, whose production in this yeast could potentially enable expression of new enzyme families.

In **Chapter 2**, the specific vitamin requirements of the popular *S. cerevisiae* laboratory strain CEN.PK113-7D were assessed by studying growth in CDMY lacking one vitamin at the time. By performing an adaptive laboratory evolution experiment in single-vitamin depleted media, mutants that were able to grow fast in the absence of either *p*ABA, pantothenic acid, thiamine, or pyridoxine were identified. The genomes of evolved single colony isolates were re-sequenced and non-synonymous mutations that had arisen during the evolution experiment were identified. Subsequently, a subset of mutations was selected and introduced in the unevolved background strain to study the effect of each modification on the strain growth rate in single-vitamin depleted media.

As previously described, *de novo* biosynthesis of thiamine, NAD^+, and pantothenate in

S. cerevisiae requires the presence of oxygen. As a result, these compounds have to be supplemented to CDMY when growing yeast under anaerobic conditions. Engineering yeast strains that are vitamin-independent in anaerobic laboratory cultures and industrial processes requires introduction of alternative oxygen-independent pathways for *de novo* synthesis of these compounds. **In Chapter 3**, alternative pathways for the oxygen-independent synthesis of pantothenate and NAD^+ were studied. First, the pathway responsible for oxygen independent synthesis of CoA (derived from pantothenate) in anaerobic gut fungi (Neocallimastigomycetes) was identified by genome analysis. Then, a candidate l-aspartate decarboxylase (Adc) gene was studied by phylogenetic analysis to determine its evolutionary origin. In a second step, the Adc gene, as well as the previously identified genes for the oxygen-independent synthesis of NAD^+ were introduced, together with previously characterized plant and insect-derived homologs, in *S. cerevisiae* knock-out strains that were devoid of critical steps in the native vitamin synthesis pathways. Physiological characterization of the engineered strains was performed in anaerobic lab-scale bioreactors and under strictly anaerobic conditions in an anaerobic chamber.

Chapter 4 describes research aimed at expansion of the cofactor repertoire of *S. cerevisiae* by introduction of a heterologous pathway *de novo* synthesis of Molybdenum cofactor (Moco). Genes involved in Moco biosynthesis were identified and functionally characterized in the nitrate-assimilating yeast *Ogataea parapolymorpha* by Cas9-mediated mutational analysis. The thus identified Moco biosynthesis gene-set, together with a previously characterized high-affinity molydbate importer, were introduced in *S. cerevisiae*. To couple Moco formation to growth, the strain was also equipped with a nitrate assimilation pathway from *O. parapolymorpha* including a Moco-dependent nitrate reductase. Engineered strains were tested in both in the presence or absence of oxygen and cultures were evolved in the laboratory for the ability to use nitrate as sole nitrogen source. Whole-genome sequencing on evolved single-colony isolates was performed to provide insights in the underlying evolutionary mechanism. In addition, the ability of strains to co-consume ammonium and nitrate, and their ability to assimilate nitrate at nM concentrations of molybdate were investigated. Finally, a possible industrial advantage of nitrate-assimilating *S. cerevisiae* was tested by performing co-cultivation experiments with the nitrate-assimilating spoilage yeast *B. bruxellensis*.

Only 27 % of the over 300 recently sequenced yeast species have been shown to carry Moco biosynthesis genes [266]. In **Chapter 5**, we studied whether Moco synthesis and Moco-dependent nitrate assimilation pathways from *O. parapolymorpha* could also be engineered into the evolutionarily distant and industrially relevant, lipid-accumulating *Yarrowia lipolytica*. No fewer than 11 genes, including 7 genes for Moco biosynthesis, 3 genes for nitrate assimilation, and 1 gene for high-affinity molybdate transport were introduced into *Yarrowia lipolytica* by CRISPR/Cas9-gene editing. The resulting strain

was subjected to adaptive laboratory evolution on nitrate-containing media. Subsequently, single-cell lines were isolated, phenotyped and their genome was re-sequenced to provide insights in the underlying genetic adaptations.

Chapter 2:
Adaptive laboratory evolution and reverse engineering of single-vitamin prototrophies in *Saccharomyces cerevisiae*

Thomas Perli, Dewi P.I. Moonen, Marcel van den Broek,
Jack T. Pronk, Jean-Marc Daran

Essentially as published in
Applied and Environmental Microbiology
2020; 86 (12) e00388-20
https://doi.org/10.1128/AEM.00388-20

Abstract

Quantitative physiological studies on *Saccharomyces cerevisiae* commonly use synthetic media (SM) that contain a set of water-soluble growth factors that, based on their roles in human nutrition, are referred to as B-vitamins. Previous work demonstrated that in *S. cerevisiae* CEN.PK113-7D, requirements for biotin could be eliminated by laboratory evolution. In the present study, this laboratory strain was shown to exhibit suboptimal specific growth rates when either inositol, nicotinic acid, pyridoxine, pantothenic acid, *para*-aminobenzoic acid (*p*ABA) or thiamine were omitted from SM. Subsequently, this strain was evolved in parallel serial-transfer experiments for fast aerobic growth on glucose in the absence of individual B-vitamins. In all evolution lines, specific growth rates reached at least 90 % of the growth rate observed in SM supplemented with a complete B-vitamin mixture. Fast growth was already observed after a few transfers on SM without myo-inositol, nicotinic acid or *p*ABA. Reaching similar results in SM lacking thiamine, pyridoxine or pantothenate required over 300 generations of selective growth. The genomes of evolved single-colony isolates were re-sequenced and, for each B-vitamin, a subset of non-synonymous mutations associated with fast vitamin-independent growth were selected. These mutations were introduced in a non-evolved reference strain using CRISPR/Cas9-based genome editing. For each B-vitamin, introduction of a small number of mutations sufficed to achieve a substantially increased specific growth rate in non-supplemented SM that represented at least 87% of the specific growth rate observed in fully supplemented complete SM.

Importance

Many strains of *Saccharomyces cerevisiae*, a popular platform organism in industrial biotechnology, carry the genetic information required for synthesis of biotin, thiamine, pyridoxine, *para*-aminobenzoic acid, pantothenic acid, nicotinic acid and inositol. However, omission of these B-vitamins typically leads to suboptimal growth. This study demonstrates that, for each individual B-vitamin, it is possible to achieve fast vitamin-independent growth by adaptive laboratory evolution (ALE). Identification of mutations responsible for these fast-growing phenotype by whole-genome sequencing and reverse engineering showed that, for each compound, a small number of mutations sufficed to achieve fast growth in its absence. These results form an important first step towards development of *S. cerevisiae* strains that exhibit fast growth on cheap, fully mineral media that only require complementation with a carbon source, thereby reducing costs, complexity and contamination risks in industrial yeast fermentation processes.

Introduction

Chemically defined media for cultivation of yeasts (CDMY) are essential for fundamental and applied research. In contrast to complex media, which contain non-defined components such as yeast extract and/or peptone, defined media enable generation of highly reproducible data, independent variation of the concentrations of individual nutrients and, in applied settings, design of balanced media for high-biomass-density cultivation and application of defined nutrient limitation regimes [12, 13]. CDMY such as Yeast Nitrogen Base (YNB) and Verduyn medium are widely used in research on *Saccharomyces* yeasts [10, 13]. In addition to carbon, nitrogen, phosphorous and sulfur sources and metal salts, these media contain a set of seven growth factors: biotin (B_7), nicotinic acid (B_3), inositol (B_8), pantothenic acid (B_5), *para*-aminobenzoic acid (*p*ABA) (formerly known as B_{10}), pyridoxine (B_6) and thiamine (B_1). Based on their water solubility and roles in the human diet, these compounds are all referred to as B-vitamins, but their chemical structures and cellular functions are very different [3]. Taking into account their roles in metabolism, they can be divided into three groups i) enzyme co-factors (biotin, pyridoxine, thiamine), ii) precursors for co-factor biosynthesis (nicotinic acid, *p*ABA, pantothenic acid) and iii) inositol, which is a precursor for phosphoinositol and glycosylphosphoinositol anchor proteins [267].

Previous studies demonstrated that *Saccharomyces* species are bradytroph; growth does not strictly depend on addition of all of these B-vitamins, but omission of individual compounds from CDMY typically results in reduced specific growth rates [19-21]. These observations imply that the term 'vitamin', which implies a strict nutritional requirement, is in many cases formally incorrect when referring to the role of these compounds in *S. cerevisiae* metabolism [267]. In view of its widespread use in yeast physiology, we will nevertheless use it in this paper.

The observation that *Saccharomyces* yeasts can *de novo* synthesize some or all of the 'B-vitamins' included in CDMY is consistent with the presence of structural genes encoding the enzymes required for their biosynthesis (Fig. 1, [267]). However, as illustrated by recent studies on biotin requirements of *S. cerevisiae* CEN.PK113-7D [7, 267], a full complement of biosynthetic genes is not necessarily sufficient for fast growth in the absence of an individual vitamin. In the absence of biotin, this strain grew extremely slowly ($\mu < 0.01$ h^{-1}), but fast biotin-independent growth could be obtained through prolonged adaptive laboratory evolution (ALE) in a biotin-free CDMY. Reverse engineering of mutations acquired by evolved strains showed that, along with mutations in the plasma-membrane-transporter genes *TPO1* and *PDR12*, a massive amplification of *BIO1* was crucial for fast biotin-independent growth of evolved strains [154]. These results illustrated the power of ALE in optimizing microbial strain performance without *a priori* knowledge of criticial genes or proteins and in unravelling the genetic basis for industrially relevant phenotypes

by subsequent whole-genome sequencing and reverse engineering [268, 269].
Elimination of vitamin requirements could enable cost reduction in the preparation of defined industrial media and fully prototrophic strains could provide advantages in processes based on feedstocks whose preparation requires heating and/or acid-treatment steps (e.g. lignocellulosic hydrolysates; [270, 271]) that inactivate specific vitamins. In addition, processes based on vitamin-independent yeast strains may be less susceptible to contamination by vitamin-auxotrophic microorganisms such as lactic acid bacteria) [272]. Thus, chassis strains able to grow fast in the absence of single or multiple vitamins would therefore be of interest for industrial application. Moreover, engineering strategies aimed at enabling fast growth and product formation in the absence of single or multiple vitamins may be relevant for large-scale industrial application of *Saccharomyces* yeasts.
The goals of the present study were to investigate whether full single-vitamin prototrophy of *S. cerevisiae* for inositol, nicotinic acid, pantothenic acid, *p*ABA, pyridoxine or thiamine could be achieved by ALE and to identify mutations that support fast growth in the absence of each of these vitamins. To this end, the laboratory strain *S. cerevisiae* CEN.PK113-7D

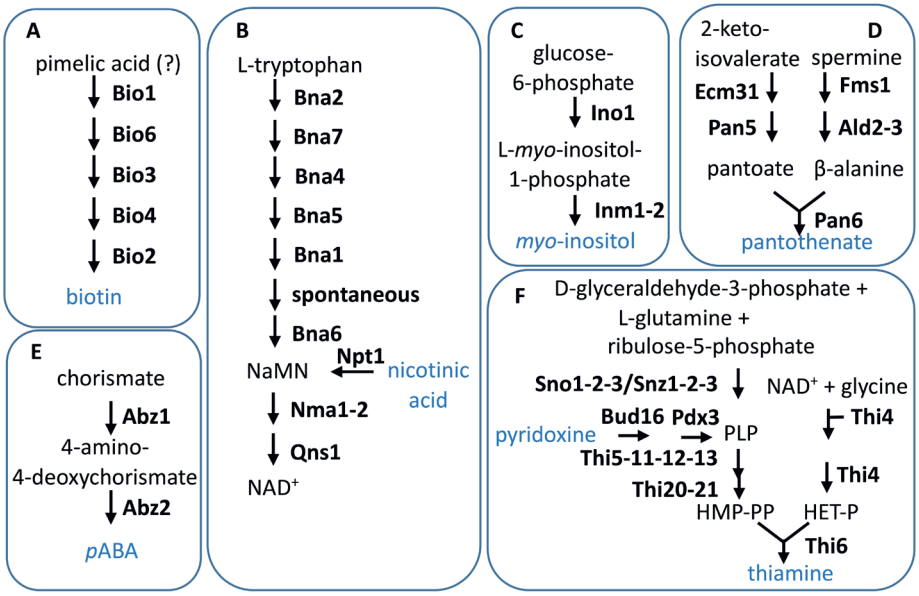


Fig. 1: Schematic representation of the *de novo* biosynthetic pathways for the B-vitamins biotin (A), nicotinic acid (B), myo-inositol (C), pantothenate (D), *p*ABA (E), pyridoxine and thiamine (F) in *S. cerevisiae* [267]. Vitamins that are usually added to the chemical defined media for cultivation of yeasts are shown in blue.

was subjected to parallel aerobic ALE experiments that encompassed serial transfer in different synthetic media, which each lacked a single B-vitamin. Independently evolved strains from each medium type were then characterized by whole-genome resequencing and the relevance of selected identified mutations was assessed by their reverse engineering in the parental non-evolved strain.

Results

Assessment of CEN.PK113-7D specific B-vitamin requirements

S. cerevisiae strains belonging to the CEN.PK lineage, which was developed in an interdisciplinary project supported by the German Volkswagen Stiftung between 1993 and 1994 [273], exhibit properties that make them good laboratory models for yeast biotechnology [274]. To provide a baseline for ALE experiments, specific growth rates of the haploid strain CEN.PK113-7D were analysed in aerobic batch cultures on complete SMD and on seven 'SMDΔ' media from which either biotin, inositol, nicotinic acid, pantothenic acid, *p*ABA, pyridoxine or thiamine was omitted. To limit interference by carry-over of vitamins from precultures, specific growth rates were measured after a third consecutive transfer on each medium (Fig. 2A).

Consistent with the presence in its genome of genes predicted to encode all enzymes involved in biosynthetic pathways for all seven vitamins (Fig. 1, [267]), strain CEN.PK113-7D grew on all SMDΔ versions. On complete SMD, a specific growth rate of 0.38 ± 0.02 h^{-1} was observed, while specific growth rates on SMDΔ lacking biotin, pantothenate, pyridoxine, thiamine or inositol were 95%, 57%, 32%, 22% and 19% lower, respectively. After three transfers, specific growth rates on SMDΔ lacking *p*ABA or nicotinic acid did not differ significantly from the specific growth rate on complete SMD (Fig. 2A). However, in SMDΔ lacking *p*ABA, growth in the first transfer was slower than in the first transfer on complete SMD (Fig. 2B). Extending the number of transfers to five, which corresponded to approximately 33 generations of selective growth, led to higher specific growth rates on several SMDΔ versions (Fig. 2A), suggesting that serial transfer selected for spontaneous faster-growing mutants.

Adaptive laboratory evolution of CEN.PK113-7D for fast growth in the absence of single vitamins.

Serial transfer in independent triplicate aerobic shake-flask cultures on each SMDΔ version was used to select mutants that grew fast in the absence of individual vitamins. Specific growth rates of evolving populations were measured after 5, 10, 23, 38 and 50 transfers and compared to the specific growth rate of strain CEN.PK113-7D grown in complete SMD.

ALE experiments were stopped once the population reached a specific growth rate equal

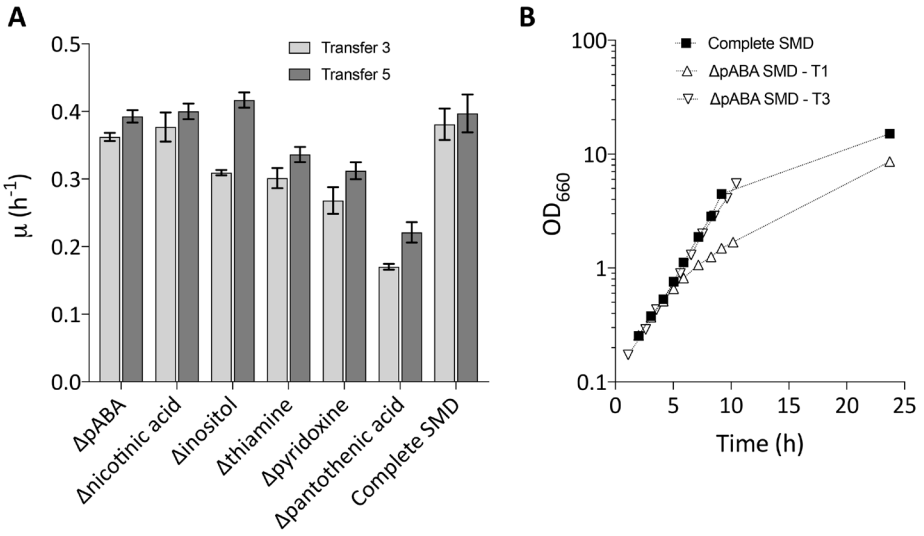


Fig. 2: Specific growth rates of *S. cerevisiae* CEN.PK113-7D in aerobic batch cultures on complete SMD and on SMD lacking single vitamins. Growth rate measurements were performed after 3 (light grey) and 5 (dark grey) consecutive transfers in the same medium. (A). Growth curve of CEN.PK113-7D in complete SMD (black square) and at transfer 1 (white upward triangle) and 3 (white downward triangle) in SMD lacking *para*-aminobenzoic acid (*p*ABA). In the latter medium a lower specific growth rate was observed at transfer 1 but upon the third transfer, the growth rate was the same as in complete SMD (B). Error bars represent the standard deviation (n = 9 for complete SMD. n = 3 for all other media).

to or higher than 0.35 h^{-1}, which represents 90-95% of the specific growth rate of strain CEN.PK113-7D on complete SMD (Fig. 2A) [275-278]. As already indicated by the specific growth rates observed after 3 and 5 transfers in SMDΔ (Fig. 2A), few transfers were required for reaching this target in SMDΔ lacking inositol, nicotinic acid or *p*ABA. Conversely, over 330 generations of selective growth were required to reach a specific growth rate of 0.35 h^{-1} on SMDΔ lacking either pantothenic acid, pyridoxine or thiamine (Fig. 3A). At least two single-cell lines were isolated from each of the three independent ALE experiments on each SMDΔ version and the fastest growing single-cell line from each experiment was selected (strains IMS0724-6 from SMDΔ lacking nicotinic acid; IMS0727-9 from SMDΔ lacking *p*ABA; IMS0730-2 from SMDΔ lacking inositol; IMS0733-5 from SMDΔ lacking pantothenate; IMS0736-8 from SMDΔ lacking pyridoxine and IMS0747-9 from SMDΔ lacking thiamine). The specific growth rates of isolates that had been independently evolved in each SMDΔ version did not differ by more than 6% (Table 1). The largest difference (5.3%) was observed for isolates IMS0733-5 evolved on SMDΔ lacking pantothenate (Fig. 3B).

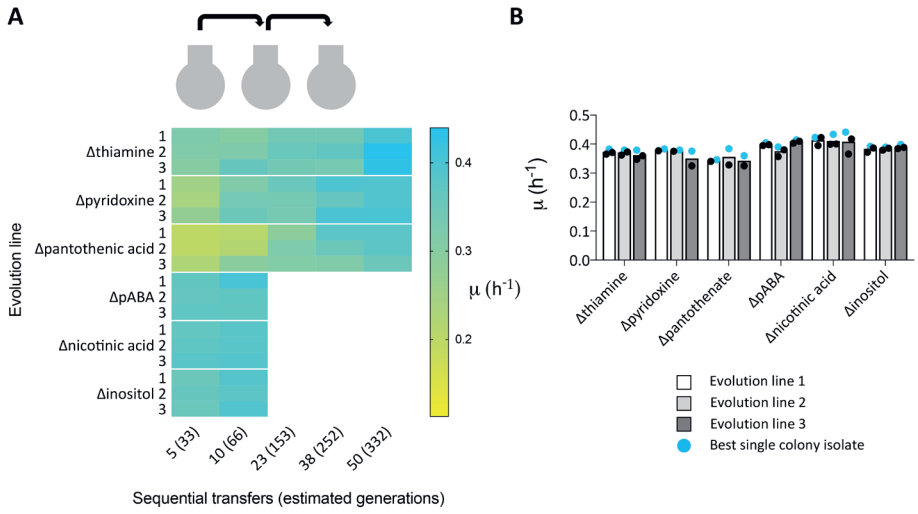


Fig. 3: Heat-map showing specific growth rates during ALE of *S. cerevisiae* CEN.PK113-7D on SMD lacking single vitamins. Aerobic serial-transfer experiments on each medium composition were performed in triplicate (rows). The specific growth rate of each evolving population was measured after a specific number of sequential transfers (columns). Yellow colour indicates slow growth while cyan indicates a specific growth rate statistically undistinguishable from the positive control (strain CEN.PK113-7D grown on SMD medium with all vitamins) (A). Specific growth rates of single colony isolates from each independent biological replicate evolution line. The fastest-growing isolates, whose genomes were resequenced, are indicated in blue (B).

Whole-genome sequencing of evolved strains and targets identification

To identify mutations contributing to vitamin independence, the genomes of the sets of three independently evolved isolates for each SMDΔ version were sequenced with Illumina short-read technology. After aligning reads to the reference CEN.PK113-7D genome sequence [225], mapped data were analysed for the presence of copy number variations (CNV) and single nucleotide variations (SNVs) that occurred in annotated coding sequences.

A segmental amplification of 34 kb (from nucleotide 802500 to 837000) on chromosome VII, which harbours *THI4*, was observed in strain IMS0749 (Fig. 4A) which had been evolved in SMDΔ lacking thiamine. *THI4* encodes a thiazole synthase, a suicide enzyme that can only perform a single catalytic turnover [121]. Segmental amplifications on chromosomes III and VIII were observed in strain IMS0725, which had been evolved in SMDΔ lacking nicotinic acid (Fig. 4B.). Since these regions are known to be prone to recombination in the parental strain CEN.PK113-7D [225, 279], their amplification is not necessarily related to nicotinic acid independence.

Table 1: Specific growth rates of best performing single colony isolates obtained from serial-transfer evolution experiments with *S. cerevisiae* CEN.PK113-7D on SMD and on SMD variants lacking individual B-vitamins. Percentage improvement over the specific growth rate of the parental strain after three transfers in the same medium is also shown (n=1 for each strain).

Strain ID	Evolution condition	Evolution replicate	Growth rate (h^{-1})	% improvement
IMS0721	Complete SMD	1	0.443	17
IMS0722	Complete SMD	2	0.423	11
IMS0723	Complete SMD	3	0.419	10
IMS0747	No thiamine	1	0.383	35
IMS0748	No thiamine	2	0.379	30
IMS0749	No thiamine	3	0.379	38
IMS0736	No pyridoxine	1	0.383	45
IMS0737	No pyridoxine	2	0.379	44
IMS0738	No pyridoxine	3	0.376	48
IMS0733	No pantothenate	1	0.346	149
IMS0734	No pantothenate	2	0.384	155
IMS0735	No pantothenate	3	0.359	159
IMS0724	No nicotinic acid	1	0.423	4
IMS0725	No nicotinic acid	2	0.434	2
IMS0726	No nicotinic acid	3	0.441	2
IMS0730	No inositol	1	0.392	12
IMS0731	No inositol	2	0.389	24
IMS0732	No inositol	3	0.399	16
IMS0727	No *p*ABA	1	0.405	6
IMS0728	No *p*ABA	2	0.389	5
IMS0729	No *p*ABA	3	0.414	4

SNV analysis was systematically performed and data from the three sequenced isolates were compared. To eliminate false positives caused by mapping artifacts, reads of the CEN.PK113-7D strains were mapped back on its own reference assembly. Identified SNVs found were systematically subtracted. SNV analysis was restricted to non-synonymous mutations in coding sequences (Table 2).

In three out of the six isolates from ALE experiments in SMDΔ lacking nicotinic acid or inositol, no non-synonymous SNVs were detected (Fig. 5). One strain (IMS0724) from a serial transfer experiment on SMDΔ lacking nicotinic acid showed SNVs in *RPG1* and *PMR1*, while a second strain (IMS0725) showed SNVs in *MTO1* and *VTH2*. A mutation in YFR054W was identified in a single strain (IMS0730) evolved for inositol-independent

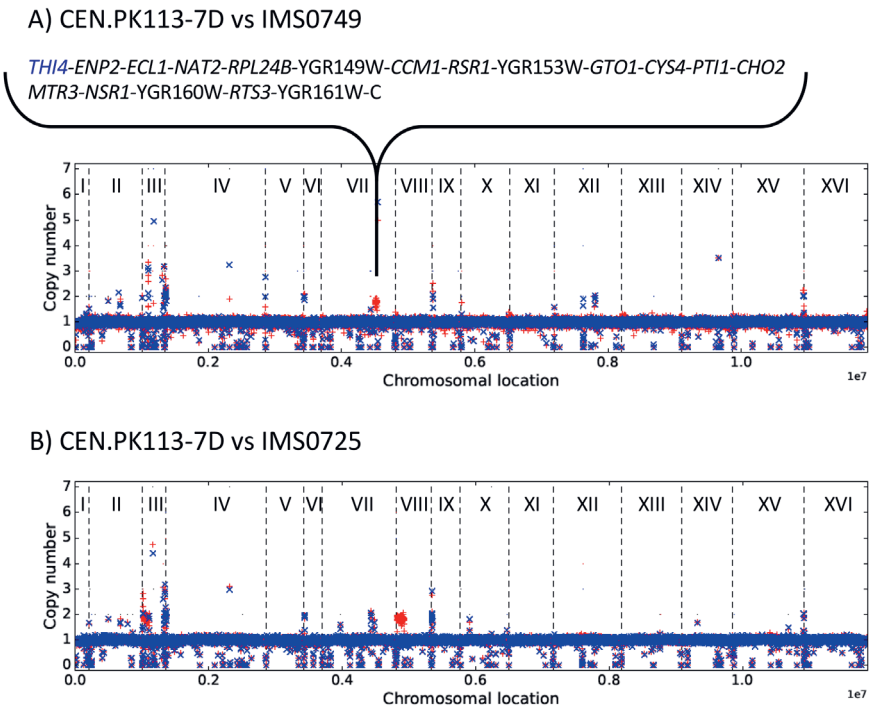


Fig. 4: Read coverages across the chromosomes of evolved isolates IMS0725 evolved for nicotinic acid prototrophy (A) and IMS0749 evolved for thiamine prototrophy (B) (in red) compared to read coverage across the chromosomes of CEN.PK113-7D (in blue). Annotated genes found in the amplified region of IMS0749 are indicated.

growth. The absence of mutations in several strains subjected to serial transfer in SMDΔ lacking nicotinic acid or inositol, is consistent with the fast growth of the parental strain CEN.PK113-7D in these media (Fig. 2A).

Sequencing of the three isolates evolved in SMDΔ lacking *p*ABA revealed only five SNVs, of which two were in *ABZ1* (strains IMS0727 and 0729) and one in *ARO7* (IMS0728), while SNVs in *NUP57* and *HTS2* were found in strains IMS0728 and IMS0729 respectively (Fig. 5). *NUP57* and *HTS2* could not be directly linked to *p*ABA metabolism. Conversely, Abz1 is an aminodeoxychorismate synthase that directs chorismate towards *p*ABA synthesis and Aro7 is a chorismate mutase that catalyses the first committed reaction towards phenylalanine and tyrosine and thereby diverts chorismate from *p*ABA synthesis (Fig.1)[280, 281]. These two SNVs therefore represented clear targets for reverse engineering.

In line with the much longer ALE experiments (approximately 332 generations), strains

Table 2: Non-conservative mutations found in single colony isolates obtained from serial transfer evolution experiments with *S. cerevisiae* CEN.PK113-7D on SMD variants lacking individual B-vitamins. Mutations that were chosen for subsequent reverse engineering experiments are shown in blue. *S. cerevisiae* strains IMS0731 and IMS0726 evolved for fast myo-inositol and nicotinic acid-independent growth, respectively, did not reveal non-conservative mutations and were not included in the table.

Gene mutated	Codon change	Amino acid change	Gene annotation
Panthotenate			
IMS0733			
AMN1	agC-agG	S67R	Antagonist of mitotic exit network protein 1
DAN4	aTc-aCc	I353T	Cell wall protein, Delayed ANaerobic 4
ERG3	Gct-Cct	A145P	Delta(7)-sterol 5(6)-desaturase, ERGosterol biosynthesis 3
ERR3	ttG-ttT	L344F	Enolase-related protein 3
ISW2	tCa-tGa	S181Stop	ISWI chromatin-remodeling complex ATPase, Imitation SWitch subfamily 2
IMS0734			
CDC15	Aca-Gca	T262A	Cell division control protein 15
RPS14A	Cca-Tca	P94S	40S ribosomal protein S14-A
TUP1	gTg-gCg	V374A	General transcriptional corepressor
RRT6	gCg-gTg	A267V	Regulator of rDNA transcription protein 6
CEG1	gCa-gTa	A4V	mRNA-capping enzyme subunit alpha
SCY1	Cct-Tct	P42S	Protein kinase-like protein SCY1
PDX1	gCa-gTa	A208V	Pyruvate dehydrogenase complex protein X component
TRM5	gCg-gTg	A106V	tRNA (guanine(37)-N1)-methyltransferase
GEF1	aGa-aTa	R637I	Anion/proton exchange transporter, Glycerol Ethanol, Ferric requiring 1
LIP2	gGc-gAc	G235D	Octanoyltransferase
HFA1	Aag-Gag	K1021E	Acetyl-CoA carboxylase, mitochondrial
UBP8	aGt-aCt	S149T	Ubiquitin carboxyl-terminal hydrolase 8
MGS1	Cca-Aca	P392T	DNA-dependent ATPase
CPT1	Gtg-Atg	V255M	Cholinephosphotransferase 1
SPE2	Gca-Aca	A278T	S-adenosylmethionine decarboxylase proenzyme
GAL11	aTt-aAt	I541N	Mediator of RNA polymerase II transcription subunit 15
CUE5	Cca-Tca	P377S	Ubiquitin-binding protein
MIP1	Gca-Aca	A630T	DNA polymerase gamma

POC4	aGc-aTc	S7I	Proteasome chaperone 4
KAP120	tTg-tCg	L582S	Importin beta-like protein
KAP120	gAc-gGc	D850G	Importin beta-like protein
SEC16	Gca-Aca	A1015T	COPII coat assembly protein
IMS0735			
TUP1	Cag-Tag	Q99Stop	General transcriptional corepressor
FMS1	Caa-Aaa	Q-33K	Polyamine oxidase
GAL11	Caa-Taa	Q383Stop	Mediator of RNA polymerase II transcription subunit 15
Pyridoxine			
IMS0736			
BAS1	cAa-cGa	Q152R	Myb-like DNA-binding protein, Basal 1
ERG5	Aga-Gga	R529G	C-22 sterol desaturase, ERGosterol biosynthesis 5
IMS0737			
BAS1	Gat-Aat	D101N	Myb-like DNA-binding protein, Basal 1
ERG5	Ggt-Tgt	G472C	C-22 sterol desaturase, ERGosterol biosynthesis 5
IMS0738			
GIP4	Tcc-Ccc	S464P	GLC7-interacting protein 4
AOS1	Gtg-Atg	V286M	DNA damage tolerance protein RHC31
ORC4	aGt-aAt	S160N	Origin recognition complex subunit 4
MSB1	Att-Ttt	I180F	Morphogenesis-related protein, Multicopy Suppression of a Budding defect 1
GCR2	Gga-Aga	G5R	Glycolytic genes transcriptional activator, GlyColysis Regulation 2
VNX1	aCa-aTa	T490I	Low affinity vacuolar monovalent cation/H(+) antiporter
MMT1	gCt-gAt	A175D	Mitochondrial Metal Transporter 1
ISF1	Tat-Gat	Y220D	Increasing Suppression Factor 1
RPM2	Gcc-Acc	A1020T	Ribonuclease P protein component, mitochondrial
BAS1	Tca-Cca	S41P	Myb-like DNA-binding protein, Basal 1
AAD14	agC-agA	S322R	Putative Aryl-Alcohol Dehydrogenase AAD14
FAS1	gaA-gaT	E1829D	Fatty Acid Synthase subunit beta
BEM2	Aac-Cac	N792H	GTPase-activating protein, Bud Emergence 2/IPL2
APL1	gGt-gTt	G6V	AP-2 complex subunit beta
DPB11	agG-agT	R699S	DNA replication regulator, DNA Polymerase B (II) 11
LSB6	Aca-Gca	T458A	Phosphatidylinositol 4-kinase, Las Seventeen Binding protein 6
EFG1	aAa-aGa	K188R	rRNA-processing protein, Exit From G1 1

CCH1	atG-atA	M828I	Calcium-Channel protein 1
RNR4	Gca-Tca	A210S	Ribonucleoside-diphosphate reductase small chain 2
GCD2	tTa-tCa	L472S	Translation initiation factor eIF-2B subunit delta
YHR219W	aAt-aGt	N61S	Putative uncharacterized protein YHR219W
CDC37	Gcc-Tcc	A275S	Hsp90 co-chaperone, Cell Division Cycle 37
SRP101	Gca-Aca	A75T	Signal recognition particle receptor subunit alpha homolog
ADE8	Gca-Aca	A142T	Phosphoribosylglycinamide formyltransferase
AIM9	gCa-gTa	A23V	Altered inheritance of mitochondria protein 9, mitochondrial
UTP20	tAt-tGt	Y1492C	U3 small nucleolar RNA-associated protein 20
RIF1	aGc-aTc	S1516I	Telomere length regulator protein, RAP1-Interacting Factor 1
PHO87	Gtc-Atc	V482I	Inorganic phosphate transporter
MAK21	tTg-tCg	L413S	Ribosome biogenesis protein, MAintenance of Killer 21
YDL176W	tCa-tAa	S186-Stop	Uncharacterized protein YDL176W
Thiamine			
IMS0747			
MAL12	Gtt-Ctt	V305L	Alpha-glucosidase, MALtose fermentation 12
CNB1	ttA-ttT	L82F	Calcineurin subunit B
PRP16	aAa-aGa	K112R	Pre-mRNA-splicing factor ATP-dependent RNA helicase
ERR3	ttG-ttT	L447F	Enolase-related protein 3
IMS0748			
PMR1	tCc-tAc	S104Y	Calcium-transporting ATPase 1
FRE2	aCt-aGt	T110S	Ferric/cupric reductase transmembrane component 2
IMS0749			
YEL074W	cAc-cCc	H66P	Putative UPF0320 protein YEL074W
CNB1	ttA-ttC	L82F	Calcineurin subunit B
MSC1	Gtt-Att	V309I	Meiotic sister chromatid recombination protein 1
ERR3	ttG-ttT	L447F	Enolase-related protein 3
***p*ABA**			
IMS0727			
ABZ1	cGt-cAt	R593H	Aminodeoxychorismate synthase
IMS0728			
ARO7	tTa-tCa	L205S	Chorismate mutase

NUP57	tCc-tTc	S396F	Nucleoporin 57
IMS0729			
ABZ1	cGt-cAt	R593H	Aminodeoxychorismate synthase
HST2	ttG-ttT	L102F	NAD-dependent protein deacetylase, Homolog of SIR Two 2
Inositol			
IMS0730			
YFR045W	Gcc-Acc	A65T	Putative mitochondrial transport protein
IMS0732			
YFR045W	Gcc-Acc	A65T	Uncharacterized mitochondrial carrier
Nicotinic acid			
IMS0724			
RPG1	Ggt-Tgt	G294C	Eukaryotic translation initiation factor 3 subunit A
PMR1	Ggt-Agt	G694S	Calcium-transporting ATPase 1
IMS0725			
MTO1	atG-atT	M356I	Mitochondrial translation optimization protein 1
VTH2	Cca-Tca	P708S	Putative membrane glycoprotein, VPS10 homolog 2
VTH2	gTT-gCC	V478A	Putative membrane glycoprotein, VPS10 homolog 2
VTH2	TtT-GtG	F477V	Putative membrane glycoprotein, VPS10 homolog 2

evolved in SMDΔ lacking thiamine, pantothenate or pyridoxine showed larger numbers of SNVs, with a maximum number of 30 SNVs in the isolate IMS0738 evolved SMDΔ lacking pyridoxine (Table 2 and Fig. 5).

Evolution on SMDΔ lacking thiamine did not yield mutations that affected the same gene in all three independently evolved isolates. However, strains IMS0747 and IMS0749 shared SNVs in *CNB1* and *ERR3*. A third isolate, strain IMS0748, contained two SNVs in *PMR1* and *FRE2*. *CNB1*, *PMR1* and *FRE2* all encode proteins that have been implicated in divalent cation homeostasis [282-286].

Isolates IMS0736 and IMS0737, which had been evolved in SMDΔ lacking pyridoxine harboured only two and three mutations, respectively, while strain IMS0738 harboured 30 mutations. All three strains carried different mutated alleles of *BAS1*, which encodes a transcription factor involved in regulation of histidine and purine biosynthesis [99, 100]. IMS0736 harboured a non- synonymous mutation causing an amino acid change position 152 (Q152R), while SNVs in strains IMS0737 and IMS0738 affected amino acids 101 (D101N) and 41 (S41P), respectively. Based on these results, *BAS1* was identified as priority target for reverse engineering.

Isolates IMS0733 and IMS0735, evolved on SMDΔ lacking pantothenic acid, carried three

and five SNVs, respectively, while isolate IMS0734 carried 21 mutations. Isolates IMS0734 and IMS0735 both carried mutations in *TUP1* and *GAL11*, resulting in different single-amino acid changes (Tup1^{V374A} Gal11^{I541N} and Tup1^{Q99stop} Gal11^{Q383stop}, respectively). *TUP1* codes for a general transcriptional corepressor [287] while *GAL11* codes for a subunit of the tail of the mediator complex that regulates activity of RNA polymerase II [288]. One of the mutations in strain IMS0733 affected *ISW2*, which encodes a subunit of the chromatin remodeling complex [289]. These three genes involved in regulatory processes were selected for reverse engineering, along with *SPE2* and *FMS1*. The latter two genes, encoding S-adenosylmethionine decarboxylase [290] and polyamine oxidase [291], are directly involved in pantothenate biosynthesis and were found to be mutated in isolates IMS0734 and IMS0735, respectively.

In summary, based on mutations in the same gene in independently evolved isolates and/or existing information on involvement of affected genes in vitamin biosynthesis, mutations in twelve genes were selected for reconstruction in the parental strain CEN.PK113-7D. These were mutated alleles of *ISW2*, *GAL11*, *TUP1*, *FMS1* and *SPE2* for panthotenate, in *BAS1* for pyridoxine, mutations in *CNB1, PMR1* and *FRE2* as well as overexpression

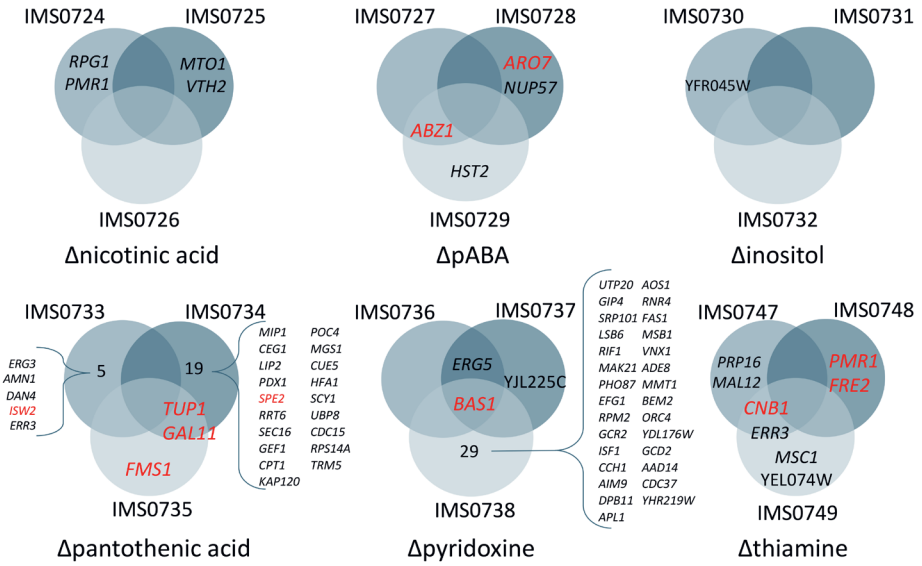


Fig. 5: Venn diagrams showing non-synonymous mutations found in coding regions of isolated strains from the different evolution experiments. Each evolution experiment was performed in triplicate. The Venn diagrams show genes that acquired nonsynonymous mutations in multiple independent evolution experiments for a specific medium as well as genes that were affected in a single replicate. Apparent mutations also found in the genome of the parent strain CEN.PK113-7D were subtracted and not shown. Target genes that were selected for reverse engineering are shown in red.

of *THI4* for thiamine and mutations in *ABZ1* and *ARO7* for *p*ABA. Since serial transfer on SMDΔ lacking nicotinic acid or inositol did not consistently yield mutations and the parental strain CEN.PK113-7D already grew fast on these media, no reverse engineering of observed mutations was observed in isolates from those experiments.

Reverse engineering of target genes mutations and overexpression

To investigate whether the selected targets contributed to the phenotypes of the evolved strains, single point mutations or single-gene overexpression cassettes were introduced in a non-evolved reference strain, followed by analysis of specific growth rate in the relevant SMDΔ variant. For most target genes, a two-step strategy was adopted, so that a single-gene knock-out mutant was constructed in the process (Fig. 6AB). For the *SPE2* mutant strains IMX2308 and IMX2289, point mutations were introduced in a single step (Fig. 6C). The *THI4*-overexpressing strains IMX2290 and IMX2291 were constructed by integrating the overexpression cassette at the YPRcTau3 locus [292] (Fig. 6D). Subsequently, multiple mutations that were found in strains evolved in the same SMDΔ version were combined into single engineered strains to test for additive or synergistic effects.

Thiamine

The specific growth rate of *S. cerevisiae* CEN.PK113-7D was only 27% lower in SMDΔ lacking thiamin than in complete SMD (Fig. 2A and 6A). Nevertheless, it took over 300 generations of selective growth to obtain evolved isolates that compensated for this difference (Fig.3 and 6A). The role of mutations in *CNB1*, *FRE2* and *PMR1* in this evolved phenotype was first investigated in the single knock-out strains IMX1721, IMX1722 and IMX1723, respectively. While deletion of *PMR1* deletion negatively affected specific growth rate on SMDΔ lacking thiamine, deletion of either *CNB1* or *FRE2* resulted in a 17% increase of the specific growth rate on this medium relative to CEN.PK113-7D. However, strains IMX1721 (*cnb1*Δ) and IMX1722 (*fre2*Δ) still grew significantly slower than the evolved isolates (Fig.6). Subsequently, the mutated alleles found in the evolved isolates were introduced at the native chromosomal locus, resulting in strains IMX1985 (*CNB1*$^{L82F)}$), IMX1986 (*PMR1*^{S104Y}) and IMX1987 (*FRE2*^{T110S}). In addition, *THI4* was overexpressed (strain IMX2290) to simulate the copy number increase observed in IMS0749. Strains IMX1987 (*FRE2*^{T110S}) and IMX2290 (*THI4*↑) grew as fast as the evolved isolates on SMDΔ lacking thiamine (0.35-0.36 h^{-1}; Fig. 5A). Combination of these mutated alleles of *PMR1* and *FRE2*, which occurred together in isolate IMS0748, as well as of the two mutations resulting in growth improvement (*FRE2*^{T110S} and *THI4*↑) was also tested. None of these combinations yielded a higher specific growth rate than observed in the evolved strains and in the reversed engineered *FRE2*^{T110S} and *THI4*↑ strains.

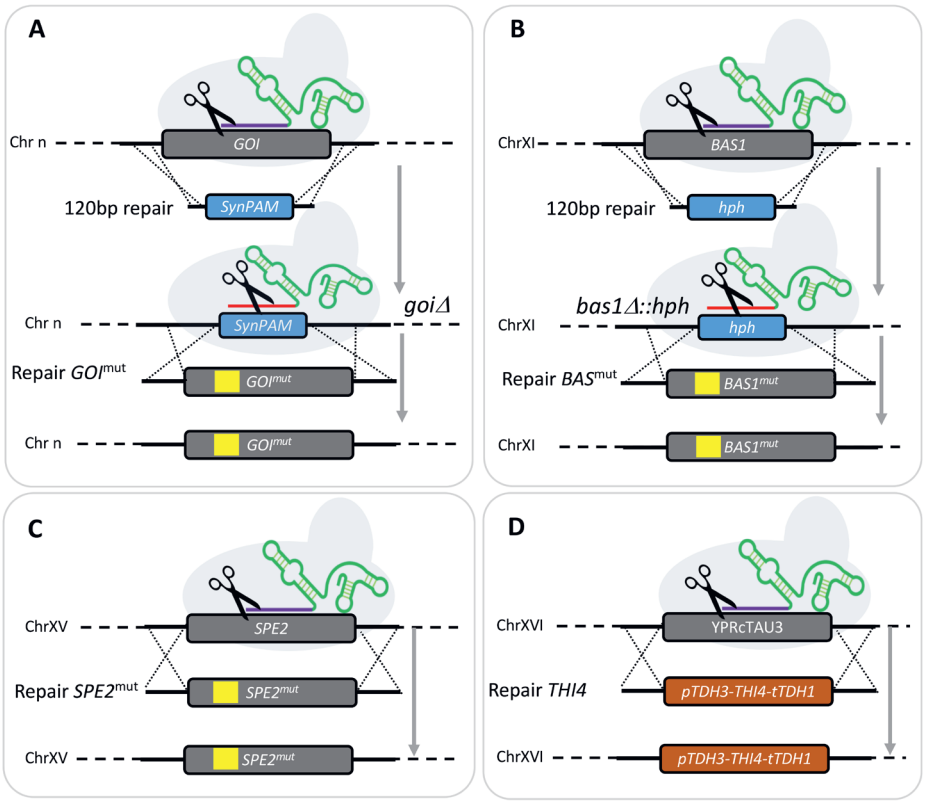


Fig. 6: Strain construction strategy for reverse engineering. Most of the single mutation strains were generated in two steps. First the gene of interest (GOI) was replaced by a synthetic 20 bp target sequence and 3 bp PAM sequence (SynPAM). In a second step, the SynPAM was targeted by Cas9 and substituted with the GOI mutant allele (A). The SynPAM approach was not successful in targeting *BAS1*. For this reason, the *BAS1* mutant strains (IMX2135-7) were constructed by first knocking out the gene by replacing it with the antibiotic marker hphNT1 that confers resistance to hygromycin. Then, in a second step, the selection marker was targeted with Cas9 and substituted with a *BAS1* mutant allele (B). In the case of *SPE2* mutant strains (IMX2289 and IMX2308), the mutant allele was swapped with the WT allele in a single step (C). The *THI4* overexpressing strains IMX2290 and IMX2291 were constructed by integrating a *THI4* overexpression cassette at the YPRcTau3 *locus* (D). SNVs are represented by yellow boxes.

para-Aminobenzoic acid

In SMDΔ lacking *p*ABA, strain CEN.PK113-7D grew 50% slower than in complete SMD. However, it took only a few transfers to achieve fast *p*ABA-independent growth. The independently evolved isolates IMX2057 and IMX1989 harboured mutations affecting genes that encode chorismate-utilizing enzymes, the precursor of *p*ABA (*ABZ1*^{N593H} and *ARO7*^{L205S}, respectively; Fig. 1). As these strains were able to grow in SMD without amino

acid supplementation, these mutations affecting did not cause a complete loss of function. However, they might well affect distribution of chorismate over *p*ABA and aromatic-amino-acid biosynthesis [280, 281]. Introduction of either *ABZ1*^{N593H} or *ARO7*^{L205S}, while replacing the corresponding wild-type allele, eliminated the slower growth observed in strain CEN.PK113-7D during the first transfer on SMDΔ lacking *p*ABA. Specific growth rates of these reverse engineered strains IMX2057 (*ABZ1*^{R593H}) and IMX1989 (*ARO7*^{L205S}) were not statistically different from those of the corresponding evolved isolates (Fig. 6B).

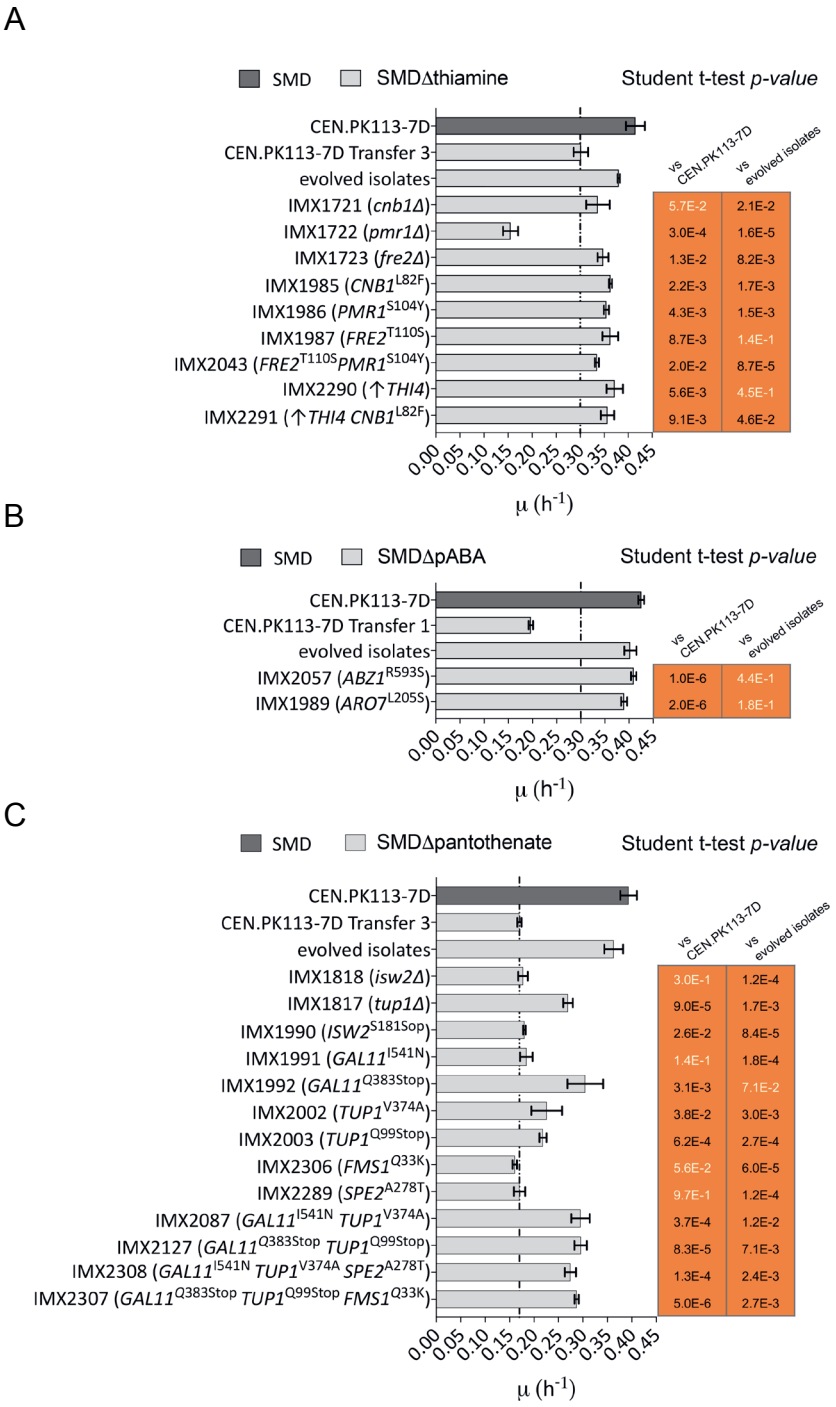
Pantothenic acid

Omission of pantothenic acid from SMD led to a 57% lower specific growth rate of strain CEN.PK113-7D than observed in complete SMD (Fig. 2 and Fig 6C). Out of a total number of 29 mutations found in three independently evolved isolates that showed fast growth in SMDΔ lacking pantothenate, SNVs in *ISW2*, *GAL11*, *TUP1*, *SPE2*, and *FMS1* were analysed by reverse engineering. Single deletion of *SPE2*, *FMS1* and *GAL11* resulted in an inability to grow on SMDΔ lacking pantothenate. This result was anticipated for the *spe2*Δ and *fms1*Δ mutants, in view of the roles of these genes in pantothenate biosynthesis. However, *GAL11* has not previously been implicated in pantothenate biosynthesis. The *gal11*Δ strain was conditional as the mutant did grow on complex YPD and SMD media. Of the remaining two deletion mutants, the *tup1Δ* strain IMX1817 showed a 68 % higher specific growth rate on SMDΔ than strain CEN.PK113-7D (Fig. 6C), while deletion of *ISW2* did not result in faster growth on this medium (Fig. 6C). Of seven SNVs that were individually expressed in the non-evolved strain background, only the *GAL11* ^{Q383Stop} mutation found in IMS0735 supported a specific growth rate of 0.33 h^{-1} on SMDΔ lacking pantothenate that was only 8% lower to that of the evolved isolates.

Combination of the *GAL11*^{Q383Stop} mutation with *TUP1*^{V374A}, *TUP1*^{Q99Stop}, and *SPE2* or *TUP1* with *FMS1* did not lead to additional improvement, indicating that the *GAL11*^{Q383Stop} mutation was predominantly responsible for the improved growth of evolved strains IMS0734 and IMS0735 in the absence of pantothenate.

Pyridoxine

Strain CEN.PK113-7D grew 35% slower on SMDΔ lacking pyridoxine than on complete SMD. Three different mutated alleles of *BAS1* were identified in strains that had been independently evolved for fast growth on the former medium (Table 2). Deletion, in a non-evolved reference strain, of *BAS1* (IMX2128) did not result in faster pyridoxine-independent growth (Fig. 6D). Individual expression of the evolved *BAS1* alleles in strain IMX2128 yielded strains IMX2135 (*BAS1*^{Q152R}), IMX2136 (*BAS1*^{D101N}) and IMX2137 (*BAS1*^{S41P}). All three *BAS1* mutant strains grew faster on SMDΔ lacking pyridoxine than strain CEN.PK113-7D, reaching specific growth rates on this medium that were not


A
SMD
SMDΔthiamine
Student t-test p-value
CEN.PK113-7D
CEN.PK113-7D Transfer 3
evolved isolates
IMX1721 (cnb1Δ)
IMX1722 (pmr1Δ)
IMX1723 (fre2Δ)
IMX1985 (CNB1L82F)
IMX1986 (PMR1S104Y)
IMX1987 (FRE2T110S)
IMX2043 (FRE2T110S PMR1S104Y)
IMX2290 (↑THI4)
IMX2291 (↑THI4 CNB1L82F)
vs CEN.PK113-7D
vs evolved isolates
5.7E-2 2.1E-2
3.0E-4 1.6E-5
1.3E-2 8.2E-3
2.2E-3 1.7E-3
4.3E-3 1.5E-3
8.7E-3 1.4E-1
2.0E-2 8.7E-5
5.6E-3 4.5E-1
9.1E-3 4.6E-2
0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45
μ (h-1)
B
SMD
SMDΔpABA
Student t-test p-value
CEN.PK113-7D
CEN.PK113-7D Transfer 1
evolved isolates
IMX2057 (ABZ1R593S)
IMX1989 (ARO7L205S)
vs CEN.PK113-7D
vs evolved isolates
1.0E-6 4.4E-1
2.0E-6 1.8E-1
0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45
μ (h-1)
C
SMD
SMDΔpantothenate
Student t-test p-value
CEN.PK113-7D
CEN.PK113-7D Transfer 3
evolved isolates
IMX1818 (isw2Δ)
IMX1817 (tup1Δ)
IMX1990 (ISW2S181Sop)
IMX1991 (GAL11I541N)
IMX1992 (GAL11Q383Stop)
IMX2002 (TUP1V374A)
IMX2003 (TUP1Q99Stop)
IMX2306 (FMS1Q33K)
IMX2289 (SPE2A278T)
IMX2087 (GAL11I541N TUP1V374A)
IMX2127 (GAL11Q383Stop TUP1Q99Stop)
IMX2308 (GAL11I541N TUP1V374A SPE2A278T)
IMX2307 (GAL11Q383Stop TUP1Q99Stop FMS1Q33K)
vs CEN.PK113-7D
vs evolved isolates
3.0E-1 1.2E-4
9.0E-5 1.7E-3
2.6E-2 8.4E-5
1.4E-1 1.8E-4
3.1E-3 7.1E-2
3.8E-2 3.0E-3
6.2E-4 2.7E-4
5.6E-2 6.0E-5
9.7E-1 1.2E-4
3.7E-4 1.2E-2
8.3E-5 7.1E-3
1.3E-4 2.4E-3
5.0E-6 2.7E-3
0.00 0.05 0.10 0.15 0.20 0.25 0.30 0.35 0.40 0.45
μ (h-1)

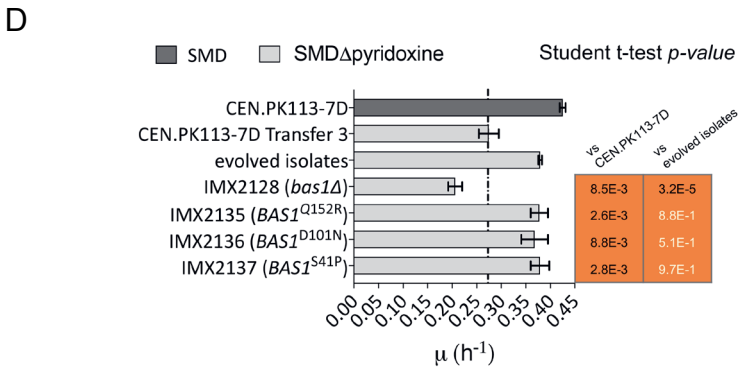


Fig. 7: Specific growth rates of engineered *S. cerevisiae* strains carrying one or multiple gene deletions or reverse engineered mutations in SMD media lacking thiamine (A), *p*ABA (B), pantothenic acid (C) and pyridoxine (D). Specific growth rates of *S. cerevisiae* CEN.PK113-7D grown in complete SMD and evolved CEN.PK113-7D in SMD medium lacking the relevant vitamin are shown as references. The specific growth rate of strain CEN.PK113-7D in SMD medium lacking the relevant vitamin is shown and highlighted by a vertical line to help to visualize improved performance of engineered strains. Error bars represent the standard deviation (n = 9 for complete SMD, n = 6 for strain IMX1721, otherwise n=3). A Student t-test was performed to compare the wild-type and evolved CEN.PK113-7D growth rate to the engineered strains growth rate and non-significant differences are indicated with white letters (p-value > 0.05).

significantly different from the average of those of evolved strains IMS736, IMS737, and IMS738 (Figure 5D). These results suggest that *BAS1*, which has previously been shown to be involved in regulation of purine and histidine biosynthesis [99, 100], may also be involved in regulation of pyridoxine biosynthesis in *S. cerevisiae.*

Discussion

Vitamin requirements of *S. cerevisiae*

Most *S. cerevisiae* genomes harbor the full complement of genes required for synthesis of the seven B-vitamins that are commonly included in chemically defined media for yeast cultivation (CDMY, for a recent review see [154, 267]). Previous studies indicated that presence of a complete set of biotin biosynthesis genes supported only slow growth on CDMY. The present study shows that, similarly, none of the other six B-vitamins included in CDMY (inositol, nicotinic acid, pantothenic acid, *p*ABA, pyridoxine and thiamine) are strictly required for growth. Remarkably, the impact of individually eliminating these six vitamins from a glucose-containing CDMY differently affected specific growth rates in aerobic, glucose-grown cultures, with growth-rate reductions varying from 0 to 57 %. It should, however, be noted that requirements for these growth factors, which for aerobic

yeast cultivation cannot be formally defined as vitamins and that their absolute and relative requirements may well be condition- and strain dependent. For example, it is well documented that synthesis of nicotinic acid by *S. cerevisiae* is strictly oxygen dependent [54]. The dataset compiled in the present study will, hopefully, serve as reference for investigating vitamin requirements of diverse natural isolates, laboratory and industrial strains and thereby help to obtain a deeper understanding of the genetics and ecology of vitamin prototrophy and vitamin biosynthesis in *S. cerevisiae.*

ALE and reverse engineering for identifying genes involved in fast B-vitamin independent growth

A serial transfer strategy was applied to select for spontaneous mutants that grew as fast in aerobic batch cultures on CDMY lacking either inositol, nicotinic acid, pyridoxine, thiamine, pantothenic acid, or *para*-aminobenzoic acid as in CDMY containing all these six vitamins as well as biotin. In the ALE experiments on media lacking nicotinic acid or inositol, fast growth was observed within a few cycles of batch cultivation and not all fast-growing strains were found to contain mutations. These observations indicated that, under the experimental conditions, the native metabolic and regulatory network of *S. cerevisiae* was able to meet cellular requirements for fast growth in the absence of these 'vitamins'.

As demonstrated in other ALE studies, performing independent replicate evolution experiments helped in identifying biologically relevant mutations upon subsequent whole-genome sequencing [268, 269]. The power of this approach is illustrated by the ALE experiments that selected for pyridoxine-independent growth, in which the independently evolved mutants IMS0736 and IMS0738 harboured 2 and 30 mutated genes, respectively, of which only *BAS1* also carried a mutation in a third, independently sequenced isolate (Fig. 5).

In total, the role of 12 genes that were found to be mutated in the ALE experiments were selected for further analysis by reverse engineering of the evolved alleles and/or deletion mutations in the parental, non-evolved genetic background (Fig. 5 and Table 2). These genes comprised three groups; i) genes encoding enzymes known or inferred to be involved in the relevant vitamin synthesis pathway (*SPE2* and *FMS1* for pantothenate, *THI4* for thiamine, *ABZ1* and *ARO7* for *p*ABA), ii) genes encoding transcriptional regulator proteins (*TUP1* and *GAL11* for pantothenate and *BAS1* for pyridoxine) and iii) non-transcriptional-regulator proteins whose functions have not previously been associated with vitamin biosynthesis (*ISW2* for pantothenate and *CNB1, PMR1* and *FRE2* for thiamine).

Of the first group of mutations defined above, only those in *SPE2* and *FMS1* were not found to contribute to faster growth in the absence of the relevant vitamin. The

second group yielded interesting information on regulation of vitamin biosynthesis in *S. cerevisiae*. In particular, the key role of mutations in *BAS1* in enabling fast pyridoxine-independent growth and the role of *GAL11* mutations in fast pantothenate-independent growth dependency provided interesting insights and leads for further research.

The *S. cerevisiae* transcriptional activator Bas1 is involved in regulation of purine and histidine [99, 100]. Interestingly, Bas1 is also involved in repression of genes involved in C1 metabolism and of *SNZ1* [102]. Snz1 is a subunit of a two-component pyridoxal-5'-phosphate synthase, which catalyses the first step of the synthesis of pyridoxal-5-phosphate, the active form of pyridoxine in *S. cerevisiae* [92]. Interrogation of the Yeastract database [293] for occurrence of transcription binding sites in promoter regions of pyridoxine-biosynthesis genes confirmed the link already established between *BAS1* and *SNZ1* [101, 102]. Moreover, this analysis revealed that all pyridoxine biosynthesis genes in *S. cerevisiae* contain a consensus Bas1 cis-regulatory binding motif (Fig. 8). Consistent with the regulatory role of Bas1 on *SNZ1* expression, Bas1 has been experimentally shown to repress transcription of genes involved in pyridoxine biosynthesis [294]. The mutations found in *BAS1* may, therefore, have attenuated Bas1-mediated repression of pyridoxine biosynthetic genes and, thereby, enabled increased pyridoxine biosynthesis.

ALE experiments in pantothenate-free medium yielded different mutations in *TUP1* and *GAL11*, two major components of the yeast regulatory machinery. *TUP1* encodes a general transcriptional repressor that, in a complex with Cyc8, modifies chromatin structure such that genes are repressed [296-298]. *GAL11* (also known as *MED15*) encodes a subunit of the mediator complex required for initiation by RNA polymerase-II and consequently plays a critical role in transcription of a large number of RNA polymerase-II dependent genes [299, 300]. Despite its involvement in general cellular transcriptional regulation, *GAL11* is not an essential gene for growth in complete medium [301]. The inability of

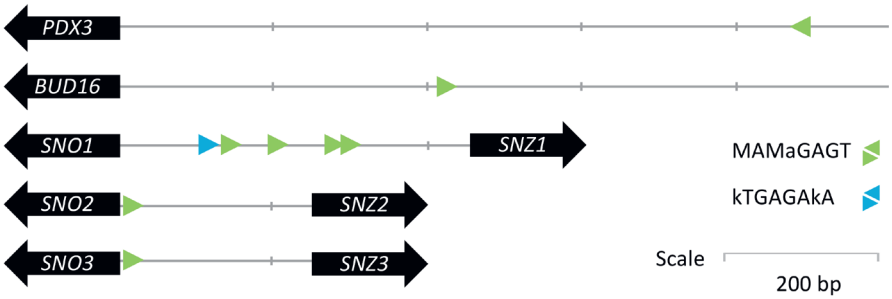


Fig. 8: Schematic representation of Bas1 binding sites in promoter regions of genes involved in pyridoxal-5-phosphate biosynthesis. The two Bas1 consensus binding sequences MAMaGAGT and kTGAGAkA (Fordyce et al., 2010) are shown in green and blue respectively. Scale bar indicates 200 bp.

a *gal11Δ* strain to grow on glucose synthetic medium without pantothenate represents the first indication for a possible involvement of Gal11 in regulation of pantothenate metabolism. Gal11 interacts with transcriptional activators through various peptidic segments including an N-terminal KIX domain . This region shows homology with the B-box motif found in the mammalian activating protein SRC-1 and is essential for recruitment of the mediator complex by other regulatory proteins (e.g. Gcn4) [302]. Of two mutations found in *GAL11*, the most potent was a nonsense mutation at nucleotide 383. In contrast to a *gal11*Δ strain, a reverse engineered strain carrying this premature stop codon grew on SMDΔ pantothenate, which indicates that the *GAL11*^{Q383Stop} allele encodes a functional peptide. Such a functional, truncated Gal11 version has not been previously described, and is sufficiently long to include a complete KIX domain (AA_9-AA_{86}) for recruitment of the RNA polymerase-II machinery by an as yet unidentified transcription factor involved in regulation of pantothenate biosynthesis. Further research is required to resolve and understand the role of the wild-type and evolved alleles of *GAL11* in regulation of pantothenate metabolism.

A third group of non-transcription factor genes had not yet been associated with the biosynthesis of vitamins. Reverse engineering of a mutation in *ISW2*, which encodes an ATP-dependent DNA translocase involved in chromatin remodeling [303] identified in the pantothenate evolution did not yield a growth improvement, but we cannot exclude that this mutation in association with *ERG3, AMN1*, *DAN4,* and *ERR3* identified in IMS0733 (Fig. 5 and 6C) could have a significant impact but systematic combinatorial analysis of the mutations was not performed.

Mutations in *CNB1*, *PMR1* and *FRE2* identified in evolved isolates all improved growth of *S. cerevisiae* in the absence of thiamine (Fig. 7A). These three genes all encode proteins involved in metal homeostasis, Fre2 is a ferric or cupric reductase [304], Cnb1 is the regulatory B-subunit of calcineurin, a Ca^{2+}/calmodulin-regulated type 2B protein phosphatase which regulates the nuclear localization of Crz1. This transcription factor influences expression of a large number of genes. Its targets include *PMR1*, which encodes a high-affinity Ca^{2+}/Mn^{2+} P-type ATPase involved in Ca^{2+} and Mn^{2+} transport into the Golgi [284, 305]. Neither of these three genes have hitherto been directly associated with thiamine. However, thiamine pyrophosphokinase (Thi80), thiamin phosphate synthase (Thi6) and hydroxymethylpyrimidine phosphate (Thi21 and Thi20) all require Mg^{2+} or Mn^{2+} as co-factors [306, 307]. At low concentration, Mn^{2+} was shown to be a stronger activator of Thi80 than Mg^{2+}[308]. In an ALE experiment with engineered xylose-fermenting assimilating *S. cerevisiae*, a non-sense mutation or deletion of *PMR1* caused selectively and strongly increased intracellular concentrations of Mn^{2+}, which was the preferred metal ion for the heterologously expressed *Piromyces* xylose isomerase [309]. Although intracellular metal ion concentrations were not measured in the current study,

the different phenotypes of a *pmr1*Δ deletion strain (IMX1722) and a *PMR1*^{S104Y} strain (IMX1986) (Fig. 7A) indicate that the latter mutation does not act through a massive increase of the intracellular Mn^{2+} concentration.

In *S. cerevisiae,* synthesis of the thiazole moiety of thiamine biosynthesis involves sulfide transfer from an active-site cysteine (Cys205) residue of the thiazole synthase Thi4. This sulfur transfer reaction is iron-dependent and generates inactive enzyme by formation of a dehydroalanine. Fe(II) plays an essential role in this sulfide transfer, which remains poorly understood [121]. Further research is needed to investigate if the *FRE2* mutation in strain IMS0749 in some way increases the efficiency of the reaction catalyzed by the energetically single-turnover enzyme Thi4 and to resolve the role of metal homeostasis in vitamin biosynthesis.

Towards mineral media for cultivation of *S. cerevisiae*

With the exception of the carbon and energy source for growth, B-vitamins are the sole organic ingredients in standard CDMY recipes for aerobic cultivation of wild-type and industrial *S. cerevisiae* strains. In view of the chemical instability of some of these compounds, vitamin solutions cannot be autoclaved along with other medium components but are usually filter sterilized. In research laboratories and in particular in industrial processes, the costs, complexity and contamination risks associated with the use of vitamins is significant. Complete elimination of vitamins from CDMY, without compromising on specific growth rate, yield or productivity, could therefore result in consideral cost and time savings as well as in improved standardization and robustness of cultivation procedures.

The present study demonstrates that, by ALE as well as introduction of small sets of defined mutations into *S. cerevisiae*, it is possible to achieve specific growth rates in single-vitamin depleted CDMY that are close or identical to those found in CDMY supplemented with a complete vitamin mixture. While these results represent a first step towards the construction of completely prototrophic growth of *S. cerevisiae* and related yeasts, further research is required which trade-offs are incurred upon simultaneous introduction of the genetic interventions identified in this study and how they can be mitigated. This issue may be particularly relevant for mutations that affect genes involved in global regulation processes [300, 310], which may interfere with other cellular processes. In addition, simultaneous high-level expression of multiple enzymes with low-catalytic turn-over numbers, with the suicide enzyme Thi4 [119-121] as an extreme example, may affect cell physiology due to the required resource allocation [311, 312].

In such cases, it may be necessary to expand metabolic engineering strategies beyond the native metabolic and regulatory capabilities of *S. cerevisiae* by expression of heterologous proteins and/or pathways with more favourable characteristics [157].

Materials and Methods

Strains, media and maintenance

The *S. cerevisiae* strains used and constructed in this study are shown in Table 3 and they all derive from the CEN.PK lineage [151, 273]. Yeast strains were grown on synthetic medium with ammonium sulfate as a nitrogen source (SM) or YP medium (10 g/L Bacto yeast extract, 20 g/L Bacto peptone) as previously described [13]. SM and YP media were autoclaved at 121°C for 20 min. Then, SM medium was supplemented with 1 ml/L of filter-sterilized vitamin solution (0.05 g/L D-(+)-biotin, 1.0 g/L D-calcium pantothenate, 1.0 g/L nicotinic acid, 25 g/L myo-inositol, 1.0 g/L thiamine hydrochloride, 1.0 g/L pyridoxol hydrochloride, 0.20 g/L 4-aminobenzoic acid). Vitamin drop-out media were prepared using vitamin solutions lacking either thiamine, pyridoxine, pantothenic acid, inositol, nicotinic acid or *para*-aminobenzoic acid, yielding SMΔthiamine, SMΔpyridoxine, SMΔpantothenic acid, SMΔinositol, SMΔnicotinic acid, SMΔ*p*ABA respectively. A concentrated glucose solution was autoclaved at 110 °C for 20 min and then added to the SM and YP medium at a final concentration of 20 g/L, yielding SMD and YPD, respectively. 500 ml shake flasks containing 100 ml medium and 100 ml shake flasks containing 20 ml medium were incubated at 30 °C and at 200 rpm in an Innova Incubator (Brunswick Scientific, Edison, NJ). Solid media were prepared by adding 1.5% Bacto agar and, when indicated, 200 mg/L G418 or 200 mg/L hygromycin. *Escherichia coli* strains were grown in LB (10 g/L Bacto tryptone, 5 g/L Bacto yeast extract, 5 g/L NaCl) supplemented with 100 mg/L ampicillin or kanamycin. *S. cerevisiae* and *E. coli* cultures were stored at -80 °C after the addition of 30% v/v glycerol.

Molecular biology techniques

PCR amplification of DNA fragments with Phusion Hot Start II High Fidelity Polymerase (Thermo Scientific, Waltham, MA) and desalted or PAGE-purified oligonucleotide primers (Sigma-Aldrich, St Louis, MO) was performed according to manufacturers' instructions. DreamTaq polymerase (Thermo Scientific) was used for diagnostic PCR. Primers used in this study are shown in Table 5. PCR products were separated by gel electrophoresis using 1 % (w/v) agarose gels (Thermo Scientific) in TAE buffer (Thermo Scientific) at 100 V for 25 min and purified with either GenElutePCR Clean-Up Kit (Sigma-Aldrich) or with Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA). Plasmids were purified from *E. coli* using a Sigma GenElute Plasmid Kit (Sigma Aldrich). Plasmids used in this study are shown in Table 4. Yeast genomic DNA was isolated with the SDS/LiAc protocol [314]. Yeast strains were transformed with the lithium acetate method [315]. Four to eight single colonies were re-streaked three consecutive times on selective media and diagnostic PCR were performed in order to verify their genotype. *E. coli* XL1-blue was

Table 3: *Saccharomyces cerevisiae* strains used in this study.

Strain ID	Relevant genotype	Parental strain	Reference
CEN.PK113-7D	MATa		[273]
CEN.PK113-5D	MATa *ura3-52*		[273]
IMX585	MATa *can1*Δ::*cas9*-natNT2 U	CEN.PK113-7D	[313]
IMS0721	*MATa* evolved in SMD colony 1	CEN.PK113-7D	This study
IMS0722	*MATa* evolved in SMD colony 2	CEN.PK113-7D	This study
IMS0723	*MATa* evolved in SMD colony 3	CEN.PK113-7D	This study
IMS0724	*MATa* evolved in Δnicotinic acidSMD colony 1	CEN.PK113-7D	This study
IMS0725	*MATa* evolved in Δnicotinic acidSMD colony 2	CEN.PK113-7D	This study
IMS0726	*MATa* evolved in Δnicotinic acidSMD colony 3	CEN.PK113-7D	This study
IMS0727	*MATa* evolved in ΔpabaSMD colony 1	CEN.PK113-7D	This study
IMS0728	*MATa* evolved in ΔpabaSMD colony 2	CEN.PK113-7D	This study
IMS0729	*MATa* evolved in ΔpabaSMD colony 3	CEN.PK113-7D	This study
IMS0730	*MATa* evolved in ΔinositolSMD colony 1	CEN.PK113-7D	This study
IMS0731	*MATa* evolved in ΔinositolSMD colony 2	CEN.PK113-7D	This study
IMS0732	*MATa* evolved in ΔinositolSMD colony 3	CEN.PK113-7D	This study
IMS0733	*MATa* evolved in Δpantothenic acidSMD colony 1	CEN.PK113-7D	This study
IMS0734	*MATa* evolved in Δpantothenic acidSMD colony 2	CEN.PK113-7D	This study
IMS0735	*MATa* evolved in Δpantothenic acidSMD colony 3	CEN.PK113-7D	This study
IMS0736	*MATa* evolved in ΔpyridoxineSMD colony 1	CEN.PK113-7D	This study
IMS0737	*MATa* evolved in ΔpyridoxineSMD colony 2	CEN.PK113-7D	This study
IMS0738	*MATa* evolved in ΔpyridoxineSMD colony 3	CEN.PK113-7D	This study
IMS0747	*MATa* evolved in ΔthiamineSMD colony 1	CEN.PK113-7D	This study
IMS0748	*MATa* evolved in ΔthiamineSMD colony 2	CEN.PK113-7D	This study
IMS0749	*MATa* evolved in ΔthiamineSMD colony 3	CEN.PK113-7D	This study
IMX1721	MATa *can1*Δ::*cas9*-natNT2 *cnb1*Δ::SynPAM	IMX585	This study
IMX1722	MATa *can1*Δ::*cas9*-natNT2 *pmr1*Δ::SynPAM	IMX585	This study
IMX1723	MATa *can1*Δ::*cas9*-natNT2 *fre2*Δ::SynPAM	IMX585	This study
IMX1817	MATa *can1*Δ::*cas9*-natNT2 *tup1*Δ::SynPAM	IMX585	This study
IMX1818	MATa *can1*Δ::*cas9*-natNT2 *isw2*Δ::SynPAM	IMX585	This study
IMX1819	MATa *can1*Δ::*cas9*-natNT2 *gal11*Δ::SynPAM	IMX585	This study
IMX1920	MATa *can1*Δ::*cas9*-natNT2 *aro7*Δ::SynPAM	IMX585	This study

IMX1985	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*CNB1*^{L82F}	IMX1721	This study
IMX1986	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*PMR1*^{S104Y}	IMX1722	This study
IMX1987	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*FRE2*^{T110S}	IMX1723	This study
IMX1988	MATa *can1*Δ::*cas9*-natNT2 *abz1*Δ::SynPAM	IMX585	This study
IMX1989	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*ARO7*^{L205S}	IMX1920	This study
IMX1990	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*ISW2*^{S181Stop}	IMX1818	This study
IMX1991	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*GAL11*^{I541N}	IMX1819	This study
IMX1992	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*GAL11*^{Q383Stop}	IMX1819	This study
IMX2002	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*TUP1*^{V374A}	IMX1817	This study
IMX2003	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*TUP1*^{Q99Stop}	IMX1817	This study
IMX2043	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*FRE2*^{T110S} *pmr1*Δ::*PMR1*^{S104Y}	IMX1986	This study
IMX2057	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*ABZ1*^{R593H}	IMX1988	This study
IMX2066	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*TUP1*^{V374A} *gal11*Δ::SynPAM	IMX2002	This study
IMX2110	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*GAL11*^{Q383Stop} *tup1*Δ::SynPAM	IMX1992	This study
IMX2127	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*GAL11*^{Q383Stop} SynPAMΔ::*TUP1*^{Q99Stop}	IMX2110	This study
IMX2128	MATa *can1*Δ::*cas9*-natNT2 *bas1*Δ::hphNT1	IMX585	This study
IMX2087	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*TUP1*^{V374A} SynPAMΔ::*GAL11*^{I541N}	IMX2066	This study
IMX2135	MATa *can1*Δ::*cas9*-natNT2 hphNT1Δ::*BAS1*^{Q152R}	IMX2128	This study
IMX2136	MATa *can1*Δ::*cas9*-natNT2 hphNT1Δ::*BAS1*^{D101N}	IMX2128	This study
IMX2137	MATa *can1*Δ::*cas9*-natNT2 hphNT1Δ::*BAS1*^{S41P}	IMX2128	This study
IMX2290	MATa *can1*Δ::*cas9*-natNT2 YPRcTau3::*pTDH3-THI4-tTDH1*	IMX585	This study

IMX2291	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*CNB1*^{L82F} YPRcTau3::*pTDH3-THI4-tTDH*1	IMX1985	This study
IMX2289	MATa *can1*Δ::*cas9*-natNT2 *SPE2*^{A278T}	IMX585	This study
IMX2292	MATa *can1*Δ::*cas9*-natNT2 *fms1*Δ::SynPAM	IMX585	This study
IMX2306	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*FMS1*^{Q33K}	IMX2292	This study
IMX2308	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*GAL11*^{Q383Stop} SynPAMΔ::*TUP1*^{V374A} *SPE2*^{A278T}	IMX2127	This study
IMX2294	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*TUP-1*^{V374A} SynPAMΔ::*GAL11*^{I541N} *FMS1*^{Q33K}::SynPAM	IMX2087	This study
IMX2307	MATa *can1*Δ::*cas9*-natNT2 SynPAMΔ::*TUP1*^{V374A} SynPAMΔ::*GAL11*^{I541N} SynPAMΔ::*FMS1*^{Q33K}	IMX2294	This study

used for chemical transformation [316]. Plasmids were then isolated and verified by either restriction analysis or by diagnostic PCR.

Laboratory evolution

Laboratory evolution of *S. cerevisiae* CEN.PK113-7D for fast growth in SMD medium lacking a single vitamin was performed by sequential transfer in aerobic shake-flask batch cultures. A frozen aliquot of strain CEN.PK113-7D was inoculated in a pre-culture shake flask containing SMD medium supplemented with all vitamins. Cells were then spun down, washed twice with sterile water and used to inoculate a second shake flask containing SMD lacking one of the vitamins. The culture was then grown until stationary phase and transferred in a third shake flask containing the same fresh medium. At each transfer, 0.2 ml culture broth were transferred to 20 ml fresh medium, corresponding to about 6.7 generations in each growth cycle. The evolution experiment was performed in SMΔthiamine, SMΔpyridoxine, SMΔpantothenic acid, SMΔinositol, SMΔnicotinic acid, SMΔ*p*ABA media. Each evolution experiment was performed in triplicate. After a defined number of transfers, intermediate strains were stocked and characterized for the growth rate. The experiment was stopped once the target specific growth rate of 0.35 h^{-1} was reached. From each evolved population, three single colonies were then isolated and stored. The specific growth rate of these single cell lines was measured to verify that they were representative of the evolved population. The best performing isolate from each evolution line was selected for whole-genome sequencing.

Table 4: Plasmids used in this study.

Plasmid	Relevant characteristics	References
pROS12	colE1^{ori} 2µm *bla hphNT1* gRNA-CAN1.Y gRNA-ADE2.Y	[268]
pROS13	colE1^{ori} 2µm *bla aph* gRNA-*CAN1*.Y gRNA-*ADE2*.Y	[268]
pUDR412	colE1^{ori} 2µm *bla* hphNT1 gRNA-ARO7 gRNA-ARO7	[275]
pYTK009	colE1^{ori} *cat pTDH3*	[317]
pYTK056	colE1^{ori} *cat tTDH1*	[317]
pYTK096	colE1^{ori} *aph URA3* 5' homology *sfGFP URA3 URA3* 3' homology	[317]
pUDR388	colE1^{ori} 2µm *bla aph* gRNA-*CNB1* gRNA-*CNB1*	This study
pUDR389	colE1^{ori} 2µm *bla aph* gRNA-PMR1 gRNA-PMR1	This study
pUDR390	colE1^{ori} 2µm *bla aph* gRNA-FRE2 gRNA-FRE2	This study
pUDR438	colE1^{ori} 2µm *bla aph* gRNA-ABZ1 gRNA-ABZ1	This study
pUDR441	colE1^{ori} 2µm *bla* hphNT1 gRNA-GAL11 gRNA-GAL11	This study
pUDR471	colE1^{ori} 2µm *bla aph* gRNA-SynPAM gRNA-SynPAM	This study
pUDR472	colE1^{ori} 2µm *bla aph* gRNA-TUP1 gRNA-TUP1	This study
pUDR473	colE1^{ori} 2µm *bla aph* gRNA-ISW2 gRNA-ISW2	This study
pUDR566	colE1^{ori} 2µm *bla aph* gRNA-BAS1 gRNA-BAS1	This study
pUDR592	colE1^{ori} 2µm *bla aph* gRNA-hphNT1 gRNA- hphNT1	This study
pUDR652	colE1^{ori} 2µm *bla aph* MX gRNA-FMS1 gRNA-FMS1	This study
pUDR651	colE1^{ori} 2µm *bla aph* gRNA-SPE2 gRNA-SPE2	This study
pUDR514	colE1^{ori} 2µm *bla aph* gRNA-YPRcTau3 gRNA-YPRcTau3	This study
pUDI180	colE1^{ori} *aph pTDH3-ScTHI4-tTDH1*	This study

Shake flask growth experiments

For specific growth rate measurements of strains (evolved populations as well as single cell lines), an aliquot was used to inoculate a shake flask containing 100 ml of fresh medium. For specific growth rate measurements of the engineered strains, a frozen aliquot was thawed and used to inoculate a 20 ml starter culture that was then used to inoculate the 100 ml flask. An initial OD_{660} of 0.1 or 0.2 was used as a starting point. The flasks were then incubated, and growth was monitored using a 7200 Jenway Spectrometer (Jenway, Stone, United Kingdom). Specific growth rates were calculated from at least four time-points in the exponential growth phase of each culture.

Table 5: Oligonucleotide primers used in this study.

Primer ID	Sequence	Product(s)
6005	GATCATTTATCTTTCACTGCGGAGAAG	gRNA pROS plasmid backbone amplification
6006	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC	gRNA pROS plasmid backbone amplification
14229	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAA-GATAAATGATCAGTAGAATTTCACCTAGACGGTTTTA-GAGCTAGAAATAGCAAGTTAAAATAAG	2µm fragment for SynPAM gRNA plasmid
13686	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAA-GATAAATGATCCTGCGGTGATAGAACCCTGGGTTTTA-GAGCTAGAAATAGCAAGTTAAAATAAG	2µm fragment for *ABZ1* gRNA plasmid
14988	CTTTTACACGATGACCTTTCGAGATTTCACAAGGGGGATAAA-GGAAGTAGAATTTCACCTAGACGTGGATATTTGTATATTATT-AGATATGTATGCAAACATTTTCTTTAGAA	*ABZ1* KO repair oligo
14989	TTCTAAAGAAAATGTTTGCATACATATCTAATAATAT-ACAAATATCCACGTCTAGGTGAAATTCTACTTCCT-TTATCCCCCTTGTGAAATCTCGAAAGGTCATCGTGTAAAAG	*ABZ1* KO repair oligo
13693	AAACCGCGAATATATAAAAACAAGC	*ABZ1* mutant allele amplification
13694	GGCACAAAACGTCATTTTCC	*ABZ1* mutant allele amplification
15075	TAATCACTCGGCAATGTGGAATTGTTACCGTGATAGCCT-TCATGCAGTAGAATTTCACCTAGACGTGGATCTTATACCAAT-TTTATGCAGGATGCTGAGTGCTATTTGTTAGC	*ARO7* KO repair oligo
15076	GCTAACAAATAGCACTCAGCATCCTGCATAAAATTGGTATAA-GATCCACGTCTAGGTGAAATTCTACTGCATGAAGGCTAT-CACGGTAACAATTCCACATTGCCGAGTGATTA	*ARO7* KO repair oligo
12052	CAGGAGTCTCTGAGCAAGGC	*ARO7* mutant allele amplification
12053	ACCATGCTAAGAGCTGCTCC	*ARO7* mutant allele amplification
15037	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAA-GATAAATGATCAGCATCAGAAGTAATAACAAGTTTTA-GAGCTAGAAATAGCAAGTTAAAATAAG	2µm fragment for *BAS1* gRNA plasmid
15584	AAACTTTTGTTGTAGCGTTTTTGCTTTTTTTTTTTTTATCG-CAGAATACATTTTATCGAGATAGGTCTAGAGATCTGT-TTAGCTTGC	Repair fragment with HphNT1 for *BAS1* KO
15585	ATTACAAAACTAATATGTTAAACAATTGAAAGATTTGTGT-TTTTTTTCGGCCTTGCCTTCAGCTCCAGCTTTTGTTCCC	Repair fragment with HphNT1 for *BAS1* KO
13687	CCTTTGACGATGTGCAACGG	Amplification *BAS1* mutant allele
13688	AACGCCCTTTGTGTTTGTGG	Amplification *BAS1* mutant allele

13520	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCTCTTGCTGGACGTATAATGGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG	2µm fragment for *CNB1* gRNA plasmid
13612	ACTCAATGGTGATCAGAATCCATAGAAGCATTTTTATTTCTTAAAAGTAGAATTTCACCTAGACGTGGGACTAGGGGACACTTCATTCATTTATGGTATGCCAATATTTTTAA	*CNB1* KO repair oligo
13613	TTAAAAATATTGGCATACCATAAATGAATGAAGTGTCCCCTAGTCCCACGTCTAGGTGAAATTCTACTTTTAAGAAATAAAAATGCTTCTATGGATTCTGATCACCATTGAGT	*CNB1* KO repair oligo
13523	GCATCAGCACTGCAGAATCG	*CNB1* mutant allele amplification
13524	GATCCCCCTTTGTGCATTGC	*CNB1* mutant allele amplification
13521	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCCATAAAAAGAGAGACCACTGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG	2µm fragment for *PMR1* gRNA plasmid
13541	CCAGCACAGACGTAAGCTTAAGTGTAAGTAAAAGATAAGATAATTAGTAGAATTTCACCTAGACGTGGTATGTCACATTTTGTGCTTTTATCGTTTTTCCTTCCTTCCCTTTA	*PMR1* KO repair oligo
13542	TAAAGGGAAGGAAGGAAAAACGATAAAAGCACAAAATGTGACATACCACGTCTAGGTGAAATTCTACTAATTATCTTATCTTTTACTTACACTTAAGCTTACGTCTGTGCTGG	*PMR1* KO repair oligo
11292	TCGCCCCGTTCTTTCCATTC	*PMR1* mutant allele amplification
11293	GGGCGAAAAGGTAAGAACGC	*PMR1* mutant allele amplification
13522	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCCATAAAAGAACATTGCACCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG	2µm fragment for *FRE2* gRNA plasmid
13539	AATAAAGTCTTTTTTATCCAAAGCTTATGAAACCCAACGAATATAAGTAGAATTTCACCTAGACGTGGTCATTTTTTACTTAAAACTAGTCATTTCATTAATAATACCTATCC	*FRE2* KO repair oligo
13540	GGATAGGTATTATTAATGAAATGACTAGTTTTAAGTAAAAAATGACCACGTCTAGGTGAAATTCTACTTATATTCGTTGGGTTTCATAAGCTTTGGATAAAAAAGACTTTATT	*FRE2* KO repair oligo
13524	GATCCCCCTTTGTGCATTGC	*FRE2* mutant allele amplification
13525	TGGCTCAATGATGCTAGTGGG	*FRE2* mutant allele amplification
12174	GCATCGTCTCATCGGTCTCATATGTCTGCTACCTCTACTGCTACTTCC	*THI4* with YTK part 3 compatible overhangs
12175	ATGCCGTCTCAGGTCTCAGGATCTAAGCAGCAAAGTGTTTCAAAATTTG	*THI4* with YTK part 3 compatible overhangs
14586	ACAGTTTTGACAACTGGTTACTTCCCTAAGACTGTTTATATTAGGATTGTCAAGACACTCCAGTTCGAGTTTATCATTATCAATAC	*THI4*↑ cassette repair for integration
14587	ATAATTATAATATCCTGGACACTTTACTTATCTAGCGTATGTTATTACTCGATAAGTGCTCGTTCAGGGTAATATATTTTAACC	*THI4*↑cassette repair for integration

13518	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAA-GATAAATGATCTGAATCTGGTGATAGCACCGGTTTTA-GAGCTAGAAATAGCAAGTTAAAATAAG	2µm fragment for *GAL11* gRNA plasmid
13533	TACTCAAAGATCAAGGATTAAAACGCTATTTCT-TTTAAATCTGCTAGTAGAATTTCACCTAGACGTGGACATTT-GAAGTTTCCATACTTTTGATACTTTTGAAGTTACTTCGT	*GAL11* KO repair oligo
13534	ACGAAGTAACTTCAAAAGTATCAAAAGTATGGAAACT-TCAAATGTCCACGTCTAGGTGAAATTCTACTAGCAGAT-TTAAAAGAAATAGCGTTTTAATCCTTGATCTTTGAGTA	*GAL11* KO repair oligo
13498	TTCGAATCGGGCCTTCCTTC	*GAL11* mutant allele amplification
13499	TGCTTGAAGTGGCACTTTGC	*GAL11* mutant allele amplification
13517	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAA-GATAAATGATCTGGAAGGGTAGACCATGACAGTTTTA-GAGCTAGAAATAGCAAGTTAAAATAAG	2µm fragment for *TUP1* gRNA plasmid
13531	TGATAAGCAGGGGAAGAAAGAAATCAGCTTTCCATCCAAAC-CAATAGTAGAATTTCACCTAGACGTGGGAACAGAACACAAAA-GGAACACTTTACAAATGTAACTAACTAAAC	*TUP1* KO repair oligo
13532	GTTTAGTTAGTTACATTTGTAAAGTGTTCCTTTTGTGTTCT-GTTCCCACGTCTAGGTGAAATTCTACTATTGGTTTGGATG-GAAAGCTGATTTCTTTCTTCCCCTGCTTATCA	*TUP1* KO repair oligo
15077	CACGCCAAGTTACCTTTCGC	*TUP1* mutant allele amplification
15078	GGAAGGGATGAATGGTGAGG	*TUP1* mutant allele amplification
13519	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAA-GATAAATGATCGAAAAAGAGAAGGCAAAACGGTTTTA-GAGCTAGAAATAGCAAGTTAAAATAAG	2µm fragment for *ISW2* gRNA plasmid
13535	CTTGTTGGTTTAAGTCGTAACAAAAGGAAAACTTACAAT-CAGATCAGTAGAATTTCACCTAGACGTGGATCATGTATTGTG-CATTAAAATAAGTGACGTGAGAGATATAATTT	*ISW2* KO repair oligo
13536	AAATTATATCTCTCACGTCACTTATTTTAATGCACAATACAT-GATCCACGTCTAGGTGAAATTCTACTGATCTGATTGTAAGT-TTTCCTTTTGTTACGACTTAAACCAACAAG	*ISW2* KO repair oligo
13496	TCACCCAGAGGCAAAAGGTG	*ISW2* mutant allele amplification
13497	TAGTTAAAGCGGCTCGACCC	*ISW2* mutant allele amplification
16598	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAA-GATAAATGATCTCAAGATTGTCTTGTTCTTGGTTTTA-GAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTAT-CAAC	2µm fragment for *FMS1* gRNA plasmid
13527	AACAAGAAGTGAGTTAATAAAGGCAAAAACAGTGGTCGT-GTGAGAAGTAGAATTTCACCTAGACGTGGAATCTAT-TTTTTCGAAATTACTTACACTTTTGACGGCTAGAAAAG	*FMS1* KO repair oligo
13528	CTTTTCTAGCCGTCAAAAGTGTAAGTAATTTCGAAAAAATA-GATTCCACGTCTAGGTGAAATTCTACTTCTCACACGACCACT-GTTTTTGCCTTTATTAACTCACTTCTTGTT	*FMS1* KO repair oligo

13525	TGGCTCAATGATGCTAGTGGG	*FMS1* mutant allele amplification
13526	AGCCAAATTGCCAAGAAAGGG	*FMS1* mutant allele amplification
16601	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCGCGTGAACGCAAATGCATCGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC	2µm fragment for *SPE2* gRNA plasmid
16602	AATAGTATTTTTCAGCGAGAATCATATTGGATGAGTATCCACATGGCGTGAACGCAAATGTATCGTGATGAAATGATAAATCGGAGTCTTGGGCCGAGTTGACATATATTTCGTCAAG	*SPE2* mutation-carrying repair oligo
16603	CTTGACGAAATATATGTCAACTCGGCCCAAGACTCCGATTTATCATTTCATCACGATACATTTGCGTTCACGCCATGTGGATACTCATCCAATATGATTCTCGCTGAAAAATACTATT	*SPE2* mutation-carrying repair oligo
12174	GCATCGTCTCATCGGTCTCATATGTCTGCTACCTCTACTGCTACTTCC	YTK-compatible end addition to *THI4* CDS
12175	ATGCCGTCTCAGGTCTCAGGATCTAAGCAGCAAAGTGTTTCAAAATTTG	YTK-compatible end addition to *THI4* CDS
12985	TGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCAAACATTCAAATATATTCCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAG	2µm fragment for YPRcTau3 gRNA plasmid
13261	AATACGAGGCGAATGTCTAGG	*THI4* integration check
13262	GCCTCCCCTAGCTGAACAAC	*THI4* integration check
13492	TACAGCTCGCTCCTTGCATC	*SPE2* mutation check
13493	GCTTGCTTGGAGGGCTTTTC	*SPE2* mutation check

DNA sequencing

Genomic DNA of strains IMS0721, IMS0722, IMS0723, IMS0724, IMS0725, IMS0726, IMS0727, IMS0728, IMS0729, IMS0730, IMS0731, IMS0732, IMS0733, IMS0734, IMS0735, IMS0736, IMS0737, IMS0738, IMS0747, IMS0748, IMS0749, IMX2128, IMX2135, IMX2136, and IMX2137 was isolated with a Blood & Cell Culture DNA Kit with 100/G Genomics-tips (QIAGEN, Hilden, Germany) according to the manufacturers' protocol. Illumina-based paired-end sequencing with 150-bp reads was performed on 300-bp insert libraries Novogene (Novogene (HK) Company Limited, Hong Kong) with a minimum resulting coverage of 50x. Data mapping was performed against the CEN.PK113-7D genome [225] where an extra chromosome containing the relative integration cassette was previously added. Data processing and chromosome copy number variation determinations were done as previously described [309, 318].

Plasmids cloning

Plasmids carrying two copies of the same gRNA were cloned by *in vitro* Gibson assembly as previously described [319]. In brief, an oligo carrying the 20 bp target sequence and homology to the backbone plasmid was used to amplify the fragment carrying the 2µm origin of replication sequence by using pROS13 as template. The backbone linear fragment was amplified by using primer 6005 and either pROS12 or pROS13 as template [313]. The two fragments were then gel purified, combined and assembled *in vitro* using the NEBuilder HiFi DNA Assembly Master Mix (New England BioLabs, Ipswich, MA) following manufacturer's instructions. Transformants were selected on LB plates supplemented with 100 mg/L ampicillin.

Primers 13520, 13521, 13522, 13686, 13518, 14229, 14271, 14272, 14848, 15037, 15728, 12985, 16598, 16601 were used to amplify the 2µm fragments targeting *CNB1*, *PMR1*, *FRE2*, *ABZ1*, *GAL11*, SynPAM, *TUP1*, *ISW2*, *BAS1*, hphNT1, YPRcTau3, *FMS1*, and *SPE2,* respectively. The fragment targeting *GAL11* was cloned in a pROS12 backbone yielding plasmid pUDR441. The fragment targeting *CNB1*, *PMR1*, *FRE2*, *ABZ1*, SynPAM, *TUP1*, *ISW2*, *BAS1*, hphNT1, YPRcTau3, *FMS1*, and *SPE2* were cloned in a pROS13 backbone yielding plasmids pUDR388, pUDR389, pUDR390, pUDR438, pUDR471, pUDR472, pUDR473, pUDR566, pUDR650, pUDR571, pUDR514, pUDR652, and pUDR651, respectively.

The plasmid carrying the expression cassette for *THI4* was cloned by golden gate assembly using the yeast toolkit parts [317]. The *THI4* coding sequence was amplified using the primer pair 12174/12175 and CEN.PK113-7D genomic DNA as a template in order to add YTK compatible ends to the gene. The PCR product was then purified and combined together with plasmids pYTK009, pYTK056, and pYTK096 in a BsaI golden gate reaction that yielded plasmid pUDI180.

Strain construction

Strains carrying the target mutations were all constructed starting from IMX585 expressing the Cas9 protein [313]. For all strain except for IMX2290, IMX2291, IMX2289 and IMX2308, a two-steps strategy was adopted where first the target gene to be mutated was removed and replaced with a synthetic and unique 20 bp target sequence + 3 bp PAM sequence (SynPAM) and then, the synthetic target sequence was targeted and replaced with the mutant gene. In the second step where the SynPAM sequence was targeted, the mutant gene flanked by about 400 bp upstream and downstream sequences was amplified by using the evolved strain genomic DNA as template. The PCR product was then gel purified and used as repair-fragment in the transformation. This strategy yielded both intermediate strains lacking the targeted gene and final strains carrying the desired mutant gene.

In the first step, IMX585 was targeted at the gene of interest by transforming the strain with the relative pUDR plasmid. The double-strand break was then repaired by co-transforming the strain with two complementary DNA oligos carrying the SynPAM sequence flanked by 60 bp homology sequences to the targeted *locus* that were previously combined at 1:1 molar ratio, boiled for 5 minutes and annealed by cooling down the solution at room temperature on the bench.

500 ng of annealed primers pair 13612/13613, 13541/13542, 13539/13540, 14988/14989, 15075/15076, 13533/13534, 13531/13532, 13535/13536, 13527/13528 were co-transformed with 500 ng pUDR388, pUDR389, pUDR390, pUDR438, pUDR412, pUDR441, pUDR472, pUDR473, pUDR652 respectively yielding IMX1721, IMX1722, IMX1723, IMX1988, IMX1820, IMX1819, IMX1817, IMX1818, IMX2292 respectively. IMX1819 and IMX1820 transformants were selected on YPD plates with 200 mg/L hygromycin while IMX1721, IMX1722, IMX1723, IMX1988, IMX1817, IMX1818, and IMX2292 transformants were selected on YPD plates with 200 mg/L G418.

The *BAS1* knock-out strain could not be obtained with the marker-free SynPAM strategy. Therefore, the hphNT1 marker cassette was amplified by using primers 15584/15585 to add 60 bp homology flanks and pROS12 as a template. The PCR fragment was then gel purified and 500 ng were co-transformed with 500 ng pUDR592 to yield IMX2128. Transformants were selected on YPD plates with 200 mg/L G418 and 200 mg/L hygromycin.

In the second step, the SynPAM target sequence in each knockout strain was targeted for the insertion of the mutant allele. The mutant gene flanked by about 400 bp upstream and downstream sequences was amplified using the evolved strain genomic DNA as template. The PCR product was then gel purified and 500 ng were co-transformed with 500 ng of pUDR471. Primer pairs 13523/13524, 11292/11293, 13525/13526, 12052/12053, 11725/11726, 13498/13499, 13498/13499, 15077/15078, 15077/15078, 13496/13497, 13527/13528 were used to amplify the mutant alleles of *CNB1*^{L82F}, *PMR1*^{S104Y}, *FRE2*^{T110S}, *ARO7*^{L205S}, *ABZ1*^{R593H}, *GAL11*^{I541N}, *GAL11*^{Q383Stop}, *TUP1*^{V374A}, *TUP1*^{Q99Stop}, *ISW2*^{S181Stop}, and *FMS1*^{Q33K}, respectively using IMS0747, IMS0748, IMS0748, IMS0728, IMS0727, IMS0734, IMS0735, IMS0734, IMS0735, IMS0733, IMS0736, IMS0735 genomic DNA as template, respectively. Transformants were selected on YPD plates with 200 mg/L G418, yielding IMX1985, IMX1986, IMX1987, IMX1989, IMX2057, IMX1991, IMX1992, IMX2002, IMX2003, IMX1990, and IMX2292, respectively. The *BAS1*^{Q152R}, *BAS1*^{D101N}, *BAS1*^{S41P} mutant alleles were amplified from IMS737, IMS738, and IMS739 genomic DNA respectively using the primer pair 13687/13688. After gel purification, 500 ng of each PCR product was co-transformed in IMX2128, together with the hphNT1 targeting plasmid pUDR650, yielding IMX2135, IMX2136, and IMX2137, respectively. The strain IMX2289 carrying the *SPE2*^{A278T} mutant allele was constructed by transforming IMX585 with the *SPE2* targeting plasmid pUDR651 together with the annealed primer pair 16602/16603

containing the desired single base change plus a synonymous mutation causing the removal of the PAM sequence. After transformation, strains IMX2135, IMX2136, IMX2137, and IMX2289 were plated on YPD plates with 200 mg/L G418 for selection.

Mutant alleles found in the same evolved strains were combined in a single strain by repeating the strategy described above but this time using a mutant strain as a starting point instead of IMX585. In this way, *GAL11*, *TUP1*, and *FMS1* were deleted in IMX2002, IMX2003, and IMX2127 respectively by co-transforming the relative gRNA plasmid and the relative dsDNA oligo pair as done for the single knock out strains, yielding the intermediate strains IMX2066, IMX2110, and IMX2294 respectively. Then, the SynPAM sequence was targeted in IMX2066, IMX2110, and IMX2294 as previously described for the single mutant strains, yielding IMX2087, IMX2127, and IMX2307 respectively. IMX2043 carrying the *PMR1*^{S104Y}-*FRE2*^{T110S} double mutation was constructed by co-transforming IMX1987 with pUDR390 and the linear fragment containing the *FRE2*^{T110S} mutant allele that was previously amplified as described above. The *SPE2*^{A278T} mutant allele was combined with the *GAL11*^{I541N} *TUP1*^{V374A} mutant alleles present in IMX2127 by co-transforming the strain with the *SPE2* targeting plasmid pUDR651 together with the annealed primer pair 16602/16603, yielding IMX2308. The *THI4* overexpression cassette was amplified by using pUDI180 as a template and primers 12174/12175. 500ng of gel-purified PCR product was co-transformed together with the YPRcTau3 targeting plasmid pUDR514 in IMX585 and IMX1985 yielding IMX2290 and IMX2291 respectively.

To verify the correct gene editing, single colonies were picked from each transformation plate and genomic DNA was extracted as previously described [314]. The targeted *locus* was amplified by PCR and run on a 1% agarose gel. Primers pair 13523/13524, 13541/13542, 13539/13540, 15077/15078, 13496/13497, 13498/13499, 12052/12053, 13523/13524, 13541/13542, 13539/13540, 13693/13694, 12052/12053, 13496/13497, 13498/13499, 13498/13499, 15077/15078, 15077/15078, 13524/13525, 13693/13694, 13498/13499, 15077/15078, 15077/15078, 13687/13688, 13498/13499, 13687/13688, 13687/13688, 13687/13688, 13261/13262, 13261/13262, 13492/13493, 13525/13526, 13525/13526, 13492/13493, 13525/13526 were used to verify the correct gene editing in IMX1721, IMX1722, IMX1723, IMX1817, IMX1818, IMX1819, IMX1920, IMX1985, IMX1986, IMX1987, IMX1988, IMX1989, IMX1990, IMX1991, IMX1992, IMX2002, IMX2003, IMX2043, IMX2057, IMX2066, IMX2110, IMX2127, IMX2128, IMX2087, IMX2135, IMX2136, IMX2137 IMX2290, IMX2291, IMX2289, IMX2292, IMX2306, IMX2308, and IMX2307 respectively. To verify the presence if the single point mutations, each PCR product was purified and Sanger sequenced (Baseclear, The Netherlands). Mutations in *BAS1* could not be verified by Sanger sequencing and therefore whole-genome re-sequencing of IMX2135, IMX2136, IMX2137 was performed as explained above for the evolved single colony isolates.

After genotyping of the transformants, correct isolates were grown in 20 ml YPD in a 50 ml vented Greiner tube at 30 °C overnight by inoculating a single colony. The next day, 1 µl was transferred to a new tube containing the same amount of medium and the sample was grown overnight. The day after, each liquid culture was restreaked to single colony by plating on YPD agar plates. Plates were incubated at 30 °C overnight and the next day single colonies were patched on both YPD and YPD plus the relative antibiotic (either G428 or hygromycin) to assess which clones have lost the gRNA plasmid. One clone for each strain that had lost the plasmid was then grown in YPD and 30 %v/v glycerol was added prior to stocking samples at -80 °C.

Data availability

The sequencing data of the evolved and of the *BAS1* deletion *Saccharomyces cerevisiae* strains were deposited at NCBI (https://www.ncbi.nlm.nih.gov/) under BioProject accession number PRJNA603441. All measurement data used to prepare the figures of the manuscript are available at the data.4TU.nl repository under the URL: https://doi.org/10.4121/uuid:53c9992f-d004-4d26-a3cd-789c524fe35c

Acknowledgments

Experiments were designed by TP, JMD and JP. Strain evolution and isolation was performed by TP. Analysis of next-generation sequencing data was performed by MvdB and TP. Reverse engineering of target mutations and phenotypical characterization of the strains was done by TP and DM. TP and JMD wrote the first version of manuscript. All authors critically read this version, provided input and approved the final version. This work has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 722287. This manuscript has also been deposited as a preprint to bioRxiv (https://www.biorxiv.org/) **doi:** https://doi.org/10.1101/2020.02.12.945287

Chapter 3:
Identifying oxygen-independent pathways for pyridine-nucleotide and Coenzyme A synthesis in anaerobic gut fungi by expression of candidate genes in yeast

Thomas Perli*, Aurin M. Vos*, Jonna Bouwknegt, Wijb J. C. Dekker, Sanne J. Wiersma, Christiaan Mooiman, Raúl A. Ortiz-Merino, Jean-Marc Daran, Jack T. Pronk

* These authors contributed equally to this work.

Essentially as published in
mBio
2021; 12 (3) e00967-21
https://doi.org/10.1128/mBio.00967-21
Supplementary material available online

Abstract

Neocallimastigomycetes are unique examples of strictly anaerobic eukaryotes. This study investigates how these anaerobic fungi bypass reactions involved in synthesis of pyridine nucleotide cofactors and coenzyme A that, in canonical fungal pathways, require molecular oxygen. Analysis of Neocallimastigomycete proteomes identified a candidate L-aspartate-decarboxylase (AdcA), and L-aspartate oxidase (NadB) and quinolinate synthase (NadA), constituting putative oxygen-independent bypasses for coenzyme A synthesis and pyridine nucleotide cofactor synthesis. The corresponding gene sequences indicated acquisition by ancient horizontal gene transfer (HGT) events involving bacterial donors. To test whether these enzymes suffice to bypass corresponding oxygen-requiring reactions, they were introduced into *fms1Δ* and *bna2Δ Sacharomyces cerevisiae* strains. Expression of *nadA* and *nadB* from *Piromyces finnis,* and *adcA* from *Neocallimastix californiae* conferred cofactor prototrophy under aerobic and anaerobic conditions. This study simulates how HGT can drive eukaryotic adaptation to anaerobiosis, and provides a basis for elimination of auxotrophic requirements in anaerobic industrial applications of yeasts and fungi.

Importance

Nicotinamide adenine dinucleotide (NAD^+) and Coenzyme A (CoA) are central metabolic cofactors, whose canonical biosynthesis pathways in fungi require oxygen. Anaerobic gut fungi of the Neocallimastigomycota phylum are unique eukaryotic organisms that adapted to anoxic environments. Analysis of Neocallimastigomycota genomes revealed that these fungi might have developed oxygen-independent biosynthetic pathways for NAD^+ and CoA biosynthesis, likely acquired through horizontal gene transfer (HGT) from prokaryotic donors. We confirmed functionality of these putative pathways under anaerobic conditions by heterologous expression in the yeast *Saccharomyces cerevisiae*. This approach, combined with sequence comparison, offers experimental insight on whether HGT events were required and/or sufficient for acquiring new traits. Moreover, our results demonstrate an engineering strategy for enabling *S. cerevisiae* to grow anaerobically in the absence of the precursor molecules pantothenate and nicotinate, thereby contributing to alleviate oxygen requirements and to move closer to prototrophic anaerobic growth of this industrially relevant yeast.

Introduction

Neocallimastigomycetes are obligately anaerobic fungi with specialised metabolic adaptations that allow them to play a key role in the degradation of recalcitrant plant biomass in herbivore guts [320]. Despite complicated cultivation techniques and lack of genetic-modification tools [321], several evolutionary adaptations of these eukaryotes to an anaerobic lifestyle have been inferred from biochemical studies [322-324]. Sequence analysis implicated extensive horizontal gene transfer (HGT) as a key mechanism in these adaptations [325-327]. For example, instead of sterols, which occur in membranes of virtually all other eukaryotes [328] and whose biosynthesis involve multiple oxygen-dependent reactions [329], Neocallimastigomycetes contain tetrahymanol [322, 325]. This sterol surrogate [330] can be formed from squalene by a squalene:tetrahymanol cyclase (STC), whose structural gene in Neocallimastigomycetes showed evidence of acquisition by HGT from prokaryotes [325, 331]. Expression of an STC gene was recently shown to enable sterol-independent anaerobic growth of the model eukaryote *Saccharomyces cerevisiae* [332].

Further exploration of oxygen-independent bypasses in Neocallimastigomycetes for intracellular reactions that in other eukaryotes require oxygen is relevant for a fundamental understanding of the requirements for anaerobic growth of eukaryotes. In addition, it may contribute to the elimination of nutritional requirements in industrial anaerobic applications of yeasts and fungi.

Most fungi are capable of *de novo* synthesis of pyridine-nucleotide cofactors (NAD^+ and $NADP^+$) and Coenzyme A (CoA) when grown aerobically. As exemplified by the facultatively anaerobic yeast *S. cerevisiae* [54], canonical fungal pathways for synthesis of these cofactors are oxygen dependent. In *S. cerevisiae,* biosynthesis of CoA involves formation of β-alanine by the oxygen-requiring polyamine oxidase Fms1 [333]. This intermediate is then condensed with pantoate to yield the CoA precursor pantothenate [162, 334] (Fig. 1, left). Similarly, the yeast kynurenine pathway for *de novo* synthesis of NAD^+ involves three oxygen-dependent reactions, catalyzed by indoleamine 2,3-dioxygenase (Bna2; EC 1.13.11.52), kynurenine 3-monooxygenase (Bna4; EC 1.14.13.9), and 3-hydroxyanthranilic-acid dioxygenase (Bna1; EC 1.13.11.6) [54] (Fig. 1, right). The Neocallimastigomycete *Neocallimastix patricianum* has been shown to grow in synthetic media lacking precursors for pyridine-nucleotide and CoA synthesis [335]. This observation indicates that at least some anaerobic fungi harbour oxygen-independent pathways for synthesizing these essential cofactors. Genomes of Neocallimastigomycetes lack clear homologs of genes encoding the oxygen-requiring enzymes of the kynurenine pathway. Instead, their genomes were reported to harbour genes encoding an L-aspartate oxidase (NadB) and quinolinate synthase (NadA), two enzymes active in the bacterial pathway for NAD^+ synthesis [325] (Fig. 1, right).

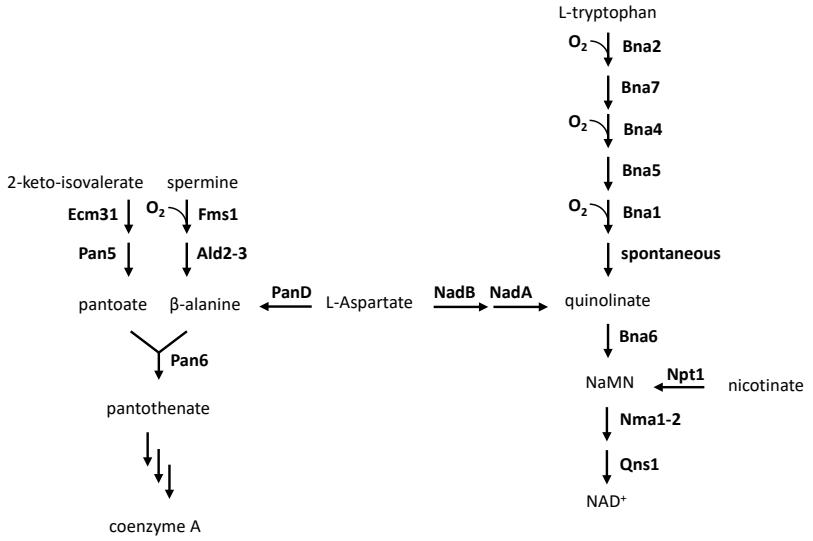


Fig. 1: CoA and NAD⁺ biosynthetic pathways in *S. cerevisiae* and oxygen-independent alternatives. CoA synthesis includes the condensation of pantoate and β-alanine. (Left) In *S. cerevisiae* β-alanine is formed from spermine in two steps using the oxygen-dependent poly-amine oxidase Fms1 . Other organisms, including Archaea, Bacteria, and insects, can by-pass this oxygen requirement by synthesizing β-alanine from aspartate using L-aspartate decarboxylase (AdcA/PanD). (Right) NAD⁺ is synthesized via the kynurenine pathway in 9 reactions starting from tryptophan, 3 of which require oxygen. Other organisms that include plants and bacteria are able to bypass this oxygen requirement by synthesizing quinolinate from aspartate using l-aspartate oxidase and quinolinate synthase (NadB and NadA, respectively).

Since bacterial and plant aspartate oxidases can, in addition to oxygen, also use fumarate as electron acceptor [336, 337], it is conceivable that NadA and NadB may allow for oxygen-independent NAD⁺ synthesis in anaerobic fungi. No hypothesis has yet been forwarded on how these fungi may bypass the oxygen requirement for the canonical fungal CoA biosynthesis route.

The goals of this study were to identify the pathway responsible for oxygen-independent synthesis of CoA in Neocallimastigomycetes and to investigate a possible role of NadA and NadB in oxygen-independent synthesis of pyridine-nucleotide cofactors. A candidate L-aspartate decarboxylase (Adc) encoding gene was identified by genome analysis of Neocallimastigomycetes and its phylogeny investigated. Candidate Neocallimastigomycete genes for L-aspartate oxidase and quinolinate synthase, previously reported to have been acquired by HGT [325], as well as the candidate Adc gene, were then functionally analysed by expression in *S. cerevisiae* strains devoid of essential steps in the native cofactor synthesis pathways. As controls, previously characterized genes involved in

oxygen-independent NAD^+ biosynthesis by *Arabidopsis thaliana* [338], and a previously characterized Adc encoding gene from the red flour beetle *Tribolium castaneum* (*TcPAND*) [339] were also expressed in the same *S. cerevisiae* strains. The results demonstrate how heterologous expression studies in yeast can provide insight into evolutionary adaptations to anaerobic growth and selective advantages conferred by proposed HGT events in Neocallimastigomycetes. In addition, they identify metabolic engineering strategies for eliminating oxygen requirements for cofactor biosynthesis in anaerobic industrial applications of *S. cerevisiae*.

Results

Identification of a candidate oxygen-independent L-aspartate decarboxylase involved in CoA synthesis in anaerobic fungi

Decarboxylation of L-aspartate to β-alanine by L-aspartate decarboxylase (Adc), an enzyme that occurs in many species across all domains of life [340], enables an oxygen-independent alternative for the canonical fungal pathway for CoA synthesis (Fig. 1). To explore its occurrence in anaerobic fungi, a set of 51 amino acid sequences of Adc homologs listed by Tomita *et al.* [340] were used as queries against all proteins from 5 Neocallimastigomycete species deposited in the TrEMBL section of the UNIPROT database. This search yielded 16 Neocallimastigomycete hits (e-value < 10^{-6}, Supplementary Table S1), 6 of which originated from *N. californiae*. Only one of these hits, A0A1Y1ZL74, did not reveal annotation errors upon RNAseq read mapping, showed the highest read coverage (Supplementary Fig. S1), and was selected as best Neocallimastigomycete Adc candidate.

The amino acid sequence A0A1Y1ZL74 (hereafter refered to as *Nc*AdcA) was used for a second round of homology search to obtain a broad set of Adc-like sequences, with a similar sequence representation of taxa across the 3 domains of life (104 sequences from Bacteria, 101 from Eukarya, and 120 from Archaea; Dataset S1). The complete set of *Nc*AdcA homologs (together with the set defined by Tomita *et al.* (2015) [340] and their Neocallimastigomycete homologs; Dataset S2) were subjected to multiple sequence alignment. A subsequent phylogenetic tree (Fig. 2; Dataset S3) showed that *Nc*Adc sequences are closely related to those of chytrid fungi (*e.g.* A0A1S8W5A4 from *Batrachochytrium salamandrivorans*) and from anaerobic bacteria (*e.g.* B8I983 from *Clostridium cellulolyticum*, currently known as *Ruminiclostridium cellulolyticum* [341] we used the former name for consistency with Uniprot identifiers). These Neocallimastigomycete, chytrid, and bacterial Adc homologs were more closely related to each other than to characterized eukaryotic Adc and bacterial PanD sequences. Furthermore, HMMER e-values obtained from using *Nc*AdcA as query against the bacterial database were most significant than when using the eukaryotic or archaeal databases

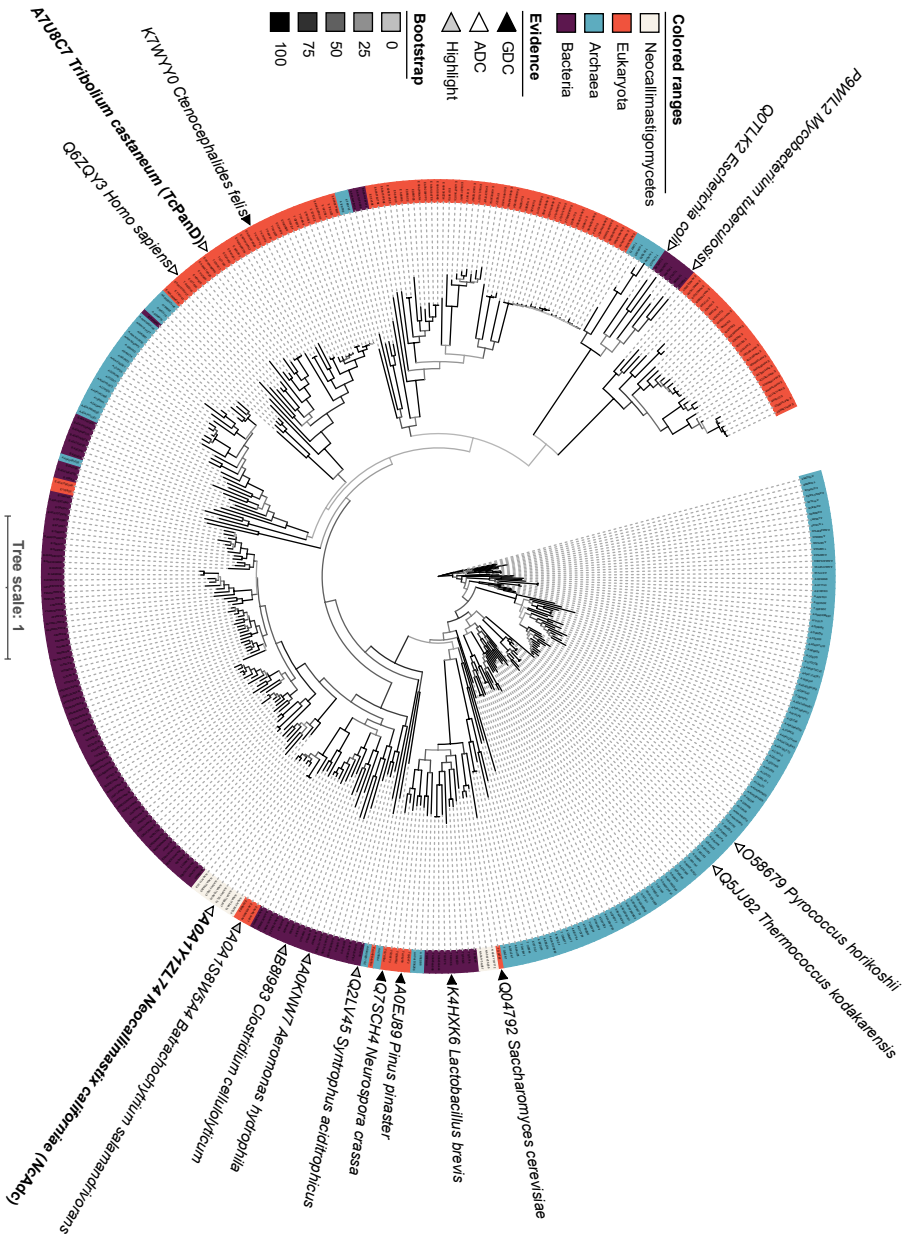


Fig. 2: Unrooted maximum likelihood phylogenetic tree of aspartate decarboxylase and glutamate decarboxylase homologs. Sequences of proteins with demonstrated enzyme activity are marked with white triangles (L-aspartate decarboxylases) or black triangles (glutamate decarboxylases). Interactive visualizations with all sequence identifiers, branch distances and bootstrap values, can be accessed in https://itol.embl.de/tree/838448017961605604402 and https://itol.embl.de/tree/838448047664161598532 3.

(Supplementary Fig. S2; Dataset S1). These results suggest that a bacterial ancestor donated an Adc-encoding sequence to a common ancestor of chytrids and Neocallimastigomycetes. To further investigate the potential bacterium-to-chytrid HGT event, a refined ortholog search and phylogenetic analysis were performed. Full proteomes of all species showing an *Nc*AdcA homolog, in addition to predicted proteomes from 6 chytrids used in a previous phylogenomic analysis [327], were retrieved and used to obtain all possible co-ortholog groups. From a total number of 103 *Nc*AdcA orthologs obtained, 85 were bacterial, 5 archaeal, and 13 eukaryotic (Table 1; Dataset S4). Eukaryotic *Nc*AdcA orthologs were only found in fungi, and 12 out of 13 were found in species from the Chytridiomycota phylum. The latter included 5 out of the 6 chytrids analysed in the phylogenomic study of Wang et al. [327], and all Neocallimastigomycetes considered in this study. Further phylogenetic analysis of the 103 *Nc*AdcA orthologs indicated a common origin for bacterial and chytrid *Nc*AdcA (Figure 3; Dataset S5). The closest bacterial relatives to *Nc*AdcA were found in the facultative anaerobe and waterborne bacterium *Aeromonas hydrophila* subsp. *hydrophila* ATCC 7966^{T}[342], and the ruminal anaerobe *C. cellulolyticum* strain H10 [341, 343]. Additional close bacterial relatives were also strict anaerobes, such as the syntrophic bacterium *Syntrophus aciditrophicus* [344] and members of the Desulfobacteraceae family [345].

The Adc bacterium-to-chytrid HGT event was futher confirmed by using Abaccus, an automated phylogeny-aware and topology-based algorithm [346]. Abaccus uses the topology of a given tree to determine taxonomic level 'jumps' (J) and 'loses' (L) between a seed sequence (*Nc*AdcA) and every other node in the tree. The tree of *Nc*AdcA orthologs resulted in J=4 and L=3, meaning that the node comprising *Nc*AdcA 'jumps' 4 taxonomic levels which could only be explained by complete loses in 3 of these taxonomic levels. These J and L values obtained for the tree of *Nc*AdcA orthologs are higher than Abaccus' default HGT cutoff values ($J \geq 2$ and $L \geq 3$), and are independent of the evolutionary model used to infer the tree (PROTGTR [347, 348], JTT [349], and LG [350]).

Comparison of bacterial PanDs (Q0TLK2 from *E. coli* and P9WIL2 from *Mycobacterium tuberculosis*) against Adcs from other bacteria (B8I983 from *C. cellulolyticum*), and eukaryotes (including A7U8C7 from *Tribolium castaneum*) showed only little sequence homology between *Nc*Adc*s,* known bacterial PanDs, and eukaryotic Adcs (Dataset S6). The only conserved region encompassed the full length of PanDs (126-139 amino acids), which represents less than 60 % of the full length of other Adc sequences (*e.g. Nc*Adc*A* is 625 amino acids long). These sequence comparisons, together with the intron-exon structures verified with RNAseq data (Supplementary Fig. S1) show *NcadcA* has acquired eukaryotic features while retaining homology to its bacterial ancestor, as is typical for prokaryotic genes acquired by fungal genomes [351].

Table 1: Summary of *Nc*Adc homology search results across domains of life.

Taxonomic rank			Species analyzed	Homologs	Orthologs
Eukarya			749	101	13
Fungi			404	48	13
	Dikarya		372	36	1
		Ascomycota	280	36	1
		Basidiomycota	92	0	0
	Fungi *incertae sedis*		32	12	12
		Blastocladiomycota	0	0	0
		Chytridiomycota	11	12	12
		Cryptomycota	1	0	0
		Microsporidia	7	0	0
		Mucoromycota	11	0	0
		Zoopagomycota	2	0	0
Bacteria			1807	101	85
Archaea			765	104	5

Neocallimastigomycete *PfnadB*, *PfnadA* and *NcadcA* genes support aerobic pyridine-nucleotide and CoA synthesis in yeast.

Neocallimastigomycetes were previously reported to have acquired an L-aspartate oxidase (*nadB*) and a quinolinate synthase gene (*nadA*) by HGT [325]. Hence, UNIPROT entries A0A1Y1V2P1 and A0A1Y1VAT1 from *Piromyces finnis* were functionally reassigned as NadA and NadB candidates and the corresponding genes were tentatively named *PfnadB* and *PfnadA*. These sequences, together with *NcadcA,* were codon-optimised and tested to bypass the corresponding oxygen-requiring reactions in *S. cerevisiae*.

The *BNA2* and *FMS1* genes of *S. cerevisiae* were deleted by Cas9-mediated genome editing. The inability of strain IMK877 (*bna2Δ*) to synthesize quinolinic acid and of strain IMX2292 (*fms1Δ*) to synthesize β-alanine was evident from their inability to grow on glucose synthetic medium (SMD) lacking nicotinic acid or pantothenate, respectively (Table 2). Strain IMK877 was used for heterologous complementation studies with codon-optimized expression cassettes for *PfnadB* and *PfnadA,* while an expression cassette for *N. californiae NcadcA* (A0A1Y1ZL74) was introduced into strain IMX2292. Congenic strains expressing previously characterized *NADB* and *NADA* genes from *Arabidopsis thaliana* (*At*NadB and *At*NadA, Q94AY1 and Q9FGS4) [338], and a previously characterized gene from *Tribolium castaneum* encoding an aspartate decarboxylase (*Tc*PanD, A7U8C7) [339] were tested in parallel.

Aerobic growth of the engineered *S. cerevisiae* strains was characterized in shake-flask cultures on SMD or on either SMDΔnic or SMDΔpan (Table 2). In contrast to the reference strain IMK877 (*bna2Δ*), *S. cerevisiae* IMX2301 (*bna2Δ PfnadB PfnadA*) grew in SMDΔnic, indicating complementation of the *bna2Δ*-induced nicotinate auxotrophy by *PfnadB* and *PfnadA*. However, the specific growth rate of the engineered strain in these

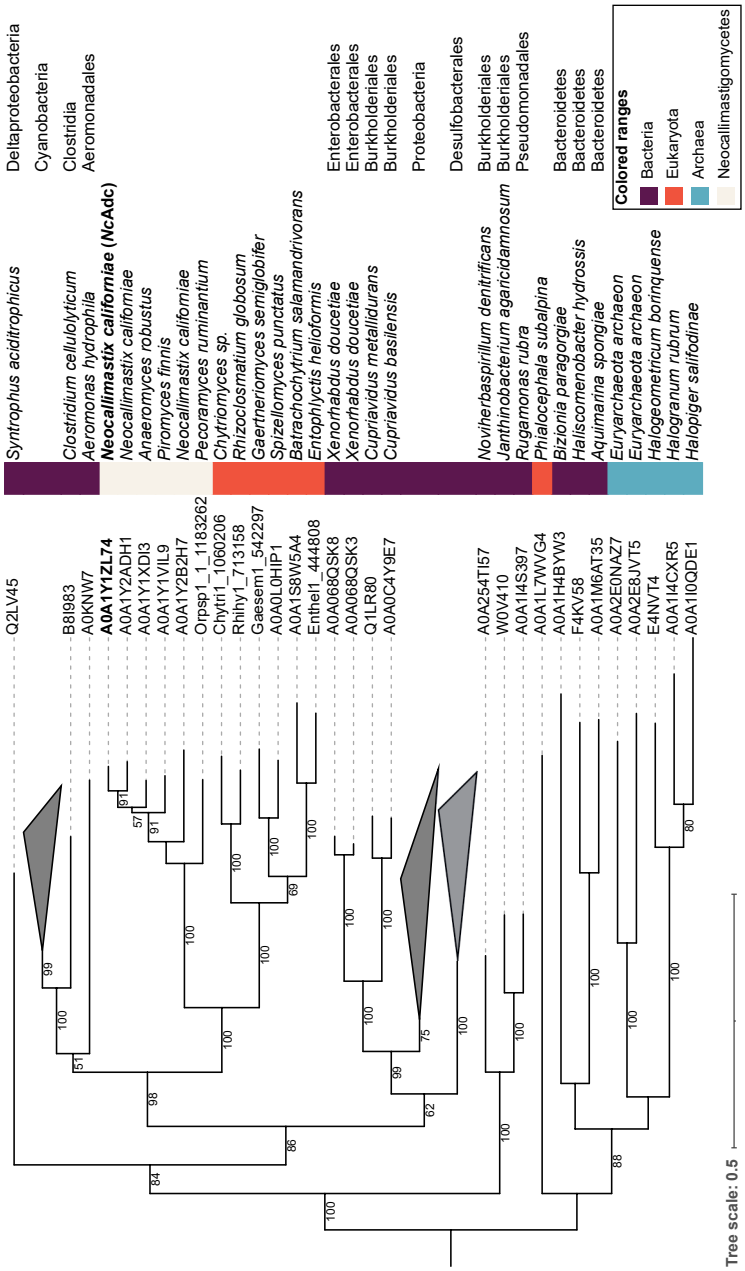


Fig. 3: Mid-rooted maximum likelihood phylogenetic tree of L-aspartate decarboxylase orthologs. Number of sequences in collapsed clades are indicated in parentheses. A summary of the search from which these sequences were obtained is presented in Table 1. An interactive visualization with all sequence identifiers, branch support, distances and bootstrap values, can be accessed in https://itol.embl.de/tree/838448026719161528015 2.

aerobic cultures was approximately 3-fold lower than that of the reference strain IMX585 (*BNA2*, Table 2). Strain IMX2302 (*bna2Δ AtNADB AtNADA*) did not grow in SMDΔnic, suggesting that the plant NadB and/or NadA proteins were either not functionally expressed or not able to complement the nicotinate auxotrophy in these aerobic yeast cultures.

Strain IMX2300 (*fms1Δ NcadcA*) grew in SMDΔpan, indicating complementation of the panthotenate auxotrophy. However, this strain reproducibly showed a lag phase of approximately 48 h upon its first transfer from SMD to SMDΔpan, and grew exponentially thereafter at a rate of 0.34 ± 0.01h^{-1}. To explore whether the lag phase of strain IMX2300 reflected selection of a spontaneous mutant, it was subjected to three sequential transfers in SMDΔpan. A single-colony isolate, IMX2300-1 from the adapted population showed a specific growth rate of 0.34 ±0.01 h^{-1} in both SMD and SMDΔpan (Table 2). Whole-genome sequencing of IMX2300-1 did not reveal any mutations in coding DNA sequences that were considered physiologically relevant in this context when compared to the non-adapted strain IMX2300 (Supplementary File 4, Bioproject accession number: PRJNA634013). This observation indicated that the lag phase of strain IMX2300 most likely reflected a physiological adaptation or culture heterogeneity rather than a mutational event [352].

The specific growth rate of *S. cerevisiae* IMX2305 (*fms1Δ TcPAND*) on SMDΔpan did not significantly differ from that of the reference strain IMX585 on SMD, and it was almost four-fold higher than the specific growth rate of the reference strain on SMDΔpan. These results are consistent with a previous study on functional expression of *TcPAND* in *S. cerevisiae* [353].

Table 2: Aerobic characterization of engineered strains. Specific growth rates of *S. cerevisiae* strains grown in SMD, SMDΔnic and SMDΔpan media. The values are average and mean deviation of data from at least two independent cultures of each strain.

Strain	SMD	SMDΔnic	SMDΔpan
IMX585 (*FMS1 BNA2*)	0.40 ± 0.01	0.40 ± 0.02	0.11 ± 0.01
IMX2292(*fms1Δ*)	0.39 ± 0.01		< 0.01
IMX2305 (*fms1Δ TcPAND*)	0.39 ± 0.01		0.39 ± 0.01
IMX2300-1 (*fms1Δ NcadcA*)	0.34 ± 0.01		0.34 ± 0.01
IMK877 (*bna2Δ*)	0.40 ± 0.01	< 0.01	
IMX2301 (*bna2Δ PfnadB PfnadA*)	0.37 ± 0.01	0.14 ± 0.01	
IMX2302 (*bna2Δ AtNADB AtNADA*)	0.40 ± 0.01	< 0.01	

Expression of Neocallimastigomycete *PfnadB*, *PfnadA*, and *NcadcA* suffice to enable anaerobic pyridine-nucleotide and CoA synthesis in yeast

To investigate whether expression of heterologous *PfnadB, PfnadA,* and *NcadcA* was sufficient to enable anaerobic growth in the absence of nicotinate and pantothenate, respectively, growth of the engineered *S. cerevisiae* strains on SMD, SMDΔnic and/or SMDΔpan was monitored in an anaerobic chamber (Fig. 4).

Growth experiments on SMDΔnic or SMDΔpan were preceded by a cultivation cycle on the same medium, supplemented with 50 g L^{-1} instead of 20 g L^{-1} of glucose to ensure complete depletion of any surplus cellular contents of pyridine nucleotides, CoA, or relevant intermediates. Indeed, upon a subsequent transfer to SMDΔnic or SMDΔpan, the reference strain IMX585 (*BNA2 FMS1*), expressing the native oxygen-dependent pathways for nicotinate and β-alanine synthesis, showed no growth (Fig. 3 panels A, B and C).

Both engineered strains IMX2301 (*bna2Δ PfnadB PfnadA*) and IMX2302 (*bna2Δ AtNADB AtNADA*) grew anaerobically on SMDΔnic. This provided a marked contrast with the aerobic growth studies on this medium, in which strain IMX2302 did not grow. Strains IMX2305 (*fms1Δ TcPAND*) and the aerobically pre-adapted IMX2300-1 (*fms1Δ NcadcA*) both grew on SMDΔpan under anaerobic conditions (Fig. 3 panels D, E and F).

Characterization of engineered yeast strains in anaerobic batch bioreactors

The anaerobic chamber experiments did not allow quantitative analysis of growth and product formation. Therefore, growth of the *S. cerevisiae* strains expressing the Neocallimastigomycetes genes, IMX2301 (*bna2Δ PfnadB PfnadA*) and IMX2300-1 (*fms1Δ NcadcA*) was studied in anaerobic bioreactor batch cultures on SMDΔnic or SMDΔpan and compared with growth of *S. cerevisiae* IMX585 (*BNA2 FMS1*) on the same media.

The reference strain IMX585, which typically grows fast and exponentially in anaerobic bioreactors when using complete SMD [354], exhibited extremely slow, linear growth on SMDΔnic and SMDΔpan (Fig. 5). Similar growth kinetics in 'anaerobic' bioreactor cultures of *S. cerevisiae* on synthetic medium lacking the anaerobic growth factors Tween 80 and ergosterol were previously attributed to slow leakage of oxygen into laboratory bioreactors [355-357].

In contrast to the reference strain IMX585, the engineered strains IMX2301 and IMX2300-1 exhibited exponential anaerobic growth on SMDΔnic and SMDΔpan, respectively (Fig. 4; Table 3). The specific growth rate of strain IMX2301 (*bna2Δ PfnadB PfnadA*) on SMDΔnic was not significantly different from that of the reference strain on complete SMD [354], indicating full complementation of the anaerobic nicotinate auxotrophy of *S. cerevisiae*. The specific growth rate of strain IMX2300-1 (*fms1Δ NcadcA*) on SMDΔpan was only 20 % lower than this benchmark (Table 3). Biomass and ethanol yields of strain

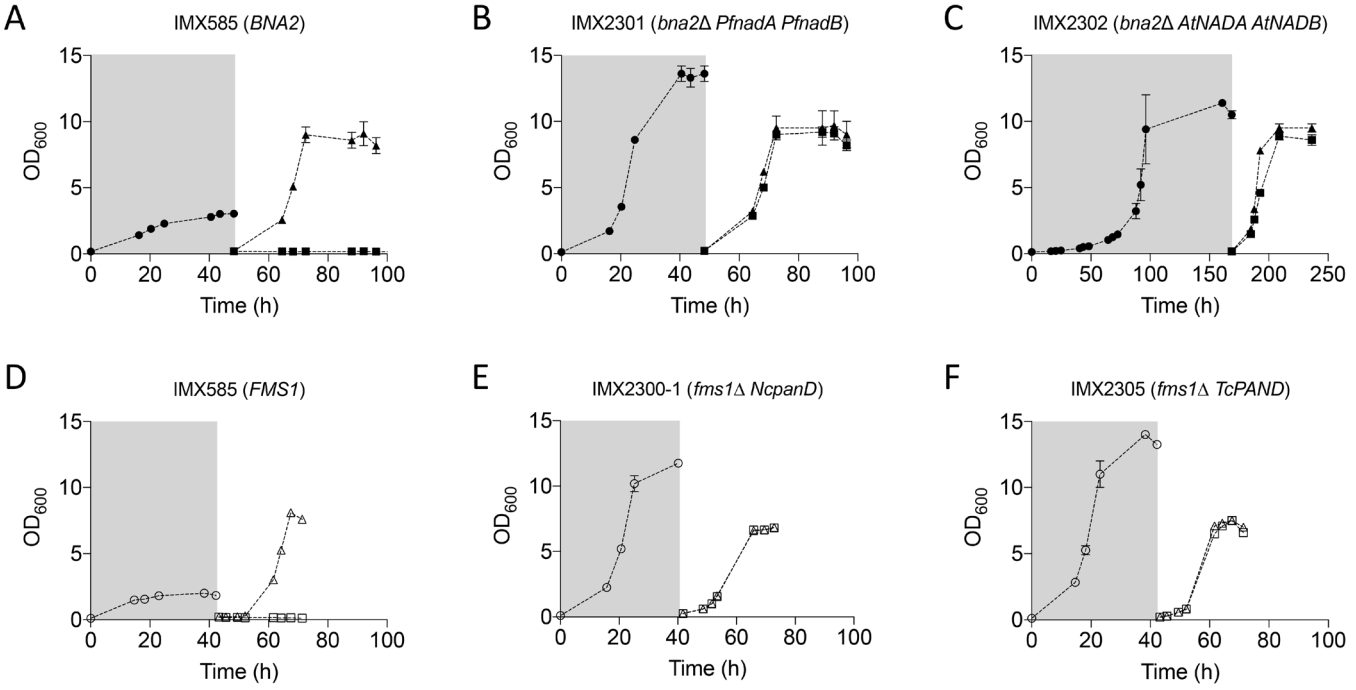


Fig. 4: Anaerobic growth of *S. cerevisiae* strains dependent or independent on supplementation of nicotinic acid (NA) or pantothenic acid (PA) in SMD medium containing Tween 80 and ergosterol. Strains IMX585 (A), IMX2301 (*bna2*Δ *PfnadB PfnadA*) (B), and IMX2302 (*bna2*Δ *AtNADB AtNADA*) (C) transferred to medium with 2 % glucose with (▲) or without (■) nicotinate after a carry-over phase in SMDΔnic containing 4 % glucose (● in grey box). Strains IMX585 (D), IMX2300-1 (*fms1*D *NcadcA*) (E), and IMX2305 (*fms1*D *TcPAND*)(F) transferred to medium with (△) or without (□) pantothenate after a carry-over phase in SMDΔpan containing 4 % glucose (○ in grey box). Anaerobic condition in the chamber were maintained using a palladium catalyst and a 5 % hydrogen concentration. Error bars represent the mean deviation of independent cultures (n=2).

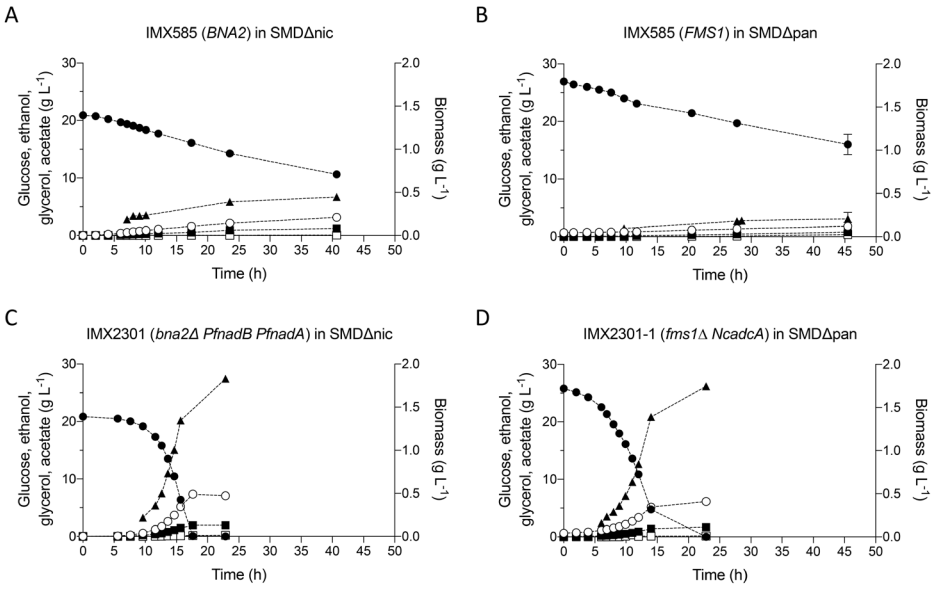


Fig. 5: Anaerobic batch cultivation of IMX585 in SMDΔnic (A) and SMDΔpan (B), IMX2301 in SMDΔnic (C) and IMX2300-1 in SMDΔpan (D). All strains were pre-grown in the corresponding medium lacking one vitamin prior to inoculation in the bioreactor to avoid carry-over effects. Values for glucose (●), ethanol (○), glycerol (▪), acetate (□) and biomass (▲) are shown over time. Error bars represent the mean deviation of independent cultures (n=2).

Table 3: Maximum specific growth rate (μ_{max}) and yields of glycerol, biomass and ethanol on glucose in anaerobic bioreactor batch cultures of *S. cerevisiae* strains IMX585, IMX2301 and IMX2300-1. Cultures were grown on SMD, SMDΔnic, or SMDΔpan, respectively, with 20 g L^{-1} glucose as carbon source (pH = 5). Growth rates and yields were calculated from the exponential growth phase. The ethanol yield was corrected for evaporation. Values represent average and mean deviation of data from independent cultures (n = 2). Carbon recovery in all fermentations was between 95 and 100%.

Strain	**IMX585*** **(FMS1 BNA2)**	**IMX2301 (bna2Δ PfnadB PfnadA)**	**IMX2300-1 (fms1Δ NcadcA)**
Medium	SMD	SMDΔnic	SMDΔpan
μ_{max} (h^{-1})	0.32 ± 0.00	0.31 ± 0.01	0.25 ± 0.00
Y glycerol/glucose (g g^{-1})	0.105 ± 0.000	0.103 ± 0.003	0.104 ± 0.000
Y biomass/glucose (g_x g^{-1})	0.094 ± 0.004	0.090 ± 0.002	0.081 ± 0.001
Y EtOH/glucose (g g^{-1})	0.372 ± 0.001	0.372 ± 0.002	0.364 ± 0.003

* data from [354]

IMX2301 grown in anaerobic batch cultures on SMDΔnic and strain IMX2300-1 grown on SMDΔpan were not significantly different from those of the reference strain IMX585 grown on complete SMD (p-value > 0.05, Table 3).

Discussion

This study shows how oxygen-independent panthotenate and nicotinate prototrophy can be conferred to the facultatively anaerobic yeast *S. cerevisiae* by heterologous expression of *NcadcA, PfnadB* and *PfnadA* genes from Neocallimastigomycetes, as well as corresponding orthologs from other species (*TcPAND, AtNADB* and *AtNADA*). These results also provide insights into how acquisition of these genes by HGT conferred selective advantage to Neocallimastigomycete ancestors under anaerobic conditions.

Results from phylogenetic analysis of Adc sequences (Fig. 2) were consistent with an earlier report on multiple evolutionary origins and variable evolutionary rates of pyridoxal-5'-phosphate-dependent enzymes, including Adc and glutamate decarboxylases [358, 359]. A separate clade of Neocallimastigomycete sequences show homology with characterised glutamate decarboxylases (e.g. Q04792 from *S. cerevisiae* and K4HXK6 from *Lactobacillus brevis*; Fig 2. These results further support acquisition of an Adc encoding DNA sequence by HGT rather than by neofunctionalization of a glutamate decarboxylase gene.

The characterised *Nc*AdcA (A0A1Y1ZL74) yielded the highest homology with orthologous sequences from chytrid fungi and anaerobic bacteria. This observation is in agreement with previous research showing that HGT events played a major role in shaping the genomes of Neocallimastigomycetes [323, 325, 326], with Firmicutes and Proteobacteria as prominent sequence donors [325]. Specifically, closer bacterial orthologs to *Nc*AdcA were found in genome sequences of *A. hydrophila* (Proteobacteria) and *C. cellulolyticum* (Firmicutes). These bacterial species are anaerobic and, considering their ecological niches (waterborne and decayed grass / ruminal fluid, respectively [341-343]) agrees with current hypotheses of these types of bacteria donating genes to anaerobic gut fungi and subsequently driving a mammalian transition to herbivory [325, 327]. Since *Nc*AdcA orthologs were found in 5 out of the 6 chytrids analysed, the Adc HGT transfer event appears to have preceded the 66 (±10) My-ago estimate for divergence of Neocallimastigomycetes from other chytrids [327], although this estimate may be contended by more recent phylogenomic analyses for the whole fungal kingdom [360].

*Nc*AdcA orthology and phylogenetic analyses revealed *Phialocephala subalpina* as the only other non-chytrid, non-Neocallimastigomycete eukaryote to have a separate Adc-like protein. This fungus is a root endophyte and was previously proposed have obtained multiple genes by HGT from bacterial donors [361]. However, A0A1L7WVG4 (PAC_06602), here identified as Adc ortholog, was not among the 21 genes of *P. subalpina* listed as likely acquired by HGT from non-fungal species. Since the phylogenetic placement

of the putative *P. subalpina* Adc was close to bacterial as well as archaeal sequences, further studies are needed to reveal its evolutionary history.

Whereas an alternative to the kynurenine pathway for NAD^+ synthesis was previously inferred from genome sequence analysis, the pathway by which Neocallimastigomycetes synthesize Coenzyme A had not previously been explored. Six pathways for synthesis of the essential CoA precursor β-alanine are known: (A) decarboxylation of L-aspartate [362], (B) transamination of malonate semialdehyde with L-glutamate as aminodonor [363] or L-alanine [364], (C) by reduction of uracil followed by hydrolysis of the resulting dihydrouracil [365], (D) oxidative cleavage of spermine to 3-aminopropanal followed by oxidation of the aldehyde group [162], (E) 2,3-aminomutase of alanine [366], and (F) addition of ammonia to acryloyl-CoA, followed by hydrolysis of the resulting CoA thioester [366]. Of these pathways, canal but option (D), in principle, can occur in the absence of oxygen. Yeasts and other filamentous fungi typically form β-alanine from spermine (pathway D), but in some species the use of pathway C was also reported [367]. While the aspartate decarboxylation route (A) has not previously been demonstrated in wild-type fungi, functional expression of bacterial and *T. castaneum Tc*PanD was used in metabolic engineering of *S. cerevisiae* to boost supply of β-alanine as a precursor for 3-hydroxypropionate production [339, 353]. Wild-type *S. cerevisiae* strains cannot grow in anaerobic environments unless supplemented with pantothenate. Expression of either *NcadcA* or *TcPAND* in an *fms1Δ S. cerevisiae* strain, which lacks the native oxygen-dependent pantothenate biosynthesis pathway, enabled growth in panthothenate-free medium under aerobic and anaerobic conditions. Although the different specific growth rates of *S. cerevisiae* strains expressing *NcadcA* or *TcPAND* indicate that changing expression levels and/or origin of ADC encoding genes may be required to achieve optimal growth, these results provide a proof of principle for a simple metabolic engineering strategy to eliminate oxygen requirements for pantothenate synthesis.

Genomic analyses previously suggested that genomes of Neocallimastigomycetes encode a putative L-aspartate oxidase (NadB) and quinolinate synthase (NadA) as alternatives to the canonical kynurenine pathway found in other fungi [325]. Additionally, Neocallimastigomycetes appear to have acquired both *nadB* and *nadA* through HGT [325]. Until now, functionality of these Neocallimastigomycete proteins in an oxygen-independent pathway for synthesis of quinolinate from L-aspartate had not been demonstrated.

Our results demonstrate that expression of *nadB* and *nadA* homologs, either from the Neocallimastigomycete *P. finnis* or from the plant *A. thaliana* [338], suffice to allow anaerobic synthesis of NAD^+ of *S. cerevisiae*. Due to the involvement of the Bna2 and Bna4 oxygenases in NAD^+ synthesis by *S. cerevisiae*, nicotinate is an essential growth factor for this yeast under anaerobic conditions [54, 368, 369]. A similar strategy was recently

successfully applied to enable oxygen-independent synthesis of pyridine nucleotides in the bacterium *Pseudomonas putida* [370]. The present study represents the first demonstration of a metabolic engineering strategy to eliminate oxygen requirements for NAD^+ synthesis in a yeast.

Functional expression of heterologous NadA quinolinate synthases in *S. cerevisiae* was observed despite the fact that these enzymes are 4Fe-4S iron-sulfur cluster proteins [371, 372], which are notoriously difficult to functionally express in the yeast cytosol [373-376]. However, earlier studies on functional expression of the 4Fe-4S activating protein of bacterial pyruvate-formate lyase [377, 378] demonstrated that low-levels of expression can occur without modification of the yeast machinery for cytosolic assembly of Fe-S clusters. The inability of *At*NadB and *At*NadA to support NAD^+ synthesis in aerobic cultures may be due to oxygen-sensitivity of the 4Fe-4S cluster in the *At*NadA quinolinate synthase domain [379]. In contrast to *Pf*NadA, *At*NadA carries an N-terminal SufE domain which, in other organisms, has been demonstrated to allow this oxygen sensitive enzyme to remain active under aerobic conditions by reconstituting its Fe-S cluster [379].

This work contributes to the understanding of how Neocallimastigomycetes adapted to their anaerobic lifestyle by acquiring genes that enable oxygen-independent synthesis of central metabolic cofactors. Experiments with engineered *S. cerevisiae* strains showed that contribution of the heterologous genes to *in vivo* oxygen-independent cofactor synthesis did not require additional mutations in the host genome. These results indicate how acquisition of functional genes by HGT, even if their expression was initially suboptimal, could have conferred an immediate advantage to ancestors of anaerobic fungi living in cofactor-limited anoxic environments. A similar approach was recently applied to study the physiological impact on *S. cerevisiae* of expressing a heterologous gene encoding squalene-tetrahymanol cyclase, which in Neocallimastigomycetes produces the sterol surrogate tetrahymanol [332]. Functional analysis by heterologous expression in *S. cerevisiae* circumvents the current lack of tools for genetic modification of Neocallimastigomycetes [321] , and can complement biochemical studies [322-324] and genome sequence analyses [325, 326].

Pantothenate and nicotinate, together with the other compounds belonging to the B-group of water-soluble vitamins, are standard ingredients of chemically defined media for aerobic and anaerobic cultivation of yeasts [267]. Further studies of the unique evolutionary adaptations of Neocallimastigomycetes may well provide additional inspiration for engineering robust fungal cell factories that operate under anaerobic conditions.

Material and Methods

Homology and phylogenetic analyses

A set of 51 amino acid sequences previously used to discriminate between L-aspartate decarboxylases (Adc) and glutamate decarboxylases [340] was re-used to identify candidate Neocallimastigomycete Adc sequences. These sequences were used as queries against a database containing all 58109 Neocallimastigomycete proteins deposited in Uniprot trembl (Release 2019_02), which represented 5 species (*Neocallimastix californiae*, *Anaeromyces robustus*, *Piromyces sp* E2, *Piromyces finnis*, and *Pecoramyces ruminatum*), and extracted according to the NCBI taxid 451455. Sequence homology was analysed using BLASTP 2.6.0+ [381] with 10^{-6} as e-value cut-off resulting in 16 Neocallimastigomycete sequences as shared hits from all 51 queries (Supplementary Table S1). Four of these sequences showing homology to experimentally characterised Adc proteins originated from *N. californiae*, and were checked for RNAseq read coverage and splicing junction support revealing A0A1Y1ZL74 as best candidate (Supplementary Fig. S1). For this purpose, Illumina libraries were obtained from the Sequence Read Archive using accession SRR7140690 [382] which were then mapped using STAR 2.6.1a_08-27 [383] against genome assembly GCA_002104975. Alignments were processed using samtools 1.3.1 [384] and visualized using Artemis [385].

A0A1Y1ZL74, also referred to as *Nc*AdcA, was used for a second round of homology search using HMMER 3.2 [386] against 3 different databases built from Uniprot Release 2019_02 to include all refseq sequences from Bacteria (taxid 2), Eukarya (taxid 2759), and Archaea (taxid 2157; TrEMBL and Swiss-Prot categories were also included in this case). Selection for hits with more than 60% alignment length over the query sequence and evalue < 10^{-6} resulted in a total of 325 sequences (103 from Bacteria, 101 from Eukaryotes, and 121 from Archaea; Dataset S1).

The set of 325 A0A1Y1ZL74 homologous sequences, together with those from Tomita *et. al.* (2015) [340], and the 16 Neocallimastigomycete sequences from above were used for further phylogenetic analyses. A total number of 387 sequences (Dataset S2) were aligned with MAFFT v7.402 [387] in "einsi" mode, alignments were trimmed with trimAl v1.2 [388] in "gappyout" mode, and then used to build a maximum likelihood phylogenetic tree with RAxML-NG 0.8.1 [347] using default parameters with the exception of the use of the PROTGTR+FO model and 100 bootstrap replicates. The resulting phylogenetic tree drawn with iTOL [389] is shown in Fig. 2, corresponding sequences and the unannotated tree are provided in Datasets S2 and S3.

Proteomes from species showing an Adc homolog were extracted into individual fasta files and used for (co-)orthology search with ProteinOrtho6 [390]. A0A1Y1ZL74 ortholog groups were then extracted and subjected to alignment, trimming, and phylogenetic analysis as described above. The resulting phylogenetic tree is shown in

Fig. 3, corresponding sequences and the unannotated tree are provided in Datasets S4 and S5.

Abaccus v1.1 [346] (https://github.com/Gabaldonlab/Abaccus) was used to search the three presented in Fig. 3 (Dataset S5) for evidence of HGT. For this purpose, the taxonomy table provided as default was supplemented with definitions for the additional chytrids considered in this study.

Multiple sequence alignment was also performed with Clustal omega 1.2.4 [391] to compare selected aminoacid sequences showing candidate and experimentally characterised Adcs, against bacterial PanDs. These sequences and alignments are shown in Dataset S6.

Strains, media and maintenance

S. cerevisiae strains used and constructed in this study (Table 4) were derived from the CEN.PK lineage [273]. Yeast cultures were routinely propagated in YP (10 g L^{-1} Bacto yeast extract [Becton, Dickinson and Co., Sparks, MD], 20 g L^{-1} Bacto peptone [Becton, Dickinson and Co]) or synthetic medium (SM) [13]. YP and SM were autoclaved at 121 °C for 20 min. SM was then supplemented with 1 mL L^{-1} of 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). Where indicated, nicotinic acid or pantothenic acid were omitted from the vitamin solution, yielding SM without nicotinic acid (SMΔnic) and SM without pantothenic acid (SMΔpan), respectively. A concentrated glucose solution was autoclaved separately for 15 min at 110 °C and added to SM and YP to a concentration of 20 g L^{-1} or 50 g L^{-1}, yielding SMD and YPD, respectively. SMD with urea or acetamide instead of ammonium sulfate (SMD-urea and SMD-Ac, respectively) were prepared as described previously [392, 393]. For anaerobic growth experiments, sterile media were supplemented with Tween 80 (polyethylene glycol sorbate monooleate, Merck, Darmstadt, Germany) and ergosterol (≥95 % pure, Sigma-Aldrich, St. Louis, MO) as described previously [357]. Yeast strains were grown in 500-mL shake flasks containing 100 mL medium or in 100-mL shake flasks containing 20 mL medium. Shake-flask cultures were incubated at 30 °C and shaken at 200 rpm in an Innova Incubator (Brunswick Scientific, Edison, NJ). Solid media were prepared by adding 15 g L^{-1} Bacto Agar (Becton, Dickinson and Co) and, when indicated, 200 mg L^{-1} G418 (Thermo Scientific, Waltham, MA). After genotyping, engineered strains were restreaked twice to select single clones. Removal of the gRNA carrying plasmid was done as previously described [313]. Stock cultures were prepared by adding glycerol to a final concentration of 33 % (v/v), frozen and stored at -80°C.

Molecular biology techniques

DNA was PCR amplified with Phusion Hot Start II High Fidelity Polymerase (Thermo Scientific) and desalted or PAGE-purified oligonucleotide primers (Sigma Aldrich) by following manufacturers' instructions. DreamTaq polymerase (Thermo Scientific) was used for diagnostic PCR. Oligonucleotide primers used in this study are listed in Supplementary Table S1. PCR products were separated by gel electrophoresis using 1 % (w/v) agarose gel (Thermo Scientific) in TAE buffer (Thermo Scientific) at 100 V for 25 min and purified with either GenElute PCR Clean-Up Kit (Sigma Aldrich) or with Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA). Plasmids were purified from *E. coli* using a Sigma GenElute Plasmid Kit (Sigma Aldrich). Yeast genomic DNA was isolated with the SDS/LiAc protocol [314]. Yeast strains were transformed with the lithium acetate

Table 4: *S. cerevisiae* strains used in this study.

Name	Relevant genotype	Parental strain	Reference
CEN.PK113-7D	MATa *URA3*	-	[273]
CEN.PK113-5D	MATa *ura3-52*	-	[273]
IMX585	MATa *can1*Δ::*Spycas9*-natNT2 *URA3*	CEN.PK113-7D	[313]
IMX581	MATa *ura3-52* *can1*Δ::*Spycas9*-natNT2	CEN.PK113-5D	[313]
IMX2292	MATa *can1*Δ::*Spycas9*-natNT2 *URA3 fms1*Δ	IMX585	[394]
IMK877	MATa *can1*Δ::*Spycas9*-natNT2 *URA3 bna2*Δ	IMX585	This study
IMX2301	MATa *can1*Δ::*Spycas9*-natNT2 *URA3 bna2*Δ *sga1*::p*TDH3-PfnadA*-t*ENO1* p*CCW12-PfnadB*-t*ENO2*	IMK877	This study
IMX2302	MATa *can1*Δ::*Spycas9*-natNT2 *URA3 bna2*Δ *sga1*::p*TDH3-AtNADA*-t*ENO1* p*CCW12-AtNADB*-t*ENO2*	IMK877	This study
IMX2293	MATa *ura3-52 can1*Δ::*Spycas9*-natNT2 *fms1*Δ	IMX581	This study
IMX2300	MATa *URA3 ura3-52*::p*TDH3-NcadcA*-t*ENO2* *can1*Δ::*Spycas9*-natNT2 *fms1*Δ	IMX2293	This study
IMX2300-1	MATa *URA3 ura3-52*::p*TDH3-NcadcA*-t*ENO2* *can1*Δ::*Spycas9*-natNT2 *fms1*Δ Colony isolate 1	IMX2300	This study
IMX2305	MATa *fms1*Δ *URA3 ura3-52*::p*RPL12b-TcPAND*-t*TDH1 can1*Δ::*Spycas9*-natNT2	IMX2293	This study

Spy: *Streptococcus pyogenes*; *Pf*: *Piromyces finnis*; *Nc*: *Neocallimastix californiae*; *At*: *Arabidopsis thaliana*; *Tc*: *Tribolium castaneum*.

method [315]. Four to eight single colonies were re-streaked three consecutive times on selective media and diagnostic PCR were performed to verify their genotype. *Escherichia coli* XL1-blue was used for chemical transformation [316]. Plasmids were then isolated and verified by either restriction analysis or by diagnostic PCR. Lysogeny Broth (LB; 10 g L^{-1} Bacto Tryptone, 5 g L^{-1}, Bacto Yeast Extract with 5 g L^{-1} NaCl) was used to propagate *E. coli* XL1-Blue. LB medium was supplemented with 100 mg L^{-1} ampicillin for selection of transformants. The overnight grown bacterial cultures were stocked by adding sterile glycerol at a final concentration of 33 % (v/v) after which samples were frozen and stored at -80 °C.

Plasmid construction

Plasmids used and cloned in this study are shown in Table 5. Plasmids carrying two copies of the same gRNA were cloned by Gibson assembly [313, 395]. In brief, an oligo carrying the gene-specific 20 bp target sequence and a homology flank to the plasmid backbone was used to amplify the fragment carrying the 2µm origin of replication sequence by using pROS13 as template. The backbone linear fragment was amplified using primer 6005 and pROS11 as template [319]. The two fragments were then gel purified and assembled *in vitro* using the NEBuilder HiFi DNA Assembly Master Mix (New England BioLabs, Ipswich, MA) following manufacturer's instructions. Transformants were selected on LB plates supplemented with 100 mg L^{-1} ampicillin or 50 mg L^{-1} kanamycin. Primer 11861 was used to amplify the 2µm fragment containing two identical gRNA sequences for targeting *BNA2*. The PCR product was then cloned in a pROS11 backbone yielding plasmid pUDR315.

The coding sequences for *AtNADA*, *AtNADB*, *PfnadA*, *PfnadB*, and *NcadcA* were codon-optimized for expression in *S. cerevisiae* and ordered as synthetic DNA through GeneArt (Thermo Fisher Scientific). The plasmids carrying the expression cassettes for *TcPAND, AtNADA*, *AtNADB*, *PfnadA* and *PfnadB* were cloned by Golden Gate assembly using the Yeast Toolkit (YTK) DNA parts [397]. These plasmids were cloned using the pYTK096 integrative backbone that carries long homology arms to the *URA3 locus* and a *URA3* expression cassette allowing for selection on SM lacking uracil. The *TcPAND* coding sequence was amplified using the primer pair 11877/11878 and pCfB-361 as template. Then, the linear *TcPAND* gene and plasmids pUD1096, pUD1097, pUD652, and pUD653 carrying the coding sequence for *AtNADA*, *AtNADB*, *PfnadA*, and *PfnadB*, respectively, were combined together with YTK-compatible part plasmids in BsaI (New England BioLabs) golden gate reactions to yield plasmid pUDI168, pUDI245, pUDE931, pUDI243, and pUDI244, respectively. A detailed list of the YTK-compatible parts used for constructing each plasmid can be found in Supplementary Table S2.

The plasmid carrying the expression cassette for *NcadcA* was cloned by Gibson

assembly. The *pTDH3* promoter, the *NcadcA* coding sequence, the *tENO2* terminator and the pYTK0096 backbone were amplified by PCR using primer pairs 16721/16722, 16723/16724, 16725/16726, and 16727/16728 respectively, using pYTK009, pUD1095, pYTK055, and pYTK096 as template, respectively. Each PCR product was then gel purified and combined in equimolar amounts in a Gibson reaction that yielded pUDI242.

Table 2: Plasmids used in this study.

Name	Characteristics	Reference
pROS10	2µm *bla* ori *URA3* gRNA-*CAN1*.Y gRNA-*ADE2*.Y	[313]
pROS11	2µm *bla* ori amdSYM gRNA-*CAN1*.Y gRNA-*ADE2*.Y	[313]
pROS13	2µm *bla* ori kanMX gRNA-*CAN1*.Y gRNA-*ADE2.Y*	[313]
pUDR119	2µm *bla* ori amdSYM gRNA-*SGA1* gRNA-*SGA1*	[396]
pYTK009	p*TDH3 cat* ColE1	[317]
pYTK010	p*CCW12 cat* ColE1	[317]
pYTK017	p*RPL18B cat* ColE1	[317]
pYTK051	t*ENO1 cat* ColE1	[317]
pYTK055	t*ENO2 cat* ColE1	[317]
pYTK056	t*TDH1 cat* ColE1	[317]
pYTK096	ConLS' *gfp* ConRE'*URA3 ntpII* ColE1 5'*URA3*	[317]
pGGKd017	ConLS' *gfp* ConRE' *URA3* 2 µm *bla* ColE1	[157]
pCfB-361	2µm *bla* ori pTEF1-*TcPAND*-tCYC1 HIS3*	[353]
pUDR652	*bla* 2µm amdSYM gRNA-*FMS1* gRNA-*FMS1*	[394]
pUD652	*bla PfnadA**	GeneArt, this study
pUD653	*bla PfnadB**	GeneArt, this study
pUD1095	*bla NcadcA**	GeneArt, this study
pUD1096	*bla AtNADA**	GeneArt, this study
pUD1097	*nptII AtNADB**	GeneArt, this study
pUDR315	*bla* 2µm *amdSYM* gRNA-*BNA2* gRNA-*BNA2*	This study
pUDI168	p*RPL18B-TcPAND**-t*TDH1 URA3 ntpII* ColE1 5'*URA3*	This study
pUDI242	p*TDH3-NcadcA**-t*ENO2 URA3 ntpII* ColE1 5'*URA3*	This study
pUDI243	p*TDH3*-Pf*NADA**-t*ENO1 URA3 ntpII* ColE1 5'*URA3*	This study
pUDI244	p*CCW12-PfnadB**-t*ENO2 URA3 ntpII* ColE1 5'*URA3*	This study
pUDI245	p*TDH3-AtNADA**-t*ENO1 URA3 ntpII* ColE1 5'*URA3*	This study
pUDE931	p*CCW12-AtNADB**-t*ENO2 URA3* 2µm *bla* ColE1	This study

Spy: *Streptococcus pyogenes*; *Pf*: *Piromyces finnis*; *Nc*: *Neocallimastix californiae*; *At*: *Arabidopsis thaliana*; *Tc*: *Tribolium castaneum*.*Codon-optimized for expression in *S. cerevisiae*.

Strain construction

S. cerevisiae strains were transformed using the LiAc/SS-DNA/PEG and CRISPR/Cas9 method [313, 315, 398]. For deletion of the *BNA2* gene, IMX585 (*can1Δ::Spycas9-natNT2*) was transformed with 500 ng of the *BNA2* targeting gRNA plasmid pUDR315 together with 500 ng of the annealed primer pair 11862/11863 as repair dsDNA oligo, yielding strain IMK877. The resulting strain was then used for the integration of the two heterologous *NADB-A* pathways. Expression cassettes for *AtNADA*, *AtNADB*, *PfnadA*, *PfnadB*, were amplified from plasmids pUDI245, pUDE931, pUDI243, pUDI244, respectively, using primer pairs 13123/13124, 13125/10710, 13123/13124, 13125/10710, respectively. Then, 500 ng of each pair of gel purified repair cassettes were co-transformed in IMK877 together with 500 ng the *SGA1* targeting gRNA plasmid, yielding IMX2302 (*sga1::AtNADA AtNADB*) and IMX2301 (*sga1::PfnadA PfnadB*).
For deletion of the *FMS1* gene, IMX581 (*can1Δ::Spycas9-natNT2 ura3-52*) was transformed with 500 ng of the *FMS1* targeting gRNA plasmid pUDR652 together with 500 ng of the annealed primer pair 13527/13528 as repair dsDNA oligo, resulting in IMX2293. Then, 500 ng of plasmids pUDI168 and pUDI242 carrying the expression cassettes for *TcPAND* and *NcadcA,* respectively, were NotI (Thermo Fisher) digested and separately transformed in IMX2293, yielding IMX2305, and IMX2300, respectively. Selection of IMX2305 and IMX2300 was done on SMD agar plate since the integration of each Adc encoding cassette also restored the *URA3* phenotype. In contrast, selection of IMK877 was done on SMD-Ac agar plates while selection of IMX2302, IMX2301, and IMX2293 was done YPD-G418 agar plates. Strains IMK877, IMX2300, IMX2302, and IMX2301 were stocked in SMD, while IMX2305 and IMX2293 were stocked in SMDΔpan and YPD, respectively.

Aerobic growth studies in shake flasks

For the determination of the specific growth rate of the engineered strains under aerobic conditions, a frozen aliquot was thawed and used to inoculate a 20 mL wake-up culture that was then used to inoculate a pre-culture in a 100 mL flask. The exponentially growing pre-culture was then used to inoculate a third flask to an initial OD_{660} of 0.2. The flasks were then incubated, and growth was monitored using a 7200 Jenway Spectrometer (Jenway, Stone, United Kingdom). Specific growth rates were calculated from at least five time-points in the exponential growth phase of each culture. Wake-up and pre-cultures of IMX2301 and IMX2302 were grown in SMDΔnic. Wake-up and pre-cultures of IMX2300 and IMX2305 were grown in SMDΔpan while wake-up and pre-cultures of IMK877 and IMX2292 were grown in SMD.

Anaerobic growth studies in shake flasks

Anaerobic shake-flask based experiments were performed in Lab Bactron 300 anaerobic workstation (Sheldon Manufacturing Inc., Cornelius, OR) containing an atmosphere of 85 % N_2, 10 % CO_2, and 5 % H_2. Flat-bottom shake flasks of 50 mL were filled with 40 mL SMD-urea media containing 50 g L^{-1} glucose as carbon source, to ensure depletion of the vitamin/growth factor of interest, and 20 g L^{-1} glucose for the first transfer. Media were supplemented with vitamins, with and without pantothenic acid or nicotinic acid as indicated, and in all cases supplemented with Tween 80 and ergosterol. Sterile medium was placed inside the anaerobic chamber 24 h prior to inoculation for removal of oxygen. Traces of oxygen were continuously removed with a regularly regenerated Pd catalyst for H_2-dependent oxygen removal placed inside the anaerobic chamber. Aerobic overnight shake-flask cultures on SMD-urea were used to inoculate the anaerobic shake flask without pantothenic acid or without nicotinic acid at an initial OD_{600} of 0.2. Cultures were cultivated at 30 °C with continuous stirring at 240 rpm on IKA KS 260 Basic orbital shaker platform (Dijkstra Verenigde BV, Lelystad, the Netherlands). Periodic optical density measurements at a wavelength of 600 nm using an Ultrospec 10 cell density meter (Biochrom, Cambridge, United Kingdom) inside the anaerobic environment were used to follow the growth over time. After growth had ceased and the OD_{600} no longer increased the cultures were transferred to SMD-urea with 20 g L^{-1} glucose at an OD_{600} of 0.2 [357].

Anaerobic bioreactor cultivation

Anaerobic bioreactor batch cultivation was performed in 2-L laboratory bioreactors (Applikon, Schiedam, the Netherlands) with a working volume of 1.2 L. Bioreactors were tested for gas leakage by applying 0.3 bar overpressure while completely submerging them in water before autoclaving. Anaerobic conditions were maintained by continuous sparging of the bioreactor cultures with 500 mL N_2 min^{-1} (≤0.5 ppm O_2, HiQ Nitrogen 6.0, Linde Gas Benelux, Schiedam, the Netherlands). Oxygen diffusion was minimized by using Fluran tubing (14 Barrer O2, F-5500-A, Saint-Gobain, Courbevoie, France) and Viton O-rings (Eriks, Alkmaar, the Netherlands). Bioreactor cultures were grown on either SMDΔpan or SMDΔnic with ammonium sulfate as nitrogen source. pH was controlled at 5 using 2 M KOH. The autoclaved mineral salts solution was supplemented with 0.2 g L^{-1} sterile antifoam emulsion C (Sigma-Aldrich). Bioreactors were continuously stirred at 800 rpm and temperature was controlled at 30 °C. Evaporation of water and volatile metabolites was minimized by cooling the outlet gas of bioreactors to 4 °C in a condenser. The outlet gas was then dried with a PermaPure PD-50T-12MPP dryer (Permapure, Lakewood, NJ) prior to analysis. CO_2 concentrations in the outlet gas were measured with an NGA 2000 Rosemount gas analyser (Emerson, St. Louis, MO). The gas analyser was calibrated with reference gas containing 3.03 % CO_2 and N6-grade N_2 (Linde Gas Benelux,

Schiedam, The Netherlands).

Frozen glycerol stock cultures were used to inoculate aerobic 100 mL shake flask cultures on either SMDΔpan or SMDΔnic. Once the cultures reached OD_{660} > 5, a second 100 mL aerobic shake-flask pre-culture on the same medium was inoculated. When this second pre-culture reached the exponential growth phase, biomass was harvested by centrifugation at 3000 *g* for 5 min and washed with sterile demineralized water. The resulting cell suspension was used to inoculate anaerobic bioreactors at an OD_{660} of 0.2.

Analytical methods

Biomass dry weight measurements of the bioreactor batch experiments were performed using pre-weighed nitrocellulose filters (0.45 µm, Gelman Laboratory, Ann Arbor, MI). 10 mL culture samples were filtrated and then the filters were washed with demineralized water prior to drying in a microwave oven (20 min at 360 W) and weight measurement. Metabolite concentrations in culture supernatants were analysed by high-performance liquid chromatography (HPLC). In brief, culture supernatants were loaded on an Agilent 1260 HPLC (Agilent Technologies, Santa Clara, CA) fitted with a Bio-Rad HPX 87 H column (Bio-Rad, Hercules, CA). The flow rate was set at 0.6 mL min^{-1} and 0.5 g L^{-1} H_2SO_4 was used as eluent. An Agilent refractive-index detector and an Agilent 1260 VWD detector were used to detect culture metabolites [309]. An evaporation constant of 0.008 divided by the volume in liters, was used to correct HPLC measurements of ethanol in the culture supernatants, taking into account changes in volume caused by sampling [399]. Statistical analysis on product yields was performed by means of an unpaired two-tailed Welch's t-test.

Whole-genome sequencing and analysis.

Genomic DNA of strains IMX2300 and IMX2300-1 was isolated with a Blood & Cell Culture DNA Kit with 100/G Genomics-tips (QIAGEN, Hilden, Germany) according to the manufacturers' instructions. The Miseq Reagent Kit v3 (Illumina, San Diego, CA), was used to obtain 300 bp reads for paired-end sequencing. Genomic DNA was sheared to an average of 550 bp fragments using an M220 ultrasonicator (Covaris, Wolburn, MA). Libraries were prepared by using a TruSeq DNA PCR-Free Library Preparation kit (Illumina) following manufacturer's instructions. The samples were quantified by qPCR on a Rotor-Gene Q PCR cycler (QIAGEN) using the Collibri Library quantification kit (Invitrogen Carlsbad, CA). Finally, the library was sequenced using an Illumina MiSeq sequencer (Illumina, San Diego, CA) resulting in a minimum 50-fold read coverage. Sequenced reads were mapped using BWA 0.7.15-r1142-dirty [400] against the CEN. PK113-7D genome [225] containing an extra contig with the relevant integration cassette. Alignments were processed using SAMtools 1.3.1 [384], and sequence variants were called

using Pilon 1.18 [401], processed with ReduceVCF 12 (https://github.com/AbeelLab/genometools/blob/master/scala/abeel/genometools/reducevcf/ReduceVCF.scala), and annotated using VCFannotator (http://vcfannotator.sourceforge.net/) against GenBank accession GCA_002571405.2 [267].

Data availability

DNA sequencing data of the *Saccharomyces cerevisiae* strains IMX2300 and IMX2300-1 were deposited at NCBI (https://www.ncbi.nlm.nih.gov/) under BioProject accession number PRJNA634013. All measurement data and calculations used to prepare Fig. 4-5 and Tables 4-5 of the manuscript are available at the 4TU.Centre for research data repository (https://researchdata.4tu.nl/) under doi: 10.4121/uuid:c3d2326d-9ddb-469a-b889-d05a09be7d97.

Acknowledgments

We thank Dr. Irina Borodina for providing us the codon-optimized *TcPAND* gene, Sabina Shrestha for constructing strain IMK877, and Dr. Toni Gabaldón for feedback on phylogenetic analyses. TP and J-MD were supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie action PAcMEN (grant agreement No 722287). AMV, WJCD, JB, SJW, RAO-M, CM and JTP were funded by an Advanced Grant of the European Research Council to JTP (grant # 694633). Authors declare that they have no conflict of interests.

Author contributions

All authors contributed to the experimental design. TP, AMV, JMD and JTP wrote a first version of the manuscript. All authors critically read this version, provided input and approved the final version. RAO-M performed the phylogenetic analysis. TP constructed the *S. cerevisiae* strains and performed the aerobic characterization. AMV, WJCD and TP performed the anaerobic chamber experiments. JB, CM, TP, AMV, and SJW performed and analysed the bioreactor experiments.

Chapter 4: Engineering heterologous molybdenum cofactor biosynthesis and nitrate assimilation pathways enables nitrate utilization by *Saccharomyces cerevisiae*

Thomas Perli, Daan N. A. van der Vorm, Mats Wassink, Marcel van den Broek, Jack T. Pronk, Jean-Marc Daran

Essentially as published in
Metabolic engineering
2021.02.004
https://doi.org/10.1016/j.ymben.2021.02.004
Supplementary material available online

Abstract

Metabolic capabilities of cells are not only defined by their repertoire of enzymes and metabolites, but also by availability of enzyme cofactors. The molybdenum cofactor (Moco) is widespread among eukaryotes but absent from the industrial yeast *Saccharomyces cerevisiae*. No less than 50 Moco-dependent enzymes covering over 30 catalytic activities have been described to date, introduction of a functional Moco synthesis pathway offers interesting options to further broaden the biocatalytic repertoire of *S. cerevisiae*. In this study, we identified seven Moco biosynthesis genes in the non-conventional yeast *Ogataea parapolymorpha* by SpyCas9-mediated mutational analysis and expressed them in *S. cerevisiae*. Functionality of the heterologously expressed Moco biosynthesis pathway in *S. cerevisiae* was assessed by co-expressing *O. parapolymorpha* nitrate-assimilation enzymes, including the Moco-dependent nitrate reductase. Following two-weeks of incubation, growth of the engineered *S. cerevisiae* was observed on nitrate as sole nitrogen source. Relative to the rationally engineered strain, the evolved derivatives showed increased copy numbers of the heterologous genes, increased levels of the encoded proteins and a 5-fold higher nitrate-reductase activity in cell extracts. Growth at nM molybdate concentrations was enabled by co-expression of a *Chlamydomonas reinhardtii* high-affinity molybdate transporter. In serial batch cultures on nitrate-containing medium, a non-engineered *S. cerevisiae* was rapidly outcompeted by the spoilage yeast *Brettanomyces bruxellensis*. In contrast, an engineered and evolved nitrate-assimilating *S. cerevisiae* strain persisted during 35 generations of co-cultivation. This result indicates that the ability of engineered strains to use nitrate may be applicable to improve competitiveness of baker's yeast in industrial processes upon contamination with spoilage yeasts.

Introduction

Catalytic activities of many enzymes strictly depend on cofactors, which comprise a chemically diverse collection of non-protein organic compounds (coenzymes) and metal ions [226, 402, 403]. In wild-type micro-organisms, cofactor requirements can strongly influence their catalytic capabilities and/or nutritional requirements. Many vitamins, which are essential organic molecules that cannot be synthesized by the organism itself and therefore have to be supplemented to growth media [3], are in fact cofactors or precursors of them. For example, in the yeast *Saccharomyces cerevisiae*, the vitamin biotin is an essential cofactor for three carboxylases (pyruvate carboxylase Pyc1 and Pyc2, urea carboxylase Dur1,2 or acetyl-CoA carboxylase Acc1) and can be taken up from growth media by the native biotin transporter Vht1 [146, 267]. *S. cerevisiae* strains belonging to the widely used S288c lineage completely lack two genes (*BIO1* and *BIO6*) required for synthesis of biotin, while many other strains of this industrially relevant yeast grow poorly in biotin-free media [7, 34, 152, 154].

When the enzyme repertoire of cells is expanded by metabolic engineering, new cofactor requirements can be introduced. However, nutritional supplementation of these new requirements in culture media may not always be possible due to either lack of supply, high costs and/or absence of membrane transporters for such compounds. In such cases, strain design should include introduction of heterologous cofactor uptake systems and/or pathways for *de novo* cofactor biosynthesis. For example, since *S. cerevisiae* lacks Ni-dependent enzymes and a Ni transporter, replacement of its ATP-dependent urease (Dur1,2) by a heterologous nickel-dependent, ATP-independent enzyme required co-expression of a Ni transporter [227]. Expansion of the organic cofactor repertoire of *S. cerevisiae* is exemplified by studies on *de novo* biosynthesis of opioids in this yeast, which required biosynthesis of tetrahydrobiopterin, the cofactor of the tyrosine hydroxylase that catalyses the first committed step of the (S)-reticuline pathway [258, 404].

The transition metal molybdenum (Mo, typically bioavailable as molybdate MoO_4^{2-}) is an essential trace element for many organisms across the three domains of life. Molybdate is typically incorporated in a tricyclic pterin-based scaffold called molybdopterin to form a molybdenum cofactor (Moco). With the notable exception of nitrogenases, which contain an FeMo cofactor, all known molybdoenzymes harbour Moco variants in their active sites [405-407]. Moco biosynthesis is conserved and extensively studied in prokaryotes and eukaryotes [405, 408, 409]. Nitrate-assimilating yeasts such as *Ogataea parapolymorpha* and *Brettanomyces bruxellensis* synthesize a Moco that is required for activity of nitrate reductase [252]. In contrast, the industrial yeast and eukaryotic model *S. cerevisiae* is devoid of Mo-dependent enzymes and cannot synthesize Moco nor assimilate nitrate [410, 411].

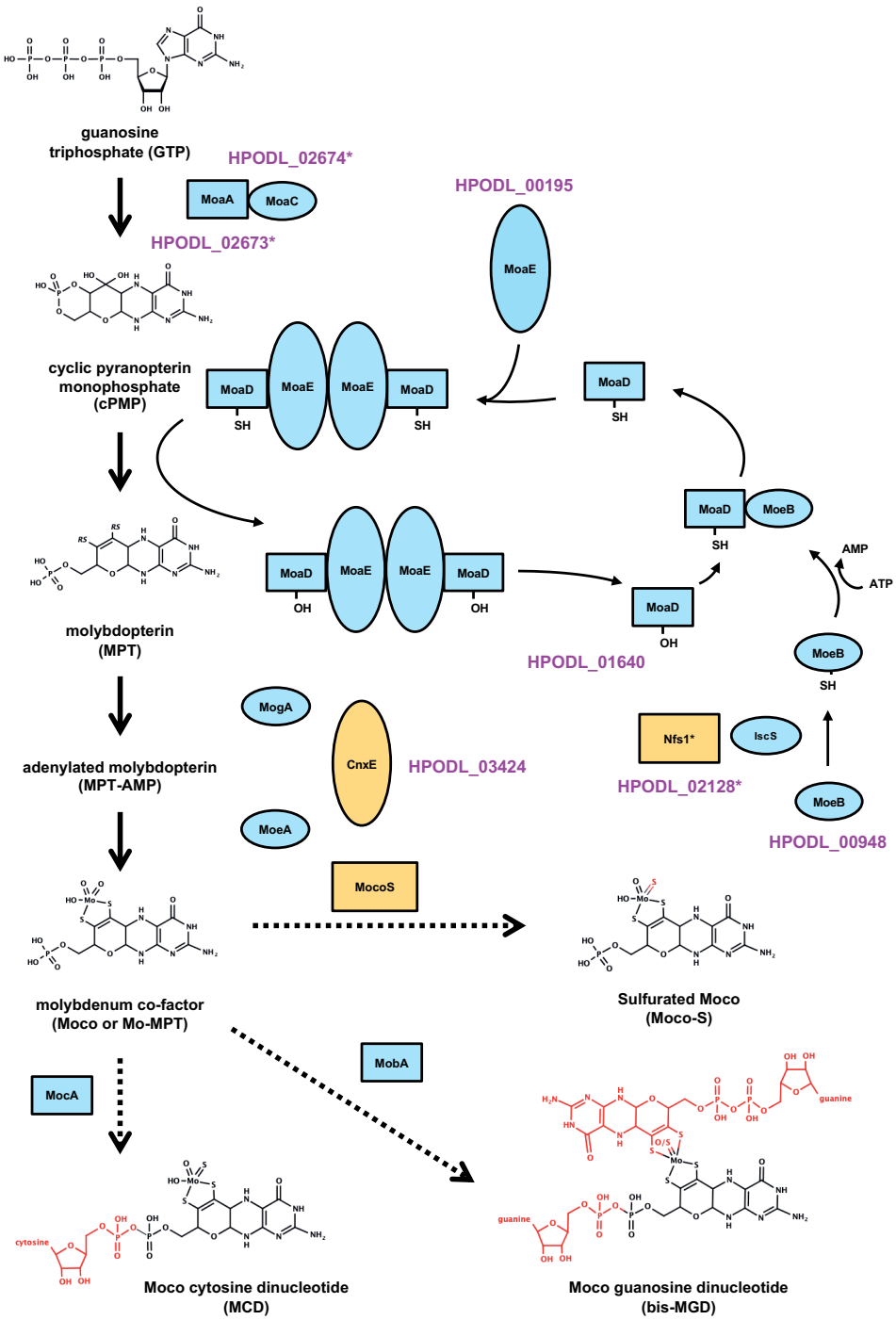
In eukaryotes, the first step of Moco synthesis (Fig. 1), which converts GTP onto cyclic

pyranopterin monophosphate (cPMP), takes place in mitochondria. After export of cPMP to the cytosol, it is first sulfurated to form molybdopterin (MPT). The MPT synthase that catalyses this sulfuration is then regenerated by a sulfur mobilization route involving the Fe-S cluster protein IscS (Nfs1 in eukaryotes), which is shared with the tRNA thiolation pathway [412]. Finally, MPT is adenylated to form MPT-AMP and, after hydrolysis of the adenylate group, molybdate is inserted into the MPT dithiolene group to form the Moco Mo-MPT [409]. Mo-MPT can be further modified in prokaryotes by addition of either cytosine or guanosine to form the Moco variants MPT-cytosine dinucleotide (MCD) or MPT-guanine dinucleotide (MGD), respectively. Prokaryotes as well as eukaryotes can further modify Moco variants by replacing one oxo ligand on the Mo atom by a sulfido ligand to form the mono-oxo Moco variant (Moco-S).

Molybdoenzymes typically use the versatile redox chemistry of MoO_4^{2-} to catalyse redox reactions, often involving oxygen transfer [413]. Based on the ligands at the Mo atom of their Moco, molybdoenzymes are divided in three families: the xanthine oxidase (XO) family, the sulfite oxidase (SO) family and the dimethyl sulfoxide reductase (DMSOR) family [414]. The XO family requires MCD or Moco-S at the catalytic site while the SO family contains Mo-MPT. Members of the DMSOR family instead require the bis-MGD cofactor, which is formed from MoMPT by first forming a bis-Mo-MPT intermediate followed by addition of GMP moieties to its two C-4 phosphates [415, 416]. Recently, bis-Mo-MPT itself has been shown to act as cofactor for the *Escherichia coli* oxidoreductase YdhV [417]. To date, over 50 Mo-containing enzymes have been purified and characterized while many genes have been predicted to encode additional, as yet uncharacterized molybdoproteins [414, 418].

Its excellent accessibility to genome editing and availability of cost-effective procedures for large-scale industrial fermentation have made *S. cerevisiae* a popular platform for production non-native low-molecular weight compounds [419, 420]. Introduction of a functional Moco biosynthesis pathway into *S. cerevisiae* would constitute an important step towards further expansion of the versatility of this yeast as a metabolic engineering

Fig. 1: Schematic representation of the molybdenum-cofactor biosynthesis pathway. GTP is first converted to cPMP by the heterodimer MoaA/MoaC, this step takes place in the mitochondria in eukaryotic cells. Then, MoaD transfers its sulfur moiety to cPMP to yield MPT. MoaD is recycled by sulfur transfer from MoeB, which is itself sulfurated by IscS. MPT is then first adenylated after which the MoO_4^{2-} oxyanion is inserted by the heterodimer MogA/MoeA. In eukaryotes, the latter reaction is catalysed by a single protein (Gephyrin). Moco can be sulfurated at the Mo site by a Moco sulfurase (MocoS) to form sulfurated Moco (Moco-S). Moreover, Moco can be further modified in prokaryotic cells by addition of either cytosine (MocA) or guanosine (MobA) to form Mo-molybdopterin cytosine dinucleotide (MCD) and bis(molybdopterin guanine dinucleotide)molybdenum (bis-MGD), respectively. *E. coli* proteins, fungal homologs, *O. parapolymorpha* homologs and Moco modifications are shown in light blue, yellow, magenta, and red, respectively. Mitochondrial proteins are indicated by an asterisk.


guanosine triphosphate (GTP)
HPODL_02674*
MoaA
MoaC
HPODL_02673*
cyclic pyranopterin monophosphate (cPMP)
HPODL_00195
MoaE
MoaD
SH
MoaE
MoaE
MoaD
SH
molybdopterin (MPT)
MoaD
OH
MoaE
MoaE
MoaD
OH
HPODL_01640
MoaD
SH
MoaD
MoeB
SH
AMP
ATP
MoeB
SH
MogA
CnxE
HPODL_03424
MoeA
adenylated molybdopterin (MPT-AMP)
Nfs1*
IscS
HPODL_02128*
MoeB
HPODL_00948
MocoS
molybdenum co-factor (Moco or Mo-MPT)
Sulfurated Moco (Moco-S)
MobA
MocA
cytosine
Moco cytosine dinucleotide (MCD)
guanine
guanine
Moco guanosine dinucleotide (bis-MGD)

platform, for example by enabling the expression of industrially relevant molybdoproteins such as nitrate reductase and/or metal-dependent formate dehydrogenases. A nitrate-assimilating *S. cerevisiae* strain could increase the robustness of industrial biotechnology processes relying on nitrate containing feedstocks. Nitrate is for instance, often found in sugarcane juice and nitrate levels have been shown to correlate with the fraction of spoilage yeast *B. bruxellensis* found in the fermented must. The inability of *S. cerevisiae* to utilize this nitrogen source has been considered as a critical factor for *B. bruxellensis* contamination in Brazilian ethanol plants [421, 422].

To explore introduction of Moco biosynthesis in *S. cerevisiae*, we first functionally analysed structural genes involved in this process in the nitrate-assimilating yeast *O. parapolymorpha* by Cas9-mediated mutational analysis [423, 424]. The identified *O. parapolymorpha* genes were expressed in *S. cerevisiae*. To enable *in vivo* analysis of the functionality of the heterologous Moco pathway, we co-expressed the *O. parapolymorpha* nitrate assimilation pathway, which includes a Moco-dependent nitrate reductase. As *S. cerevisiae* is not known to harbour a specific molybdate transporter [425], a high-affinity molybdate transporter from *Chlamydomonas reinhardtii* (*Cr*Mot1) was also included in strain designs. Aerobic and anaerobic cultures of engineered *S. cerevisiae* strains were tested and evolved in the laboratory for the ability to use nitrate as sole nitrogen source, followed by whole-genome sequencing of evolved strains to identify relevant mutations. In addition, co-consumption of nitrate and ammonium, as well as the ability to assimilate nitrate at nM concentrations of molybdate were investigated. A possible industrial advantage of nitrate-assimilating *S. cerevisiae* was investigated in co-cultivation experiments with the nitrate-assimilating spoilage yeast *B. bruxellensis*.

Material and Methods

Strains, media and maintenance

Yeast strains used and constructed in this study are shown in Table 1. *O. parapolymorpha* strains were derived from the DL-1 strain [426]. *B. bruxellensis* strain CBS 2499 [427] was obtained from the Westerdijk Institute (Utrecht, The Netherlands). All *S. cerevisiae* strains were derived from the CEN.PK lineage [225, 273]. Yeast strains were grown on either YP (10 g/L Bacto yeast extract, 20 g/L Bacto peptone) or SM medium [13] with either 5 g/L KNO_3, 5 g/L $(NH_4)_2SO_4$, 0.6 g/L acetamide, 0.8 g/L NH_4NO_3, or 2.3 g/L urea (SM_{NO3}, SM_{Amm}, SM_{Ac}, SM_{AN}, and SM_{urea}, respectively) as sole nitrogen source. For *O. parapolymorpha* cultures grown on SMD_{NO3}, KNO_3 was substituted with 4.25 g/L $NaNO_3$. In all SM media variants, with the exception of SM_{Amm}, 6.6 g/L K_2SO_4 was added as a source of sulfate [428]. YP or SM media were autoclaved at 121 °C for 20 min prior to addition of 1 ml/L of filter-sterilized vitamin solution [13]. For anaerobic growth experiments, sterile media were supplemented with Tween 80 (polyethylene

Table 1: Yeast strains used in this study.

Name	Relevant genotype	Parental strain	Reference
O. parapolymorpha CBS 11895 (DL-1)	Wild type		[430]
B. bruxellensis CBS 2499	Wild-type		[427]
S. cerevisiae CEN.PK113-7D	MATa *SUC2 MAL2-8*^{c}		[273]
S. cerevisiae CEN.PK113-5D	MATa *ura3-52 SUC2 MAL2-8*^{c}		[273]
IMX585	MATa *can1*::*Spycas9*-natNT2 *SUC2 MAL2-8*^{c}	CEN.PK113-7D	[313]
IMD019	HPODL_02673^{C155CA}	CBS 11895	This study
IMD020	HPODL_02674^{G172GA}	CBS 11895	This study
IMD021	HPODL_00948^{G235GA}	CBS 11895	This study
IMD022	HPODL_00195^{C126CAT}	CBS 11895	This study
IMD023	HPODL_03424^{C229CT}	CBS 11895	This study
IMD025	*OpYNR1*^{G397GC}	CBS 11895	This study
IMD027	HPODL_01640^{C112CA}	CBS 11895	This study
IMX1777	MATa *can1*::*Spycas9*-natNT2 *sga1*:: *ScTDH3*p-HPODL_02673-*ScENO1*t *ScCCW12*p-HPODL_02674-*ScSSA1*t *ScPGK1*p-HPODL_00195-*ScADH1*t *ScHHF2*p-HPODL_01640-*ScPGK1*t *ScTEF2*p-HPODL_03424-*ScTDH1*t *ScPGM1*p-HPODL_00948-*ScPYK1*t *ScHHF1*p-HPODL_02128-*ScFBA1*t	IMX585	This study
IMX1778	MATa *can1*::*Spycas9*-natNT2 *sga1*:: *ScTDH3*p-HPODL_02673-*ScENO1*t *3ScCCW12*p-HPODL_02674-*ScSSA1*t *ScPGK1*p-HPODL_00195-*ScADH1*t *ScHHF2*p-HPODL_01640-*ScPGK1*t *ScTEF2*p-HPODL_03424-*ScTDH1*t *ScPGM1*p-HPODL_00948-*ScPYK1*t *ScHHF1*p-HPODL_02128-*ScFBA1*t *ScFBA1*p-*CrMOT1*-*ScTEF2*t	IMX585	This study

IMX1779	MATa *can1*::*Spycas9*-natNT2 *sga1*:: *ScFBA1*p-*CrMOT1*-*ScTEF2*t	IMX585	This study
IMX1780	*MATa can1*::*Spycas9*-natNT2 *sga1*::*ScTEF1*p-*OpYNT1*-*ScPDC1*t *ScRPL18b*p-*OpYNR1*-*ScGPM1*t *ScTPI1*p-*OpYNI1*-*ScTPI1*t	IMX585	This study
IMX1781	MATa *can1*::*Spycas9*-natNT2 *sga1*:: *ScTDH3*p-HPODL_02673-*ScENO1*t *ScCCW12*p-HPODL_02674-*ScSSA1*t *ScPGK1*p-HPODL_00195-*ScADH1*t *ScHHF2*p-HPODL_01640-*ScPGK1*t *ScTEF2*p-HPODL_03424-*ScTDH1*t *ScPGM1*p-HPODL_00948-*ScPYK1*t *ScHHF1*p-HPODL_02128-*ScFBA1*t *ScFBA1*p-*CrMOT1*-*ScTEF2*t *ScTEF1*p-*OpYNT1*-*ScPDC1*t *ScRPL18b*p-*OpYNR1*-*ScGPM1*t *ScTPI1*p-*OpYNI1*-*ScTPI1*t	IMX585	This study
IMX1782	MATa *can1*::*Spycas9*-natNT2 *sga1*:: *ScTDH3*p-HPODL_02673-*ScENO1*t *ScCCW12*p-HPODL_02674-*ScSSA1*t *ScPGK1*p-HPODL_00195-*ScADH1*t *ScHHF2*p-HPODL_01640-*ScPGK1*t *ScTEF2*p-HPODL_03424-*ScTDH1*t *ScPGM1*p-HPODL_00948-*ScPYK1*t *ScHHF1*p-HPODL_02128-*ScFBA1*t *ScTEF1*p-*OpYNT1*-*ScPDC1*t *ScRPL18b*p-*OpYNR1*-*ScGPM1*t *ScTPI1*p-*OpYNI1*-*ScTPI1*t	IMX585	This study
IMS816	MATa *can1*::*Spycas9*-natNT2 *sga1*:: *ScTDH3*p-HPODL_02673-*ScENO1*t *ScCCW12*p-HPODL_02674-*ScSSA1*t *ScPGK1*p-HPODL_00195-*ScADH1*t *ScHHF2*p-HPODL_01640-*ScPGK1*t *ScTEF2*p-HPODL_03424-*ScTDH1*t *ScPGM1*p-HPODL_00948-*ScPYK1*t *ScHHF1*p-HPODL_02128-*ScFBA1*t *ScFBA1*p-*CrMOT1*-*ScTEF2*t *ScTEF1*p-*OpYNT1*-*ScPDC1*t *ScRPL18b*p-*OpYNR1*-*ScGPM1*t *ScTPI1*p-*OpYNI1*-*ScTPI1*t (Adapted for growth on nitrate - colony 1)	IMX1781	This study
IMS817	MATa *can1*::*Spycas9*-natNT2 *sga1*:: *ScTDH3*p-HPODL_02673-*ScENO1*t *ScCCW12*p-HPODL_02674-*ScSSA1*t *ScPGK1*p-HPODL_00195-*ScADH1*t *ScHHF2*p-HPODL_01640-*ScPGK1*t *ScTEF2*p-HPODL_03424-*ScTDH1*t *ScPGM1*p-HPODL_00948-*ScPYK1*t *ScHHF1*p-HPODL_02128-*ScFBA1*t *ScTEF1*p-*OpYNT1*-*ScPDC1*t *ScRPL18b*p-*OpYNR1*-*ScGPM1*t *ScTPI1*p-*OpYNI1*-*ScTPI1*t (Adapted for growth on nitrate - colony 1)	IMX1782	This study

IMS815	MATa *can1*::*Spycas9*-natNT2 *sga1*:: *ScTDH3*p-HPODL_02673-*ScENO1*t *ScCCW12*p-HPODL_02674-*ScSSA1*t *ScPGK1*p-HPODL_00195-*ScADH1*t *ScHHF2*p-HPODL_01640-*ScPGK1*t *ScTEF2*p-HPODL_03424-*ScTDH1*t *ScPGM1*p-HPODL_00948-*ScPYK1*t *ScHHF1*p-HPODL_02128-*ScFBA1*t *ScFBA1*p-*CrMOT1*-*ScTEF2*t *ScTEF1*p-*OpYNT1*-*ScPDC1*t *ScRPL18b*p-*OpYNR1*-*ScGPM1*t *ScTPI1*p-*OpYNI1*-*ScTPI1*t (Adapted for growth on nitrate - colony 2)	IMX1781	This study
IMS818	MATa *can1*::*Spycas9*-natNT2 *sga1*:: *ScTDH3*p-HPODL_02673-*ScENO1*t *ScCCW12*p-HPODL_02674-*ScSSA1*t *ScPGK1*p-HPODL_00195-*ScADH1*t *ScHHF2*p-HPODL_01640-*ScPGK1*t *ScTEF2*p-HPODL_03424-*ScTDH1*t *ScPGM1*p-HPODL_00948-*ScPYK1*t *ScHHF1*p-HPODL_02128-*ScFBA1*t *ScTEF1*p-*OpYNT1*-*ScPDC1*t *ScRPL18b*p-*OpYNR1*-*ScGPM1*t *ScTPI1*p-*OpYNI1*-*ScTPI1*t (Adapted for growth on nitrate - colony 2)	IMX1782	This study
IMS819	MATa *can1*::*Spycas9*-natNT2 *sga1*:: *ScTDH3*p-HPODL_02673-*ScENO1*t *ScCCW12*p-HPODL_02674-*ScSSA1*t *ScPGK1*p-HPODL_00195-*ScADH1*t *ScHHF2*p-HPODL_01640-*ScPGK1*t *ScTEF2*p-HPODL_03424-*ScTDH1*t *ScPGM1*p-HPODL_00948-*ScPYK1*t *ScHHF1*p-HPODL_02128-*ScFBA1*t *ScFBA1*p-*CrMOT1*-*ScTEF2*t *ScTEF1*p-*OpYNT1*-*ScPDC1*t *ScRPL18b*p-*OpYNR1*-*ScGPM1*t *ScTPI1*p-*OpYNI1*-*ScTPI1*t (Adapted for growth on nitrate - colony 3)	IMX1781	This study
IMS821	MATa *can1*::*Spycas9*-natNT2 *sga1*:: *ScTDH3*p-HPODL_02673-*ScENO1*t *ScCCW12*p-HPODL_02674-*ScSSA1*t *ScPGK1*p-HPODL_00195-*ScADH1*t *ScHHF2*p-HPODL_01640-*ScPGK1*t *ScTEF2*p-HPODL_03424-*ScTDH1*t *ScPGM1*p-HPODL_00948-*ScPYK1*t *ScHHF1*p-HPODL_02128-*ScFBA1*t *ScTEF1*p-*OpYNT1*-*ScPDC1*t *ScRPL18b*p-*OpYNR1*-*ScGPM1*t *ScTPI1*p-*OpYNI1*-*ScTPI1*t (Adapted for growth on nitrate - colony 3)	IMX1782	This study

glycol sorbate monooleate, Merck, Darmstadt, Germany) and ergosterol (≥95 % pure, Sigma-Aldrich, St. Louis, MO) as described previously [357]. A concentrated glucose solution was autoclaved at 110 °C for 20 min and then added to the YP and SM medium at a final concentration of 20 g/L, yielding SMD and YPD, respectively. For testing the essentiality of a heterologously expressed high-affinity molybdenum transporter, the Mo concentration in the medium was lowered from 1.6 µM to 16 nM, yielding $SMD_{NO3\text{-}LowMo}$. 500-ml Shake flasks containing 100 mL medium and 100-mL shake flasks containing 20 mL medium were incubated at 30 °C and 200 rpm in an Innova Incubator (Brunswick Scientific, Edison, NJ). Solid media were prepared by adding 1.5 % (w/v) Bacto agar and, where indicated, 200 mg/L G418 or 200 mg/L hygromycin. *Escherichia coli* strains were grown in LB (10 g/L Bacto tryptone, 5 g/L Bacto yeast extract, 5 g/L NaCl) supplemented with 100 mg/L ampicillin or kanamycin. To discriminate between *S. cerevisiae* and *B. bruxellensis* in competition experiments [429], filter-sterilized bromocresol green (Sigma-Aldrich) at a final concentration of 88 mg/L was added to either SMD_{NO3} or SMD_{Amm} agar medium ($SMD_{NO3\text{-}blue}$ and $SMD_{Amm\text{-}blue}$, respectively). *S. cerevisiae* and *E. coli* cultures were stored at -80 °C after the addition of 30 % v/v glycerol.

Molecular biology techniques

Primers used in this study are shown in Table 2. DNA was amplified using Phusion Hot Start II High Fidelity Polymerase (Thermo Scientific, Waltham, MA) and desalted or PAGE-purified oligonucleotide primers (Sigma-Aldrich) according to manufacturers' instructions. Diagnostic PCR reactions were performed with DreamTaq polymerase (Thermo Scientific). PCR products were separated by gel electrophoresis on a 1 % (w/v) agarose gel (Thermo Scientific) in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA; Thermo Scientific) and purified with either a GenElutePCR Clean-Up Kit (Sigma-Aldrich) or a Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA). Plasmids were isolated from *E. coli* and *S. cerevisiae* using a Sigma GenElute HP plasmid miniprep kit (Sigma-Aldrich) or a Zymoprep Yeast Plasmid Miniprep II (Zymo Research), respectively, and verified by either restriction digestion or diagnostic PCR. *E. coli* XL1-blue was used for transformation [316]. Yeast genomic DNA used for diagnostic PCR reactions was isolated by using the SDS/LiAc protocol [314]. *S. cerevisiae* transformation was performed with the LiAc method [315] while *O. parapolymorpha* transformation was performed by electroporation [423, 431]. Four to eight colonies were re-streaked on selective medium to select for single clones and diagnostic PCRs were performed to verify the correct genotypes.

Table 2: Primers used in this study

Primer number	Primer sequence	Product(s)
12251	TGCGCCTTGATACGTGC	HPODL_02673_InDelCheck_fwd
12252	AAATAAGAAGGAGAAACATGCAGG	HPODL_02674_InDelCheck_fwd And internal junction 1 check
12253	ACATCCTCCCTCAAGTAGTAGCC	HPODL_00948_InDelCheck_fwd
12254	GACTGGTGTTAGACAAACCGG	HPODL_00195_InDelCheck_fwd
12255	TGCTCGACCATCTCGAGC	HPODL_03424_InDelCheck_fwd
12257	CACCATGGTCGGAAGAACC	*YNR1*_InDelCheck_fwd
12259	GCGTAAACAACATGTCCACC	HPODL_01640_InDelCheck_fwd
12260	GTATGTCCTCGTATGAGACCAGC	HPODL_02673_InDelCheck_rev
12261	CGATTGAGAGAGCTTTTTGGC	HPODL_02674_InDelCheck_rev
12262	CCTGTTCAGAGAAAGAGAAGCC	HPODL_00948_InDelCheck_rev
12263	GGACGTACTGCGAAATCTGG	HPODL_00195_InDelCheck_rev
12264	GATTACTTCTGGAGCTGGCG	HPODL_03424_InDelCheck_rev
12266	ATGTAATTCCTCACGAACTTTGG	*YNR1*_InDelCheck_rev
12268	AAGCCGGGTCTTCTTTCC	HPODL_01640_InDelCheck_rev
12863	GGTCACCCATGTATGCTGGAAATCTGCTCGTCA	pYTK096_backbone_gibson_ pUDI189_FWD
12864	GATAATGATAAACTCGAACTGCGTTGTATTG-CGACGAATTG	pYTK096_backbone_gibson_ pUDI189_REV
12865	TCGTCGCAATACAACGCAGTTCGAGTTTATCATTATC	pYTK009_promoter_gibson_ pUDI189_FWD
12866	CAGCAGTAATGAGGATATCATAGATCTTTTGTTTGT-TTATGTGTGTTTATTC	pYTK009_promoter_gibson_ pUDI189_REV
12867	CATAAACAAACAAAAGATCTATGATATCCTCAT-TACTGCTGAGGC	HPODL_02673_insert_gibson_ pUDI189_FWD
12868	AAAGCTCTCGAGTTAGGATTCATCCTCCAATTAAAAT-CATCG	HPODL_02673_insert_gibson_ pUDI189_REV
12869	TGATTTTAATTGGAGGATGAATCCTAACTCGAGAGCT-TTTGATTAAG	pYTK051_terminator_gibson_ pUDI189_FWD
12870	CGAGCAGATTTCCAGCATACATGGGTGACCAA	pYTK051_terminator_gibson_ pUDI189_FWD _REV
12871	GCATCGTCTCATCGGTCTCATATGGTTGCAATTCAT-GAAAAAGA	HPODL_02674_Insert_golden-gate_pUDI190_FWD
12872	TGCCGTCTCAGGTCTCAGGATCTATTTGAAGATGGTTGA-TAGATCTATGTC	HPODL_02674_Insert_golden-gate_pUDI190_REV
12873	GCATCGTCTCATCGGTCTCATATGTCCATCTTTGTAGA-TATTACTGATAAGC	HPODL_00195_Insert_golden-gate_pUDI191_FWD

12874	TGCCGTCTCAGGTCTCAGGATTTAGGTGCGACTAAGCACGTTAG	HPODL_00195_Insert_goldengate_pUDI191_REV
12875	GCATCGTCTCATCGGTCTCATATGGTCGCAGTTGCTATCGA	HPODL_01640_Insert_goldengate_pUDI192_FWD
12876	TGCCGTCTCAGGTCTCAGGATTTATCCACTTGAAACTGGCGG	HPODL_01640_Insert_goldengate_pUDI192_REV
12879	GCATCGTCTCATCGGTCTCATATGACTGTTGGTATCTTGGTTGTATCA	HPODL_03424_Insert_goldengate_pUDI194_FWD
12880	TGCCGTCTCAGGTCTCAGGATTCACACATAGATCTGGTCGATGAGA	HPODL_03424_Insert_goldengate_pUDI194_REV
12881	GCATCGTCTCATCGGTCTC	*CrMOT1*_insert_goldengate_pUDI195_FWD
12882	TGCCGTCTCAGGTCTCAGGATTTAAGCTCTACCACCTCTAGCAAAAAC	*CrMOT1*_insert_goldengate_pUDI195_REV
12883	GGATGGCGAAAGGATACGCTGGAAATCTGCTCGTCAG	pGGK d017_backbone_*in vivo* assembly_pUDE796_FWD
12884	CCTGTCAAAGTATCACCGTTGTATTGCGACGAATTGC	pGGKd017_backbone_*in vivo* assembly_pUDE796_REV
12885	TCGTCGCAATACAACGGTGATACTTTGACAGGAGC	pGGKp116_promoter_*in vivo* assembly_pUDE796_FWD
12886	CTCATTTAAGGACAAAGACATATATTGTAATATGTGTGTTTGTTTGGATT	pGGKp116_promoter_*in vivo* assembly_pUDE796_REV
12887	ACAAACACACATATTACAATATATGTCTTTGTCCTTAAATGAGTACCTTCG	HPODL_00948_insert_*in vivo* assembly_pUDE796_FWD
12888	AATCATGATTCTTTTTGGATTTAGTAAATAGGGAAGTTTGGGTCTATCTG	HPODL_00948_insert_*in vivo* assembly_pUDE796_REV
12889	CCAAACTTCCCTATTTACTAAATCCAAAAAGAATCATGATTGAATG	pGGKp038_terminator_*in vivo* assembly_pUDE796_FWD
12890	ACGAGCAGATTTCCAGCGTATCCTTTCGCCA	pGGKp038_terminator_*in vivo* assembly_pUDE796_REV
12891	GAGTTCGCGGCTGGAAATCTGCTCGTCAG	pYTK096_backbone_gibson_pUDI197_FWD
12892	GTAAGGCCCCAAGACGTTGTATTGCGACGAATTG	pYTK096_backbone_gibson_pUDI197_REV
12893	GCAATTCGTCGCAATACAACGTCTTGGGGCCTTACCACC	pYTK015_promoter_gibson_pUDI197_FWD
12894	CGATCCTGAACCTGTACATAGATCTATTTTACTATATTATATTTGTTGCTTGT	pYTK015_promoter_gibson_pUDI197_REV
12895	CAACAAATATAATATAGTAAAATAGATCTATGTACAGGTTCAGGATCGGA	HPODL_02128_Insert_gibson_FWD
12896	CAATTAATTTGAATTAACGGATTCAATGTCCTGCCCATTCG	HPODL_02128_Insert_gibson_REV

12897	TGGGCAGGACATTGAATCCGTTAATTCAAATTAATTGATATAGTTTT	pGGKp040_terminator_gibson_pUDI197_FWD
12898	ACGAGCAGATTTCCAGCCGCGAACTCCAA	pGGKp040_terminator_gibson_pUDI197_REV
12899	GCATCGTCTCATCGGTCTCATATGCGACTTTCTACCTTATGGGA	*YNT1*_insert_goldengate_FWD_pUDI198
12900	TGCCGTCTCAGGTCTCAGGATTCAAATTTCCGCTTTCCTAGG	*YNT1*_insert _goldengate_REV_pUDI198
12901	GCATCGTCTCATCGGTCTCATATGGACTCTGTTGTCACTGAGGTG	*YNR1*_insert _goldengate_FWD_pUDI199
12902	TGCCGTCTCAGGTCTCAGGATTCAGAAGTACACTACATACTGTTTATCCAAA	*YNR1*_insert _goldengate_REV_pUDI199
12903	AGAAGTGTACCGGCTGGAAATCTGCTCGTCAG	pGGKd017_backbone_*in vivo* assembly_pUDE797_FWD
12904	ATCTCTGGGTCTTCGTTGTATTGCGACGAATTG	pGGKd017_backbone_*in vivo* assembly_pUDE797_REV
12905	CGTCGCAATACAACGAAGACCCAGAGATGTTGT	pGGKp114_promoter_*in vivo* assembly_pUDE797_FWD
12906	GAGGAACAGAACAAGTCATATTTTAGTTTATGTATGTGTTTTTTGTAGTTATAG	pGGKp114_promoter_*in vivo* assembly_pUDE797_REV
12907	CAAAAAACACATACATAAACTAAAATATGACTTGTTCTGTTCCTCCCTT	*YNI1*_insert_*in vivo* assembly_pUDE797_FWD
12908	TTTTTATATAATTATATTAATCGGATTTACCAGTCGAACGATATTGCTTTG	*YNI1*_insert_*in vivo* assembly_pUDE797_REV
12909	CGTTCGACTGGTAAATCCGATTAATATAATTATATAAAAATATTATCTTCTTTTC	pGGKp042_terminator_*in vivo*_assembly_pUDE797_FWD
12910	CACTGACGAGCAGATTTCCAGCCGGTACACTTCTGAGTAAC	pGGKp042_terminator_*in vivo* assembly_pUDE797_REV
13123	TTTACAATATAGTGATAATCGTGGACTAGAGCAAGATTTCAAATAAGTAACAGCAGCAAACAGTTCGAGTTTATCATTATCAATACTG	SGA1_homology_pUDI189_cassette_integration_fwd
13124	ATAGCATAGGTGCAAGGCTCTCGCCGCTTGTCGAGCTATTGGCATGGATGTGCTCCCTAAATACATGGGTGACCAAAAGAGC	SHR1__homology_pUDI189_cassette_integration_rev
13125	TTAGGGAGCACATCCATGCCAATAGCTCGACAAGCGGCGAGAGCCTTGCACCTATGCTATCACCCATGAACCACACGG	SHR1__homology_pUDI190_cassette_integration_fwd
13126	TCAGCGTGTTGTAATGATGCGCCATGAATTAGAATGCGTGATGATGTGCAAAGTGCCGTCATAAAATTAAAGTAGCAGTACTTCAACCATTAG	SHR2_homology_pUDI190_cassette_integration_rev
13127	GACGGCACTTTGCACATCATCACGCATTCTAATTCATGGCGCATCATTACAACACGCTGAGTGAGTAAGGAAAGAGTGAGGAACT	SHR2__homology_pUDI191_cassette_integration_fwd

13128	GCTACATCTTCCGTACTATGCTGTAGTCTCATGGTCGAGTTCTATTGCTGTTCGGCGGCAGAAATGGGGAGCGATTTG	SHR3_homology_pUDI191_cassette_integration_rev
13129	TGCCGCCGAACAGCAATAGAACTCGACCATGAGACTACAGCATAGTACGGAAGATGTAGCTGTGGAGTGTTTGCTTGGATTCT	SHR3_homology_pUDI192_cassette_integration_fwd
13130	CTCCACTGTACTGCATGTAGCATTCGCCGATCTGCATGATGTGTGACATTCTGCTATCGGACATAGAAATATCGAATGGGAAAAAAAAAC	SHR4_homology_pUDI192_cassette_integration_rev
13131	CCGATAGCAGAATGTCACACATCATGCAGATCGGCGAATGCTACATGCAGTACAGTGGAGTTGATAGGTCAAGATCAATGTAAACAAT	SHR4_homology_pUDI194_cassette_integration _fwd
13132	TGAGAGCTTGTGATAACTGCTCGCCAGTTGTGGTGATCTCCCAGTCGGTGTAGCAGCAATCGTTCAGGGTAATATATTTTAACCG	SHR6_homology_pUDI194_cassette_integration _rev
13133	ATTGCTGCTACACCGACTGGGAGATCACCACAACTGGCGAGCAGTTATCACAAGCTCTCAGTGATACTTTGACAGGAGCTATATCATG	SHR6_homology_pUDE796_cassette_integration_fwd
13134	GGTGAATTGAGAGCTATCCTATATTATAGCAGATGCCGGGTATGCAGCTTGGTAGAATGCGTATCCTTTCGCCATCCTGATA	SHR7_homology_pUDE796_cassette_integration _rev
13135	GCATTCTACCAAGCTGCATACCCGGCATCTGCTATAATATAGGATAGCTCTCAATTCACCTCTTGGGGCCTTACCACC	SHR7_homology_pUDI197_cassette_integration_fwd
13136	CTCAGCCTTAGCCAATATGATCATGTCGTTGCGTCTCGGACCATCTAGTCTACTCTGAAGCGCGAACTCCAAAATGAGC	SHR8_homology_pUDI197_cassette_integration_rev
13570	TATATTTGATGTAAATATCTAGGAAATACACTTGTGTATACTTCTCGCTTTTCTTTTATTCGCGAACTCCAAAATGAGC	SGA1_homology_pUDI197_cassette_integration_rev
13573	CAGTGACGTGAGTGCCATCTGCAGGTCATGTGATGCTATCAGCTACACTGCCAGCAATGACGCGAACTCCAAAATGAGC	SHR9_homology_pUDI197_cassette_integration_ rev
13138	TTTACAATATAGTGATAATCGTGGACTAGAGCAAGATTTCAAATAAGTAACAGCAGCAAATGAACAACAATACCAGCCTTCC	SGA1_homology_pUDI196_cassette_integration_fwd
13139	CTTCAGAGTAGACTAGATGGTCCGAGACGCAACGACATGATCATATTGGCTAAGGCTGAGTGAACAACAATACCAGCCTTCC	SHR8_homology_pUDI196_cassette_integration__fwd
13571	TATATTTGATGTAAATATCTAGGAAATACACTTGTGTATACTTCTCGCTTTTCTTTTATTAGGAAACGTAAATTACAAGGTATATACATACG	SGA1_homology_pUDI196_cassette_integration_rev
13141	CAGTGACGTGAGTGCCATCTGCAGGTCATGTGATGCTATCAGCTACACTGCCAGCAATGAAGGAAACGTAAATTACAAGGTATATACATACG	SHR9_homology_pUDI196_cassette_integration_rev

13142	TTTACAATATAGTGATAATCGTGGACTAGAGCAAGATTTCAAATAAGTAACAGCAGCAAACCTTGCCAACAGGGAGTTC	SGA1_homology_pUDI198_cassette_integration_fwd
13143	TCATTGCTGGCAGTGTAGCTGATAGCATCACATGACCTGCAGATGGCACTCACGTCACTGCCTTGCCAACAGGGAGTTC	SHR9_homology_pUDI198_cassette_integration_fwd
13144	ACGCAATATCGGCCATCGTGCGAGTGTCTCAAACTATCTGTATGCAAATTCGTGCGTGTGCAGTGTTCCTTAATCAAGGATACCTC	SHR10_homology_pUDI198_cassette_integration rev
13145	CACACGCACGAATTTGCATACAGATAGTTTGAGACACTCGCACGATGGCCGATATTGCGTAAGAGGATGTCCAATATTTTTTTTAAG	SHR10_homology_pUDI199_cassette_integration_fwd
13146	TCAGACAATTCTATACGCGGACTGATATGGCAGAAGCTAGGAGACGTTATGCGATCTTAGCATTAAACTACGATGTAAACATCAAGG	SHR11_homology_pUDI199_cassette_integration _rev
13147	CTAAGATCGCATAACGTCTCCTAGCTTCTGCCATATCAGTCCGCGTATAGAATTGTCTGAAAGACCCAGAGATGTTGTTGTC	SHR11_homology_pUDE797_cassette_integration_fwd
13572	TATATTTGATGTAAATATCTAGGAAATACACTTGTGTATACTTCTCGCTTTTCTTTTATTCGGTACACTTCTGAGTAACCCATATAG	SGA1_homology_pUDE797_cassette_integration_rev
3372	GCCCAAATCGGCATCTTTAAATG	Internal_junction_1_check_REV
13727	CCAATTGGTGCGGCAATTG	Internal_junction_2_check_FWD
13728	AAACAAATCACGAGCGACGG	Internal_junction_2_check_REV
13729	GTTGCTTTCTCAGGTATAGCATGAGG	Internal_junction_3_check_FWD
13730	GCGAAACTCTCGGTCTAGTACC	Internal_junction_3_check_REV
13731	CTTTTCTCTTTCCCCATCCTTTACG	Internal_junction_4_check_FWD
13732	CGCCGTCACAAACAACC	Internal_junction_4_check_REV
13733	GTTATGGCGAGAACGTCGG	Internal_junction_5_check_FWD
13734	GCATCACTGCATGTGTTAACCG	Internal_junction_5_check_REV
13735	TCCAATTGTCGTCATAACGATGAGG	Internal_junction_6_check_FWD
13736	GATCCTGGCCGTAATATCTCTCC	Internal_junction_6_check_REV
13737	GTCGGCTCTTTTCTTCTGAAGG	Internal_junction_7_check_FWD
13738	TTAGGGCTTGCGTCAGC	Internal_junction_7_check_REV
13739	ATGTCCTCCAACTCGGC	Internal_junction_8_check_FWD
13740	CGGAGTCCGAGAAAATCTGG	Internal_junction_8_check_REV
13741	GAATTGGCTTAAGTCTGGGTCC	Internal_junction_9_check_FWD
13742	CGTTCTCAAGACGTGGTCC	Internal_junction_9_check_REV
13743	TTTTCAGCCTGTCGTGGTAGC	Internal_junction_10_check_FWD
13744	AGGGAATAAGTAGGGTGATACCGC	Internal_junction_10_check_REV

7806	ACTCGAAGCAGTTCAGAACG	5' External_junction_1_2_3_ check_FWD
4369	GAGGCACATCTGCGTTTCAGG	5' External_junction_1_check_REV
5026	CGTATTACGATAATCCTGCTGTC	5' External_junction_2_check_REV
8410	CGACGAAGAAAAAGAAACGAGG	5' External_junction_3_check_REV
2372	TATTGGTCGGCTCTTTTCTTCTG	3' External_junction_1_check_ FWD
5389	GTTCTTCCTTGCGTTATTCTTCTG	3' External_junction_2_check_ FWD
2375	TGAGCCACTTAAATTTCGTGAATG	3' External_junction_3_check_ FWD
7331	GAGACTCGCATGAGAACATC	3' External_junction_1_2_3_ check_REV
10886	AAGCATCGTCTCATCGGTCTCAATCCAAAAAGAATCATGATTGAATGAAGATATT	*ScPYK1t*_YTK_fwd
10887	TTATGCCGTCTCAGGTCTCACAGCGTATCCTTTCGCCATCCTG	*ScPYK1t*_YTK_rev
10765	AAGCATCGTCTCATCGGTCTCAATCCGATTAATATAATTATATAAAAATATTATCTTCTTTTC	*ScTPI1t*_YTK_fwd
10766	TTATGCCGTCTCAGGTCTCACAGCCGGTACACTTCTGAGTAAC	*ScTPI1t*_YTK_rev
10757	AAGCATCGTCTCATCGGTCTCAATCCGTTAATTCAAATTAATTGATATAGTTTTTTAATG	*ScFBA1t*_YTK_fwd
10758	TTATGCCGTCTCAGGTCTCACAGCCGCGAACTCCAAAATGAGC	*ScFBA1t*_YTK_rev
10773	AAGCATCGTCTCATCGGTCTCAATCCGCGATTTAATCTCTAATTATTAGTTAAAG	*ScPDC1t*_YTK_fwd
10774	TTATGCCGTCTCAGGTCTCACAGCCAGTGTTCCTTAATCAAGGATACC	*ScPDC1t*_YTK_rev
10759	AAGCATCGTCTCATCGGTCTCAATCCGTCTGAAGAATGAATGATTTGATG	*ScGPM1t*_YTK_fwd
10760	TTATGCCGTCTCAGGTCTCACAGCCATTAAACTACGATGTAAACATC	*ScGPM1t*_YTK_rev

Identification of Moco biosynthetic genes

tBLASTn (BLOSUM62 scoring matrix, gap costs of 11 for existence and 1 for extension) analysis was performed to identify putative Moco biosynthetic genes in *O. parapolymorpha* DL-1 [432]. The protein sequences of the *E. coli* molybdopterin-cofactor biosynthesis enzymes MoaA (P30745; GTP 3',8-cyclase), MoaC (P0A738; cyclic pyranopterin monophosphate synthase), MoeB (P12282; molybdopterin-synthase adenylyltransferase), IscS (P0A6B7; cysteine desulfurase), MoaD (P30748; molybdopterin synthase sulfur carrier subunit), MoaE (P30749; molybdopterin synthase catalytic subunit), MogA

(P0AF03; molybdopterin adenylyltransferase) and MoeA (P12281; molybdopterin molybdenumtransferase) were used as query against the *O. parapolymorpha* transcriptome dataset with accession number SRX365635 (https://www.ncbi.nlm.nih.gov//sra?term=SRX365635) [426]. Identified coding sequences were manually annotated in the *O. parapolymorpha* genome sequence (PRJNA60503) and checked for the presence of alternative in-frame start codons upstream of the annotated region. Protein identity was calculated using the Clustal Omega protein alignment tool [433].

Plasmid construction

Plasmids used in this study are shown in Table 3. Plasmids carrying two copies of the same gRNA targeting one of the putative Moco biosynthetic genes in *O. parapolymorpha* were cloned by BsaI Golden Gate assembly as previously described [423, 434]. In brief, synthetic dsDNA strings including a BsaI and ribozyme-flanked 20 bp target sequence were *de novo* synthesized and cloned in plasmids by GeneArt (Thermo Scientific). Then, each of the plasmids pUD697, pUD698, pUD699, pUD700, pUD701, pUD703, pUD704, and pUD705 carrying the gRNA sequence targeting HPODL_02673, HPODL_02674, HPODL_00948, HPODL_00195, HPODL_03424, *OpYNR1*, HPODL_02128, and HPODL_01640, respectively, was combined in a 'one pot' BsaI Golden Gate reaction [435] together with the backbone carrying plasmid pUDP002 (Addgene plasmid number #103872) [423] to yield plasmids pUDP093 ($gRNA_{HPODL\_02673}$) pUDP094 ($gRNA_{HPODL\_02674}$), pUDP095 ($gRNA_{HPODL\_00948}$), pUDP096 ($gRNA_{HPODL\_00195}$), pUDP097 ($gRNA_{HPODL\_03424}$), pUDP099 ($gRNA_{OpYNR1}$), pUDP100 ($gRNA_{HPODL\_02128}$), and pUDP101 ($gRNA_{HPODL\_01640}$), respectively.

In order to assemble plasmids with promoter-gene-terminator expression modules, new promoters and terminator parts compatible with the Golden Gate based yeast toolkit (YTK) were cloned [317]. For this purpose, terminator parts from *S. cerevisiae* were amplified with primers having YTK-compatible ends and *S. cerevisiae* CEN.PK113-7D genomic DNA as template. Primer pairs 10886/10887, 10765/10766, 10757/10758, 10773/10774, and 10759/10760 were used to amplify *ScPYK1t*, *ScTPI1t*, *ScFBA1t*, *ScPDC1t*, and *ScGPM1t*, respectively and purified PCR products were used in a BsmBI Golden Gate reaction together with the pUD565 entry vector to yield pGGKp040, pGGKp042, pGGKp046, and pGGKp048, respectively. Plasmids carrying *ScFBA1p*, *ScTPI1p*, *ScGPM1p* flanked by YTK compatible ends, were *de novo* synthesized by GeneArt (Thermo Scientific) and named pGGKp104, pGGKp114, and pGGKp116, respectively. Promoter fragments of glycolytic genes were selected to be 800 bp long while terminator length was selected to be 300 bp.

Plasmids carrying a transcriptional unit for expression in *S. cerevisiae* were cloned by either Golden Gate assembly, Gibson assembly or *in vivo* homologous recombination in yeast. All coding sequences were amplified from *O. parapolymorpha* DL-1 genomic

Table 3: Plasmids used in this study.

Name	Characteristics	Reference
pUDP002	*ori bla* panARS(OPT) *AgTEF1p-hph-AgTEF1t ScTDH3p*^{BsaI} ^{BsaI}*ScCYC1t AaTEF1p-Spycas9*$^{D147Y\ P411T}$*-ScPHO5t*	[423]
pYTK096	3'*URA3* ConLS' *gfp* ConRE'*URA3 ntpII* ColE1 5'*URA3*	[317]
pGGKd017	ConLS' *gfp* ConRE' *URA3* 2 µm *bla* ColE1	[157]
pYTK009	*cat* ColE1 ^{BsaI}-*ScTDH3*p-^{BsaI}	[317]
pYTK010	*cat* ColE1 ^{BsaI}-*ScCCW12*p-^{BsaI}	[317]
pYTK011	*cat* ColE1 ^{BsaI}-*ScPGK1*p-^{BsaI}	[317]
pYTK012	*cat* ColE1 ^{BsaI}-*ScHHF2*p-^{BsaI}	[317]
pYTK013	*cat* ColE1 ^{BsaI}-*ScTEF1*p-^{BsaI}	[317]
pYTK014	*cat* ColE1 ^{BsaI}-*ScTEF2*p-^{BsaI}	[317]
pYTK015	*cat* ColE1 ^{BsaI}-*ScHHF1*p-^{BsaI}	[317]
pYTK017	*cat* ColE1 ^{BsaI}-*ScRPL18b*p-^{BsaI}	[317]
pYTK051	*cat* ColE1 ^{BsaI}-*ScENO1*t-^{BsaI}	[317]
pYTK052	*cat* ColE1 ^{BsaI}-*ScSSA1*t-^{BsaI}	[317]
pYTK053	*cat* ColE1 ^{BsaI}-*ScADH1*t-^{BsaI}	[317]
pYTK054	*cat* ColE1 ^{BsaI}-*ScPGK1*t-^{BsaI}	[317]
pYTK055	*cat* ColE1 ^{BsaI}-*ScENO2*t-^{BsaI}	[317]
pYTK056	*cat* ColE1 ^{BsaI}-*ScTDH1*t-^{BsaI}	[317]
pUDR119	2µm amdSYM *SNR52*p-gRNA.*SGA1*.Y-*SUP4*t	[437]
pROS13	2µm *bla* kanMX gRNA-*CAN1*.Y gRNA-*ADE2*.Y	[313]
pUD565	*cat* ColE1	[438]
pUD697	*bla* ColE1 ^{BsaI}HH-gRNA$_{OpHPODL_02673}$-HDV^{BsaI}	GeneArt, this study
pUD698	*bla* ColE1 ^{BsaI}HH-gRNA$_{OpHPODL_02674}$-HDV^{BsaI}	GeneArt, this study
pUD699	*bla* ColE1 ^{BsaI}HH-gRNA$_{OpHPODL_00948}$-HDV^{BsaI}	GeneArt, this study
pUD700	*bla* ColE1 ^{BsaI}HH-gRNA$_{OpHPODL_00195}$-HDV^{BsaI}	GeneArt, this study
pUD701	*bla* ColE1 ^{BsaI}HH-gRNA$_{OpHPODL_03424}$-HDV^{BsaI}	GeneArt, this study
pUD703	*bla* ColE1 ^{BsaI}HH-gRNA$_{OpYNR1}$-HDV^{BsaI}	GeneArt, this study
pUD704	*bla* ColE1 ^{BsaI}HH-gRNA$_{OpHPODL_02128}$-HDV^{BsaI}	GeneArt, this study
pUD705	*bla* ColE1 ^{BsaI}HH-gRNA$_{OpHPODL_01640}$-HDV^{BsaI}	GeneArt, this study
pUD728	*bla* ColE1 ^{BsaI}-*CrMOT1*-^{BsaI}	GeneArt, this study
pUDP093	*bla* ColE1 panARS(OPT) *AgTEF1*p-hph-*AgTEF1*t *ScTDH3*p-HH- gRNA$_{OpHPODL_02673}$-HDV-*ScCYC1*t *AaTEF1*p-*Spycas9*$^{D147Y\ P411T}$-*ScPHO5*t	This study
pUDP094	*bla* ColE1 panARS(OPT) *AgTEF1*p-hph-*AgTEF1*t *ScTDH3*p-HH- gRNA$_{OpHPODL_02674}$-HDV-*ScCYC1*t *AaTEF1*p-*Spycas9*$^{D147Y\ P411T}$-*ScPHO5*t	This study

pUDP095	*bla* ColE1 panARS(OPT) *AgTEF1*p-hph-*AgTEF1*t *ScTDH3*p-HH- gRNA$_{OpHPODL_00948}$-HDV-*ScCYC1*t *AaTEF1*p-*Spycas9*$^{D147Y\ P411T}$-*ScPHO5*t	This study
pUDP096	*bla* ColE1 panARS(OPT) *AgTEF1*p-hph-*AgTEF1*t *ScTDH3*p-HH- gRNA$_{OpHPODL_00195}$-HDV-*ScCYC1*t *AaTEF1*p-*Spycas9*$^{D147Y\ P411T}$-*ScPHO5*t	This study
pUDP097	*bla* ColE1 panARS(OPT) *AgTEF1*p-hph-*AgTEF1*t *ScTDH3*p-HH- gRNA$_{OpHPODL_03424}$-HDV-*ScCYC1*t *AaTEF1*p-*Spycas9*$^{D147Y\ P411T}$-*ScPHO5*t	This study
pUDP099	*bla* ColE1 panARS(OPT) *AgTEF1*p-hph-*AgTEF1*t *ScTDH3*p-HH- gRNA$_{OpYNR1}$-HDV-*ScCYC1*t *AaTEF1*p-*Spycas9*$^{D147Y\ P411T}$-*ScPHO5*t	This study
pUDP100	*bla* ColE1 panARS(OPT) *AgTEF1*p-hph-*AgTEF1*t *ScTDH3*p-HH-gRNA$_{OpHPODL_02128}$-HDV-*ScCYC1*t *AaTEF1*p-*Spycas9*$^{D147Y\ P411T}$-*ScPHO5*t	This study
pUDP101	*bla* ColE1 panARS(OPT) *AgTEF1*p-hph-*AgTEF1*t *ScTDH3*p-HH-gRNA$_{OpHPODL_01640}$-HDV-*ScCYC1*t *AaTEF1*p-*Spcas9*$^{D147Y\ P411T}$-*ScPHO5*t	This study
pGGKp040	*cat* ColE1 ^{BsaI}-*ScPYK1*t-^{BsaI}	This study
pGGKp042	*cat* ColE1 ^{BsaI}-*ScTPI1*t-^{BsaI}	This study
pGGKp045	*cat* ColE1 ^{BsaI}-*ScPDC1*t-^{BsaI}	This study
pGGKp046	*cat* ColE1 ^{BsaI}-*ScFBA1*$^{t\text{-}Bs}$aI	This study
pGGKp048	*cat* ColE1 ^{BsaI}-*ScGPM1*t-^{BsaI}	This study
pGGKp104	*cat* ColE1 ^{BsaI}-*ScFBA1*p-^{BsaI}	This study
pGGKp114	*cat* ColE1 ^{BsaI}-*ScTPI1*p-^{BsaI}	This study
pGGKp116	*cat* ColE1 ^{BsaI}-*ScGPM1*p-^{BsaI}	This study
pUDI189	3'*URA3* ConLS' *ScTDH3*p-HPODL_02673-*ScENO1*t ConRE'*URA3 ntpII* ColE1 5'*URA3*	This study
pUDI190	3'*URA3* ConLS' *ScCCW12*p-HPODL_02674-*ScSSA1*t ConRE'*URA3 ntpII* ColE1 5'*URA3*	This study
pUDI191	3'*URA3* ConLS' *ScPGK1*p-HPODL_00195-*ScADH1*t ConRE'*URA3 ntpII* ColE1 5'*URA3*	This study
pUDI192	3'*URA3* ConLS' *ScHHF2*p-HPODL_01640-*ScPGK1*t ConRE'*URA3 ntpII* ColE1 5'*URA3*	This study
pUDI193	3'*URA3* ConLS' *ScTEF1*p-HPODL_00337-*ScENO2*t ConRE'*URA3 ntpII* ColE1 5'*URA3*	This study
pUDI194	3'*URA3* ConLS' *ScTEF2*p-HPODL_03424-*ScTDH1*t ConRE'*URA3 ntpII* ColE1 5'*URA3*	This study
pUDI195	3'*URA3* ConLS' *ScFBA1*p-*CrMOT1*-*ScTEF2*t ConRE'*URA3 ntpII* ColE1 5'*URA3*	This study

pUDE796	ConLS' *ScGPM1*p-HPODL_00948-*ScPYK1*t ConRE' *URA3* 2 µm *bla* ColE1	This study
pUDI197	3'*URA3* ConLS' *ScHHF1*p-HPODL_02128-*ScFBA1*t ConRE'*URA3 ntpII* ColE1 5'*URA3*	This study
pUDI198	3'*URA3* ConLS' *ScTEF1*p-*OpYNT1*-*ScPDC1*t ConRE'*URA3 ntpII* ColE1 5'*URA3*	This study
pUDI199	3'*URA3* ConLS' *ScRPL18b*p-*OpYNR1*-*ScGPM1*t ConRE'*URA3 ntpII* ColE1 5'*URA3*	This study
pUDE797	ConLS' *ScTPI1*p-*OpYNI1*-*ScTPI1*t ConRE' *URA3* 2 µm *bla* ColE1	This study
pUDR653	2 µm *bla* kanMX gRNA-*OpYNR1*.Y gRNA-*OpYNR1*.Y	This study

DNA (gDNA) as template, except for the *Chlamydomonas reinhardtii MOT1* gene, which was codon optimized for expression in *S. cerevisiae* and *de novo* synthesized by GeneArt (Thermo Scientific) resulting in plasmid pUD728. Expression cassettes for HPODL_02674, HPODL_00195, HPODL_01640, HPODL_03424, *CrMOT1*, *OpYNT1*, and *OpYNR1* were constructed *in vitro* by Golden Gate assembly [317]. First, primer pairs 12871/12872, 12873/12874, 12875/12876, 12879/12880, 12881/12882, 12899/12900, and 12901/12902 were used to amplify the HPODL_02674, HPODL_00195, HPODL_01640, HPODL_03424, *CrMOT1*, *OpYNT1*, and *OpYNR1* coding sequence, respectively, to add the YTK part 3 compatible ends. Then, each linear DNA carrying the coding sequence for HPODL_02674, HPODL_00195, HPODL_01640, HPODL_03424, *CrMOT1*, *OpYNT1*, and *OpYNR1* was combined together with the backbone plasmid pYTK096 and the respective promoter/terminator part plasmid pair pYTK010/pYTK052, pYTK011/pYTK055, pYTK012/pYTK054, pYTK014/pYTK056, pGGKp104/pGGKp038, pYTK013/pGGKp045, and pYTK017/pGGKp048, in a BsaI Golden gate reaction to yield pUDI190, pUDI191, pUDI192, pUDI194, pUDI195, pUDI198, and pUDI199, respectively. The expression cassettes for HPODL_02673 and HPODL_02128 were constructed using *in vitro* Gibson assembly [395]. *ScTDH3p* promoter, HPODL_02673 coding sequence, *ScSSA1t* terminator and backbone were amplified using primer pairs 12865/12866, 12867/12868, 12869/12870, 12863/12864 and pYTK009, *O. parapolymorpha* DL-1 gDNA, pYTK055 and pYTK096 as template, respectively. PCR products were then combined in equimolar amounts in an *in vitro* Gibson assembly reaction with NEBuilder HiFi DNA Assembly Master Mix (New England Biolabs, Ipswich, MA) that yielded plasmid pUDI189. Similarly, *ScHHF1p* promoter, HPODL_02128 coding sequence, *ScFBA1t* terminator and backbone fragments were amplified with primer pairs 12893/12894, 12895/12896, 12897/12898, 12891/12892 and pYTK015, *O. parapolymorpha* DL-1 gDNA, pGGKp040 and pYTK096 as template, respectively. Equimolar amounts of these PCR products were then combined in an *in vitro* Gibson assembly reaction that yielded plasmid pUDI197. The expression

cassettes for HPODL_00948 and *OpYNI1* (HPODL_02386) were constructed using *in vivo* assembly in *S. cerevisiae* [436]. *ScGPM1p* promoter, HPODL_00948 coding sequence, *ScPYK1t* terminator and backbone were amplified using primer pairs 12885/12886, 12887/12888, 12889/12890, 12883/12884 and *O. parapolymorpha* DL-1 gDNA, pGGKp038, and pGGKd017 as template, respectively. *S. cerevisiae* CEN.PK113-5D (*MATa ura3-52*) was then co-transformed with equimolar amounts of PCR products to yield pUDE796. Similarly, *ScTPI1p* promoter, HPODL_02386 (*OpYNI1*) coding sequence, *ScTPI1t* terminator and backbone were amplified using primer pairs 12893/12894, 12895/12896, 12897/12898, 12891/12892 and pGGKp114, *O. parapolymorpha* DL-1 gDNA, pGGKp040, and pGGKd017 as template, respectively. *S. cerevisiae* CEN.PK113-5D (*MATa ura3-52*) was then transformed with equimolar amounts of PCR products to yield pUDE797.

Strain construction

O. parapolymorpha strains carrying a single gene disruption were obtained by transformation with the gRNA- and Cas9-carrying plasmid followed by prolonged incubation in selective media as previously described [423]. *O. parapolymorpha* DL-1 strain (CBS 11895) was individually transformed with plasmids pUDP093, pUDP094, pUDP095, pUDP096, pUDP097, pUDP099, and pUDP101 to yield strains IMD019 (HPODL_02673^{C155CA}), IMD020 (HPODL_02674^{G172GA}), IMD021 (HPODL_00948^{G235GA}), IMD022 (HPODL_00195^{C126CAT}), IMD023 (HPODL_03424^{C229CT}), IMD025 (*OpYNR1*^{G397GC}), and IMD027 (HPODL_01640^{C112CA}), respectively. Editing at HPODL_02673, HPODL_02674, HPODL_00948, HPODL_00195, HPODL_03424, *OpYNR1*, HPODL_01640 was verified by PCR amplification of each *locus* using primer pairs 12251/12260, 12252/12261, 12253/12262, 12254/12263, 12255/12264, 12257/12266, and 12258/12268, respectively. Resulting DNA fragments were purified, and Sanger sequenced (Baseclear, Leiden, The Netherlands) to check for the presence of INDELs.

S. cerevisiae strains carrying different combinations of the Moco, Mo-transport, and nitrate modules were obtained by co-transforming strain IMX585 (*MATa Cas9*) with the gRNA^{SGA1} targeting plasmid pUDR119 together equimolar amounts of each expression cassette that was previously amplified by PCR to add unique 60 bp homology flanks. Transformants were selected by plating on SMD_{Ac} solid medium [428]. Correct integration of expression cassettes was assessed by PCR amplification of each recombination junction. Primers used for integration fragments and junction-check PCR reactions are given in Supplementary Figures S1-6. Following genotyping of transformants, gRNA-carrying plasmids were cured [313]. For each transformation, one correctly genotyped clone was stocked at -80 °C and named IMX1777 (Moco), IMX1778 (Moco, Mo importer), IMX1779 (Mo importer), IMX1780 (nitrate), IMX1781 (Moco, Mo importer, nitrate), IMX1782 (Moco, nitrate).

O. parapolymorpha spot-plate assay

Frozen aliquots of *O. parapolymorpha* strains IMD019-23, IMD025 and IMD027, as well as of reference strains *S. cerevisiae* CEN.PK113-7D and *O. parapolymorpha* DL-1 were thawed and used to inoculate 20 mL SMD_{Amm} flask cultures. Once OD_{660} reached a value above 5, cultures were spun down at 3000 *g* for 5 min. Cell pellets were washed thrice with sterile demineralized water and resuspended to an OD_{660} of 1. For each strain, 10 µL aliquots of the resulting suspension was spotted on either $SMD_{Amm\ or}$ SMD_{NO3} agar plates. Photographs were taken after 48 h incubation at 30 °C.

Aerobic shake flask experiments

To adapt engineered *S. cerevisiae* strains IMX1777-1782 for growth on nitrate, they were inoculated in triplicate in 20 mL SMD_{NO3} in 100 mL flasks until, after approximately 2 weeks, OD_{660} reached a value above 5. If no growth was observed after two weeks, cultures were discarded. Each grown culture was restreaked on an SMD_{NO3} agar plate to yield single colonies. One single colony from each independent adaptation experiment was inoculated in 100 mL SMD_{NO3} and stored at -80 °C. Adaptation of strain IMX1781 resulted in independently evolved isolates IMS815, IMS816, and IMS819 while adaptation of IMX1782 resulted in evolved isolates IMS817, IMS818, and IMS821.
For the determination of the specific growth rates of evolved strains IMS815-819, IMS821, and of *O. parapolymorpha* DL-1 and *B. bruxellensis* CBS 2499, frozen stock cultures were used to inoculate 20 mL starter cultures. These were subsequently used to inoculate 100 mL SMD_{NO3} flask cultures to initial OD_{660} values between 0.1 and 0.2. Growth of these cultures was monitored with a 7200 Jenway Spectrophotometer (Jenway, Stone, United Kingdom). Specific growth rates were calculated from at least five time points in the exponential growth phase of each culture.

Anaerobic growth experiments

Anaerobic growth of the engineered *S. cerevisiae* strain IMS816 and the wild-type *B. bruxellensis* strain CBS 2499 was studied in a Lab Bactron 300 anaerobic workstation (Sheldon Manufacturing Inc., Cornelius, OR) containing an atmosphere of 85 % N_2, 10% CO_2 and 5% H_2. Exponentially growing aerobic cultures were used to inoculate anaerobic starter cultures at a OD_{600} of about 0.2. These starter cultures were grown in 50-mL shake flasks containing 40 mL of SMD_{NO3} supplemented with 40 g/L glucose and used inoculate a second anaerobic culture on SMD_{NO3} with 20 g/L glucose. Anaerobic cultures were incubated at 30 °C and shaken at 240 rpm on an IKA KS 260 Basic orbital platform (Dijkstra Verenigde BV, Lelystad, The Netherlands). A regularly regenerated Pd catalyst for H_2-dependent oxygen removal was placed inside the anaerobic chamber. Optical density at 600 nm was periodically measured using a Ultrospec 10 spectrophotometer (Biochrom,

Cambridge, United Kingdom). Sterile media was placed inside the anaerobic chamber at least 24 h prior to inoculation to ensure removal of residual oxygen. When indicated, SMD_{NO3} media were supplemented with 1 ml/L of a concentrated hemin solution that was prepared by adding 0.05% (w/v) hemin (Sigma Aldrich) to a 1:1 ethanol:water solution with 50 mM NaOH. As a negative control for oxygen leaks, a parallel culture of *S. cerevisiae* CEN.PK113-7D strain on SMD_{Urea} without the anaerobic growth factors Tween 80 and ergosterol was included in all anaerobic growth experiments [357].

Competitive cultivation

Frozen stock cultures of *S. cerevisiae* strains IMX585 (*MATa SpyCas9*), IMS816 (Moco - Mo importer - Nitrate) and of *B. bruxellensis* CBS 2499 were used to inoculate 20 mL starter cultures, which were subsequently used to inoculate 100 mL flask cultures on SMD_{NO3}. Upon reaching mid-exponential phase ($1 < OD_{660} < 5$), these cultures were centrifuged at 3000 *g* for 5 min and washed three times in demineralized water. Cells were then resuspended in SMD_{NO3} and co-inoculated at an initial OD_{660} of 0.1 in 100 mL shake-flask cultures on SMD_{NO3}. Triplicate co-cultures were prepared for strain pairs IMX585/CBS 2499 and IMS816/CBS 2499. Flasks were incubated for 48 h prior to plating diluted samples on $SMD_{NO3\text{-}blue}$ and $SMD_{Amm\text{-}blue}$ plates. Plates were incubated for 4 days at 30 °C and then two weeks at 4 °C to develop bromocresol green staining prior to imaging and colony counting (Supplementary Figure S7).

Whole-genome sequencing

Genomic DNA of strains IMX1781, IMX1782, IMS815, IMS816, IMS817, IMS818, and IMS821 was isolated with a Blood & Cell Culture DNA Kit with 100/G Genomics-tips (QIAGEN, Hilden, Germany) following manufacturer's instructions. Illumina-based paired-end sequencing with 150-bp reads was performed on 300-bp insert libraries (Novogene Company Limited, Hong Kong, China) with a minimum resulting coverage of 50 x. Data mapping was performed using bwa 0.7.15-r1142-dirty against the CEN.PK113-7D genome [225] to which an extra contig containing the relevant integration cassette had been previously added. Data processing and chromosome copy number variation determinations were done as previously described [318, 439].

In vitro nitrate reductase activity measurements from cell extract

Frozen stock cultures of *S. cerevisiae* strains IMX1780, IMX1781 and IMS816 were used to inoculate 20 mL starter cultures on SMD_{urea}, which were then used to inoculate 100-mL shake flask cultures on the same medium, to an initial OD_{660} of 0.2. Shake flasks were incubated for 24 or 48 h, until the OD_{660} exceeded 30. Cultures were then centrifuged at 3000 *g* for 5 min and supernatant was discarded. Lysis buffer was prepared by dissolving

1 tablet of complete ULTRA EDTA-free protease inhibitor cocktail (Roche, Basel, Switzerland) in 10 mL ice-cold 100 mM potassium phosphate buffer (pH 7). Cell pellets were resuspended in 1.5 ml lysis buffer and transferred to 1.5 ml bead-beating tubes along with 0.75 g of 400-600 µm acid-washed glass beads (Sigma Aldrich) per tube. Cells were disrupted by six 1-min cycles at 5 m/s speed in a Fast-Prep 24 cell homogenizer (MP Biomedicals, Santa Ana, CA), with 5-min cooling on ice between cycles. Samples were then centrifuged at 14000 *g* and at 4 °C for 10 min. Supernatant was collected in 10 mL centrifuge tubes, diluted by adding 2 mL ice-cold lysis buffer and centrifuged at 20000 *g* and at 4 °C for 1 h. Clear supernatant were then transferred in clean 15 mL plastic tubes and kept on ice prior to analysis. Nitrate-reductase activity was measured by monitoring either NADH or NADPH consumption at 340 nm using a spectrophotometer (Jasco, Easton, MA). Reactions were performed at 30 °C, in 100 mM phosphate buffer pH7. Reaction mixtures included 20 µM FAD, and either 50 or 100 µl of clarified cell extract. After addition of 200 µM NADH or NADPH, background activity was monitored, after which 0.005, 0.05, 1 or 2 mM KNO_3 was added to initiate the reaction. Reaction rates were corrected based on an extinction coefficient of NADH and NADPH of 6.22 mM^{-1} cm^{-1} at 340 nm and corrected for the background activity in the absence of nitrate. Protein contents of cell extracts were quantified with a Quick Start Bradford Assay (Bio-Rad Laboratories, Hercules, CA) following manufacturer's instructions. Specific activities of nitrate reductase in cell extracts were expressed in µmol NAD(P)H min^{-1} (mg protein)$^{-1}$.

Proteome analysis

Starter cultures on 20 mL SMD_{urea} were inoculated with frozen stock cultures of strains IMX1781 and IMS816 and used to inoculate two independent 100 mL flask cultures for each strain at an initial OD_{660} of 0.2. Once these cultures reached and OD_{660} of 4, 1 ml broth was collected and centrifuged at 3000 *g* for 5 min. The cell pellet, which had a volume approximately 60 µl was then subjected to protein extraction and trypsin digestion [440]. Prior to analysis, peptides were resuspended in 30 µl of 3 % acetonitrile/0.01 % trifluoroacetic acid and peptide concentrations were measured with a Nanodrop spectrophotometer (Thermo Scientific) set at 280 nm. One µg of sample was injected into a CapLC system (Thermo Scientific) coupled to an Orbitrap Q-exactive HF-X mass spectrometer (Thermo Scientific). After capture of samples, at a flow rate of 10 µl/min on a precolumn (µ-precolumn C18 PepMap 100, 5µm, 100Å), peptides were separated on a 15-cm C18 easy spray column (PepMap RSLC C18 2µm, 100Å, 150 µmx15cm) at a flow rate of 1.2 µL/min and with a 60-min continuous gradient from 4 % to 76 % acetonitrile in water . Data analysis was performed using Proteome discover 2.4 (Thermo Scientific) with fixed modifications set to carbamidomethyl (C), variable modifications set to oxidation of methionine residues, search mass tolerance set to 20 ppm, MS/MS

tolerance set to 20 ppm, trypsin selected as hydrolytic enzyme and allowing one missed cleavage. False discovery rate was set at 0.1% and the match between runs window was set to 0.7 min. Quantification was exclusively based on unique peptides and normalization between samples was based on total peptide amount. A protein database consisting of the *S. cerevisiae* S288c proteome amino-acid sequences together with sequences of the heterologously expressed proteins was used for protein searches. For each strain analyses were performed on independent biological duplicate samples.

Analytical methods

Metabolite concentrations in culture supernatants were analysed by high-performance liquid chromatography (HPLC) on an Agilent 1260 HPLC (Agilent Technologies, Santa Clara, CA) fitted with a Bio-Rad HPX 87 H column (Bio-Rad). The flow rate was set at 0.6 mL min^{-1}, 0.5 g L^{-1} H_2SO_4 was used as eluent and the column temperature was 65 °C. An Agilent refractive-index detector and an Agilent 1260 VWD detector were used for metabolite quantification [309]. Nitrate, nitrite and ammonium concentrations culture supernatants were measured with a Hach DR3900 spectrophotometer and Hach kits LCK 339, LCK 341, and LCK 304 (Hach Lange, Düsseldorf, Germany), according to the manufacturer's instructions.

Statistical analysis

Statistical significance of differences between measurements from replicate cultures were calculated by using a two-tailed t-test assuming unequal variances (Welch's correction).

Data availability

All measurement data and calculations used to prepare Fig. 2-7 and Supplementary figure S7-8 of the manuscript are available at the 4TU.Centre for research data repository (https://researchdata.4tu.nl/) under doi: 10.4121/13194518. DNA sequencing data of *Saccharomyces cerevisiae* strains IMX1781-2, IMS815-19, and IMS821 were deposited at NCBI (https://www.ncbi.nlm.nih.gov/) under BioProject accession number PRJNA658462. Mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://www.proteomexchange.org/) via the PRIDE partner repository with the dataset identifier PXD020472.

Results

Identification of Moco biosynthesis genes in *O. parapolymorpha*

As a nitrate-assimilating yeast, *O. parapolymorpha* DL-1 can express a functional nitrate reductase (NR). Its genome should therefore carry a full complement of Moco biosynthesis genes, but these have not yet been annotated or characterized. A tBLASTn search of *O. parapolymorpha* DL-1 transcriptome data [426] for orthologs of seven *E. coli* Moco biosynthesis genes yielded strong hits (E value < $1.0e^{-14}$) with six queries (Figure 1, Table 4). A seventh, MoaD, yielded only a weak hit (E-value score 1.4 and 23.5 % sequence identity; Supplementary Table S1) with transcript HPODL_03424 (CnxE). However, *Ec*MogA and *Ec*MoeA showed strong similarities with the 5' and 3' ends, respectively, of the same coding sequence. This observation indicated that, similar to the situation in other eukaryotes, a single *O. parapolymorpha* protein carries MPT adenylyltransferase and molybdenumtransferase domains [441, 442].

The six identified coding sequences were manually annotated in the *O. parapolymorpha* genome sequence (PRJNA60503) and checked for presence of alternative in-frame start codons. *Ec*MoaA (GTP 3',8-cyclase) orthologs such as human MOCS1A and *Arabidopsis thaliana* Cnx2 are known to be iron-sulfur cluster proteins that localize to the mitochondria

Table 4: tBLASTn analysis of *E. coli* Moco-biosynthesis-related proteins versus *O. parapolymorpha* transcriptome.

Query protein (Uniprot ID)	Protein annotation	Filamentous fungi ortholog gene name	Gene name of first hit in *O. parapolymorpha* (E value - % query cover)	Proposed yeast gene name
E. coli MoaA (P30745)	*GTP 3',8-cyclase*	*cnxA*	*HPODL_02673* ($5e^{-64}$ - 93)	*OpCNX1*
E. coli MoaC (P0A738)	*Cyclic pyranopterin monophosphate synthase*	*cnxB*	*HPODL_02674* ($1e^{-39}$ - 87)	*OpCNX2*
E. coli MoeB (P12282)	*Molybdopterin-synthase adenylyltransferase*	*cnxF*	*HPODL_00948* ($7e^{-50}$ - 97)	*OpCNX4*
E. coli IscS (P0A6B7)	*Cysteine desulfurase*	*NFS1*	*HPODL_02128* ($3e^{-172}$ - 99)	*OpNFS1*
E. coli MoaD (P30748)	*Molybdopterin synthase sulfur carrier subunit*	*cnxG*	*HPODL_01640* (1.4 - 95)	*OpCNX5*
E. coli MoaE (P30749)	*Molybdopterin synthase catalytic subunit*	*cnxH*	*HPODL_00195* ($4e^{-15}$ - 72)	*OpCNX6*
E. coli MogA (P0AF03)	Molybdopterin adenylyltransferase	*cnxE (E)*	*HPODL_03424* ($4e^{-21}$ - 76)	*OpCNX3*

[443, 444]. Sequence analysis of the *Ec*MoaA ortholog HPODL_02673 (CnxA) indicated that an N-terminal mitochondrial signal peptide sequence had been missed in the original annotation.

Individual disruption mutants of six of the *O. parapolymorpha* candidate genes (HPODL_02673, HPODL_02674, HPODL_00948, HPODL_01640, HPODL_00195,

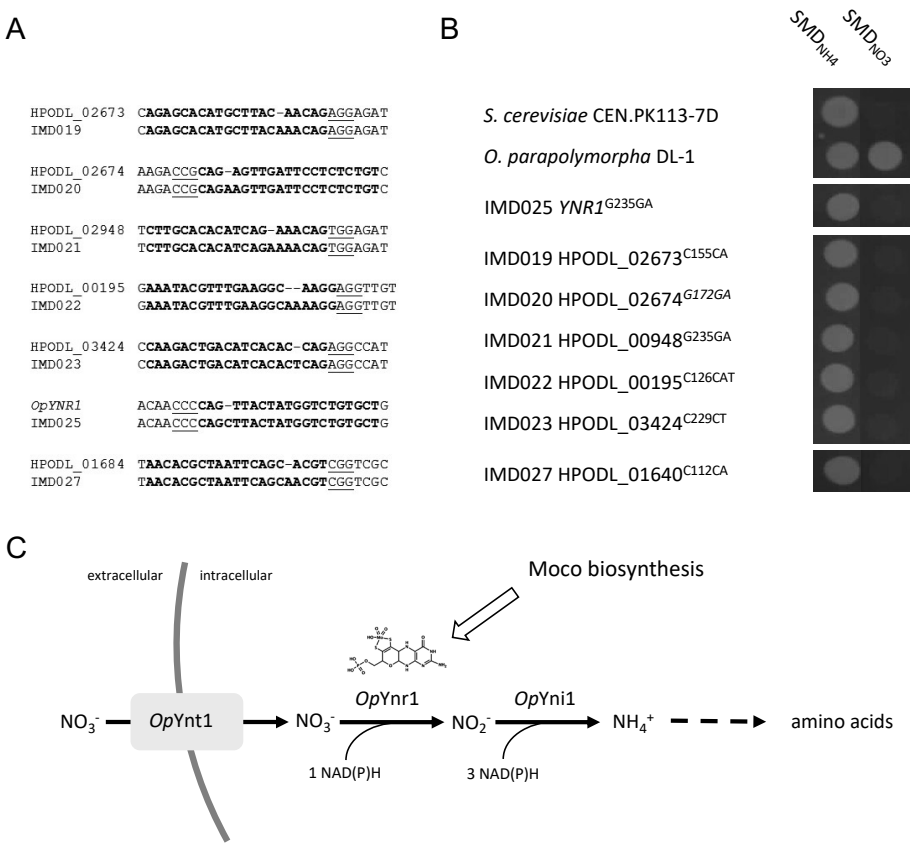


Fig. 2: Frameshift mutations in putative Moco biosynthesis genes impair nitrate assimilation in *O. parapolymorpha*. (A) Sanger-sequencing results showing the presence of -frameshift mutations in *O. parapolymorpha* strains after targeted *Spy*cas9-directed double-strand breaks in candidate Moco biosynthesis genes. The 20 bp gRNA targeting sequences are shown in bold, PAM sequences are underlined. (B) Spot plate of the wild-type *O. parapolymorpha* and mutant strains on SMD with either ammonium (SMD_{NH4}) or nitrate (SMD_{NO3}) as sole nitrogen source. As a control, the NR *OpYNR1* gene encoding nitrate reductase was also targeted and mutated. Pictures were taken after 24 h incubation at 30 °C. All strains were spotted on the same agar plate and then re-arranged in the photograph. (C) Schematic representation of the nitrate assimilation pathway including a high-affinity nitrate transporter (*Op*Ynt1), a Moco-dependent NR (*Op*Ynr1), and a nitrite reductase (*Op*Yni1). The dashed line represents multiple enzyme-catalysed reactions.

and HPODL_03424) were successfully constructed by introducing frameshifts within the first 30 % of each coding sequence with *Spy*Cas9 (Fig. 2A). Attempts to disrupt of HPODL_02128 (*OpNFS1*) were not successful, suggesting that, as in *S. cerevisiae*, *Op*Nfs1 is essential due its roles in iron-sulfur cluster biosynthesis and tRNA thiolation [301]. The ability of disruption mutants to use nitrate as sole nitrogen source was tested by spot-plate assays (Fig. 2B). The *S. cerevisiae* reference strain CEN.PK113-7D and the *O. parapolymorpha* DL-1 reference strains, as well as the NR-deficient *O. parapolymorpha* strain IMD025 (*OpYNR1*^{G235GA}) were included as controls. As expected, the DL-1 strain, but not strains CEN.PK113-7D and IMD025 grew on SMD_{NO3}.

Consistent with involvement of the six candidate genes in Moco synthesis, the corresponding *O. parapolymorpha* disruption mutants (IMD019, IMD020, IMD021, IMD022, IMD023, and IMD027) did not grow on SMD_{NO3} (Fig. 2C).

Design of Moco biosynthesis and nitrate assimilation in *S. cerevisiae*

Since *S. cerevisiae* does not naturally express molybdenum-dependent enzymes, Moco synthesis by this yeast may not only require functional expression of Moco biosynthesis genes, but also of a molybdate transporter. Heterologous genes required for Moco biosynthesis and nitrate assimilation were therefore grouped in three functional modules (1) Moco biosynthesis (Moco) comprising HPODL_02128 (*OpNFS1*), HPODL_02673 (*OpCNX1*), HPODL_02674 (*OpCNX2*), HPODL_00948 (*OpCNX4*), HPODL_01640 (*OpCNX5*), HPODL_00195 (*OpCNX6*), and HPODL_03424 (*OpCNX3*), (2) Molybdate high-affinity transport (Mo-transport) consisting of *CrMOT1* from the unicellular green alga *Chlamydomonas reinhardtii* [425] and (3) Nitrate assimilation (nitrate) comprising of *O. parapolymorpha* genes encoding a high-affinity nitrate transporter (*OpYNT1*), nitrate reductase (*OpYNR1*), and nitrite reductase (*OpYNI1*). Theoretically, an engineered *S. cerevisiae* strain expressing these three modules should be able to grow with nitrate as the sole nitrogen source. Although *S. cerevisiae* has a native *NFS1* gene, *Sc*Nfs1 predominantly localizes to the mitochondria [445]. In contrast, human Nfs1 contributes to Moco biosynthesis in the cytosol [446]. To ensure a sufficient activity of Nfs1 in the cytosol of *S. cerevisiae*, *OpNFS1* was included in the Moco module. Each module was integrated individually or in combination with other modules at the *SGA1* locus on chromosome IX in one single transformation (Figure 3A). This genomic locus has been previously shown to be a suitable integration site for single or multiple genes expression modules [277, 313].

Growth of *S. cerevisiae* on nitrate required amplification of the Moco biosynthesis and nitrate assimilation pathway genes

Transformation of the different modules resulted in *S. cerevisiae* strains IMX1777 (Moco), IMX1778 (Moco, Mo-transport), IMX1779 (Mo-transport), IMX1780 (nitrate), IMX1781 (Moco, Mo-transport, nitrate) and IMX1782 (Moco, nitrate) (Figure 3B, Figure S1-6).

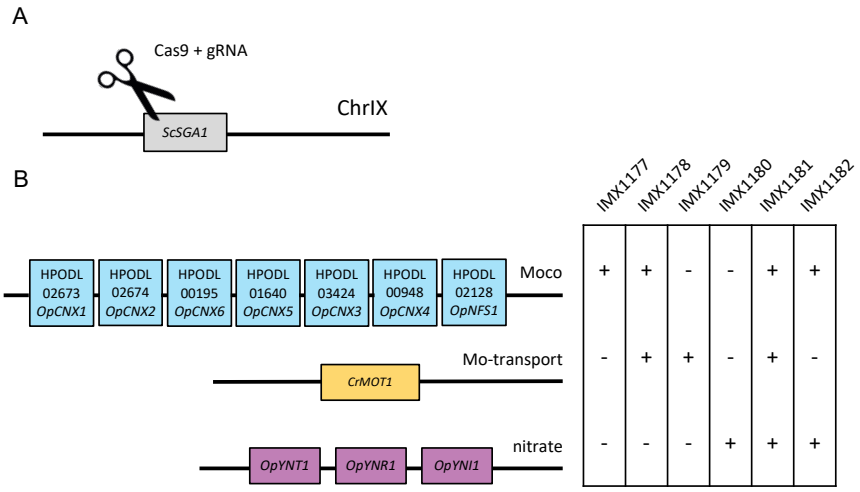


Fig.3: Schematic overview of *S. cerevisiae* strain construction. All genes were integrated by CRISPR/Cas9 in one step at the *SGA1 locus* on chromosome IX (A). One or more of the moco, Mo-transport, and nitrate modules were integrated, resulting in strains IMX1177, IMX1178, IMX1179, IMX1180, IMX1181, and IMX1182 (B).

After a short adaptive laboratory evolution of two weeks, only strains IMX1781 (Moco, Mo-transport, nitrate) and IMX1782 (Moco, nitrate) grew on synthetic medium with nitrate, indicating that next to adaptation, expression of the Moco and nitrate modules was essential for nitrate assimilation, while, under these conditions, high-affinity Mo-transport was dispensable.

To further investigate the genetic basis of this adaptation, evolved populations derived from IMX1781 and IMX1782 were each inoculated in triplicate shake-flask cultures on SM_{NO3}. After reaching stationary phase, single-colony isolates were obtained from these cultures and named IMS815-6, and IMS819 (derived from IMX1781), and IMS817-8, and IMS821 (derived from IMX1782). Whole-genome sequencing showed a disomy or trisomy of chromosome IX, which harboured the *SGA1* locus at which the heterologous genes were integrated (Fig. 4), in five of these six isolates. This change in chromosomal copy number was not observed in a culture of the parental strains IMX1781 and IMX1782 grown on complex YPD medium. Strain IMS819, which did not show aneuploidy, but had lost mitochondrial DNA and was therefore not used in further experiments because of its inability to respire.

To assess the impact of the observed changes in copy number of chromosome IX on expression levels of the heterologous proteins, strain IMX1781, which contains all three modules (Moco, Mo-transport, nitrate) and the derived isolate IMS816 were analysed by

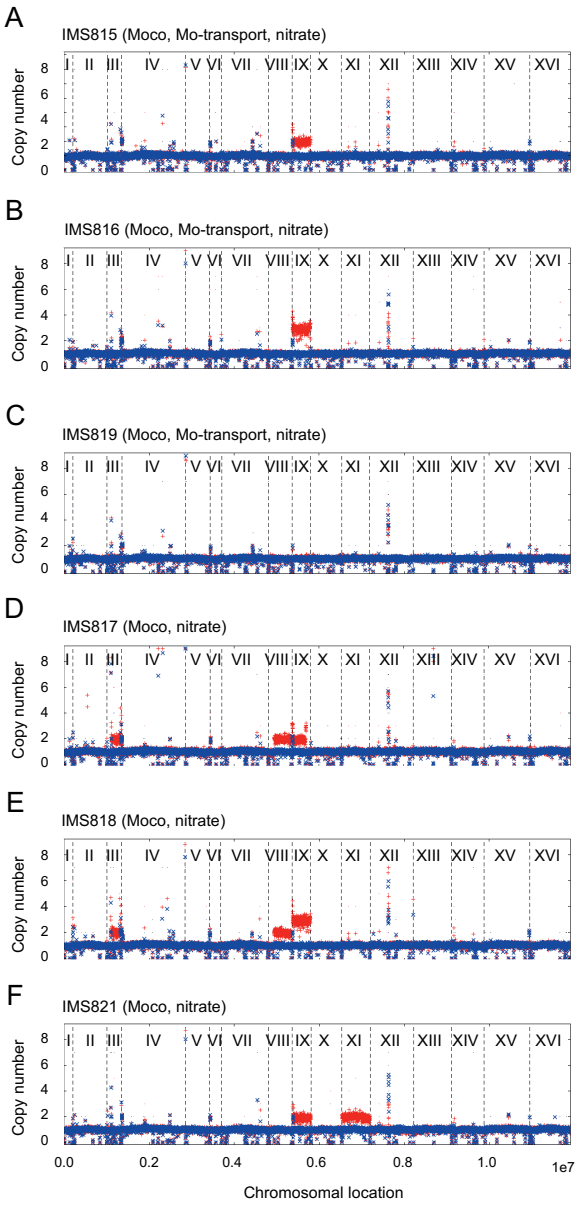


Fig. 4: Chromosomal copy number variations in engineered *S. cerevisiae* strains evolved for growth in SMD with 50 mM KNO_3 as sole nitrogen source. *S. cerevisiae* strains IMS815 (A), IMS816 (B) and IMS819 (C) were evolved starting from strain IMX1781 (Moco, Mo-transport, nitrate) while strains IMS817 (D), IMS818 (E) and IMS821 (F) were evolved starting from strain IMX1782 (Moco, nitrate). Copy numbers of chromosomes and chromosomal regions were calculated from sequence data with the Magnolya algorithm [318]. Results for the parental unevolved strain and the evolved isolate are shown in blue and red, respectively. Individual chromosomes are indicated by Roman numerals and separated by dashed lines.

untargeted proteomics (Fig. 5A). The heterologously expressed proteins were all detected in both strains, except for *Op*Nfs1 which was not detected in the unevolved strain IMX1781 (Moco, Mo-transport, nitrate). Statistical analysis of normalized peptide counts showed that levels of five Moco biosynthetic proteins (*Op*Cnx2, *Op*Cnx3, *Op*Cnx4, *Op*Cnx6, and *Op*Nfs1) were significantly higher (P-value < 0.05) in the evolved isolate IMS816 than in the parental strain IMX1781.

NADPH- and NADH-dependent NR activity was assayed in cell extracts of strains IMX1780 (nitrate), IMX1781 (Moco, Mo-transport, nitrate), and IMS816 (evolved IMX1781) (Fig. 5B). Cell extracts of strain IMX1780, which lacks the Moco module,

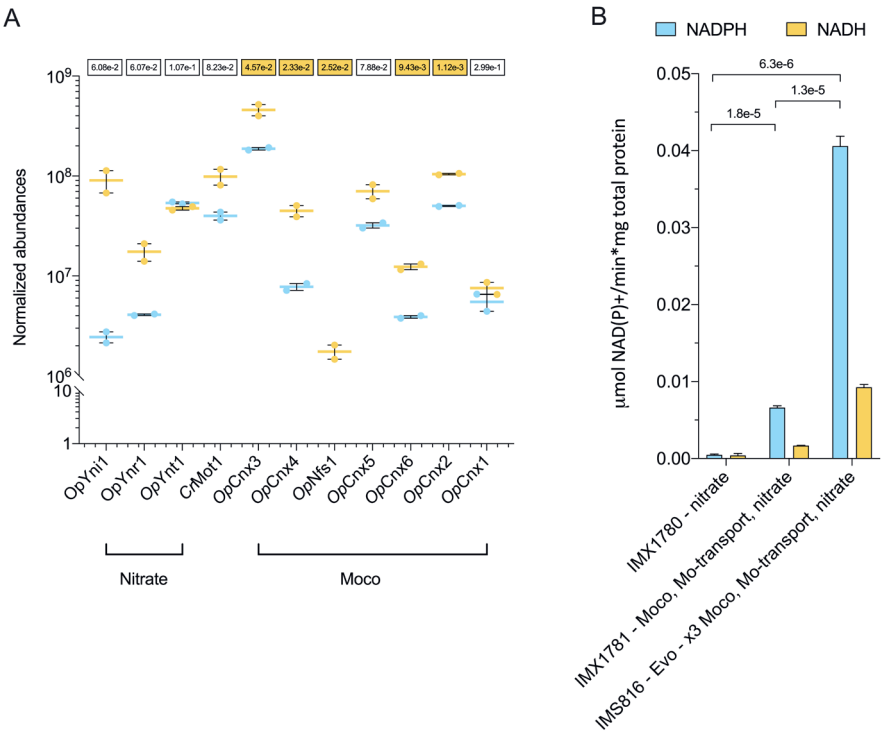


Fig. 5: Evolved nitrate-assimilating *S. cerevisiae* strains show increased Moco biosynthesis and nitrate-assimilation protein expression levels and higher *in vitro* nitrate-reductase activity. (A) Normalized abundances of heterologously expressed proteins in strains IMX1781(Moco, nitrate, Mo-transport, light blue) and IMS816 (Evolved IMX1781, x3 Moco, x3 nitrate, x3 Mo-transport, yellow) measured by LC-MS. P-values from a two-tailed Welch's t-test are shown above each tested pair and highlighted in yellow when P-value < 0.05. (B) Nitrate reductase activity in cell extracts derived from overnight cultures of IMX1780 (nitrate), IMX1781 (Moco, nitrate, Mo-transport), and IMS816 (Evolved IMX1781, x3 Moco, x3 nitrate, x3 Mo-transport) grown on SMD_{urea}. Statistical analysis was based on a two-tailed Welch's t-test and P-values are reported for tested pairs. Error bars represent the standard error of the mean of biological replicates (n=2 for panel A, n=3 for panel B).

showed no significant NR activity with either redox cofactor. In contrast, extracts from strains IMX1781 and IMS816 both showed NR activity, with a ca. 5-fold increased activity in the latter strain. Activities observed with NADPH as electron donor were approximately four-fold higher than with NADH. These results indicate that strain IMX1781 already expressed a functional nitrate assimilation pathway and that its cultivation SM_{NO3} provided a strong selective pressure for amplification of the heterologous gene cassettes, leading to increased protein expression and enzyme capacity.

Growth characteristics of engineered nitrate-assimilating *S. cerevisiae*

Specific growth rates of the evolved *S. cerevisiae* isolates IMS815-8 and IMS821 measured in shake-flask cultures on SMD_{NO3} ranged from 0.10 to 0.17 h^{-1} (Fig. 6A). These specific growth rates are two- to three-fold lower than that of a congenic reference strain on SMD with ammonium as nitrogen source [394]. Compared to natural nitrate-assimilating yeasts, *S. cerevisiae* strains IMS816 and 817 grew faster than *B. bruxellensis* CBS 2499 (specific growth rate of 0.1 h^{-1} on SMD_{NO3}) but up to 2.5-fold slower than *O. parapolymorpha* CBS 11895 (specific growth rate of 0.25 h^{-1} on SMD_{NO3}). During exponential growth on SMD_{NO3}, nitrate consumption by fastest growing nitrate-assimilating *S. cerevisiae* strains IMS816 (evolved IMX1781, Moco, Mo-transport, nitrate) IMS817 (evolved IMX1782, Moco, nitrate) occurred without detectable accumulation of either nitrite or ammonium (Fig. 6B-C). Release of small amounts of ammonium in late stationary phase cultures was tentatively attributed to protein turnover and/or cell lysis.

To test whether expression of a high-affinity Mo-transporter was essential at low extracellular molybdate concentrations, strains IMS816 and IMS817 were inoculated in SMD_{NO3} with a 100-fold lower MoO_4^{2-} concentration than the reference medium (16 nM instead of 1.6 µM, Fig.6D). After two weeks of incubation, only strain IMS816 started growing on the low-molybate medium and, upon transfer to the same medium, instantaneously grew exponentially at a rate of 0.11 ± 0.01 h^{-1}. This observation indicated that, after an adaptation period, the Mo-transport module was required for growth at low molybdate concentrations.

The ability of strain IMS817 (Moco, nitrate) to co-consume nitrate and ammonium was tested in shake-flask cultures on SMD_{AN}, which contained 10 mM NH_4NO_3 as nitrogen source (Fig. 6E). Although ammonium and nitrate were consumed at different rates, nitrate and ammonium were consumed simultaneously. Nitrate was completely consumed (residual concentration < 0.1 mM) and a high specific growth rate (0.30 ± 0.01 h^{-1}) was observed throughout the exponential growth phase.

In addition to Moco, NR requires a flavin adenine dinucleotide and heme *b* as cofactors [447, 448]. In *S. cerevisiae*, heme *b* is synthesised via an oxygen-dependent pathway [449, 450]. To test whether the nitrate-assimilating *S. cerevisiae* strain IMS816 (Moco, Mo-

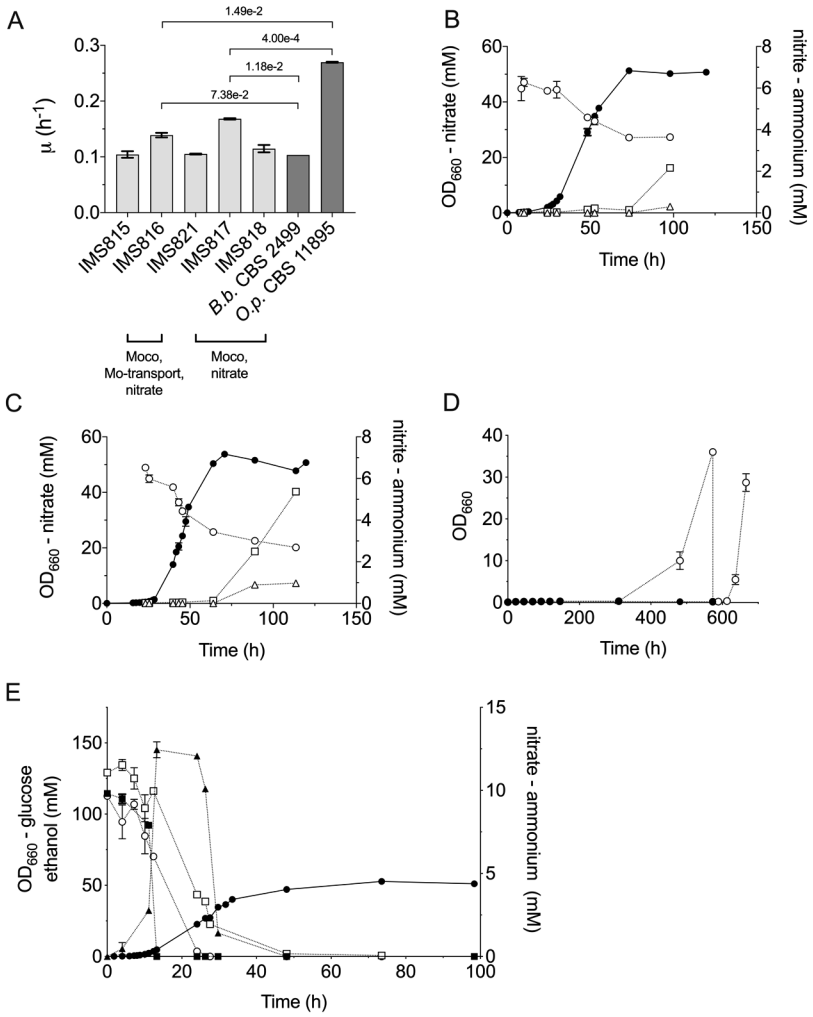


Fig. 6: Aerobic characterization of engineered and evolved nitrate-assimilating *S. cerevisiae* strains. (A) Specific growth rates in aerobic shake-flask cultures of evolved *S. cerevisiae* strains IMS815, IMS816, IMS817, IMS818, IMS821, *B. bruxellensis* CBS 2499, and *O. parapolymorpha* CBS 11895 on SMD_{NO3}. Growth curves of aerobic shake-flask cultures of *S. cerevisiae* strains IMS816 (Evolved IMX1781, x3 Moco, x3 nitrate, x3 Mo-transport, B) and IMS817 (Evolved IMX1782, x2 Moco, x2 nitrate, C) in SMD_{NO3}. Symbols indicate biomass (●) and nitrate (○), nitrite (□) and ammonium (△). (D) Growth curves of IMS816 (Evolved IMX1781, x3 Moco, x3 nitrate, x3 Mo-transport, ○) and IMS817 (Evolved IMX1782, x2 Moco, x2 nitrate, ●) in $SMD_{NO3\text{-}LowMo}$ containing 16 nM MoO_4^{2-}. (E) Growth curve in aerobic shake-flask cultures of *S. cerevisiae* IMS817 (Evolved IMX1782, x2 Moco, x2 nitrate) on SMD_{AN} containing 10 mM NH_4NO_3 as nitrogen source. Symbols indicate OD_{660} (●), glucose (■), ethanol (▲), ammonium (○), and nitrate (□). Statistical analysis was based on a two-tailed Welch's t-test and P-values are reported for tested pairs. Error bars represent the standard error of the mean of independent cultures (n=3 except for panel D and CBS 2499, and CBS 11895 in panel A where n=2).

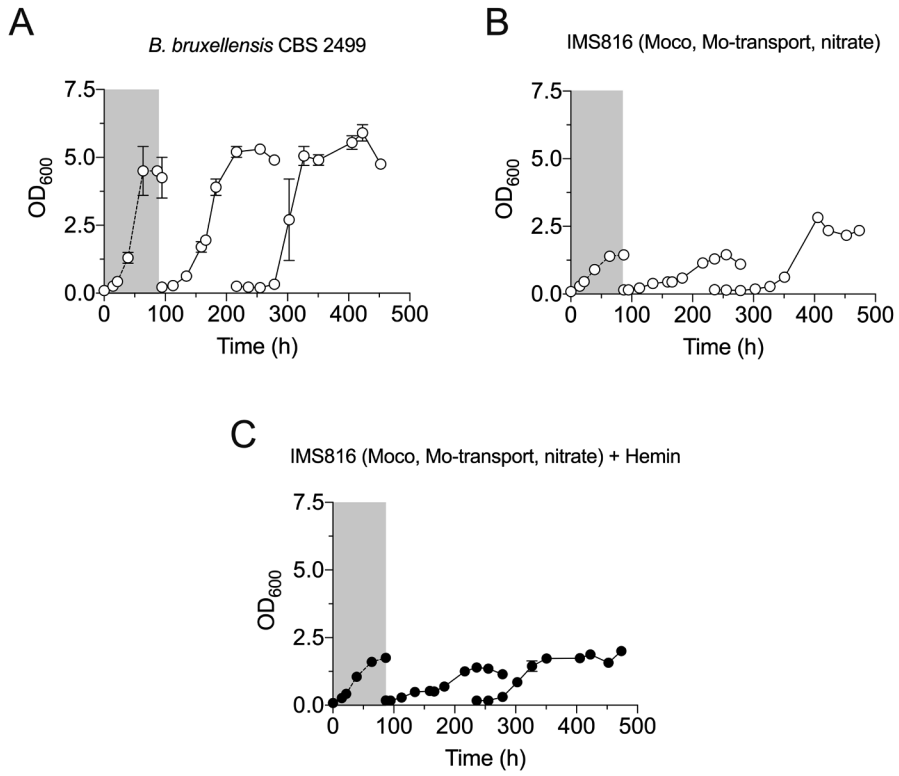


Fig. 7: Anaerobic growth of the engineered and evolved nitrate-assimilating *S. cerevisiae* strain IMS816 and *B. bruxellensis* CBS2499 on glucose synthetic medium with nitrate as sole nitrogen source. Strain CBS 2499 (A) and IMS816 (Evolved IMX1781, x3 Moco, x3 nitrate, x3 Mo-transport, B-C) were sequentially transferred in SMD_{NO3} supplemented with Tween 80 and ergosterol with (●) or without (○) addition of hemin. A first anaerobic batch-cultivation cycle to deplete possible heme *b* introduced with the aerobically pre-grown inoculum is highlighted by a grey box. The absence of oxygen leaks in the anaerobic chamber was verified by near-absence of growth of *S. cerevisiae* CEN.PK113-7D on SMD_{Urea} without supplementation of the anaerobic growth factors Tween 80 and ergosterol (Supplementary Figure S7). Error bars represent the standard error of the mean of independent cultures (n=2).

transport, nitrate, evolved) was nevertheless able to assimilate nitrate under anaerobic condition, its growth on SMD_{NO3} was studied in an anaerobic chamber. The wild-type *B. bruxellensis* strain CBS 2499, which was previously reported to grow anaerobically on nitrate [451] was included as a reference. Although *B. bruxellensis* CBS 2499 reproducibly showed anaerobic growth after three consecutive transfers in SMD_{NO3}, the cultures reached only low OD values. *S. cerevisiae* strain IMS816 grew slower and reached final OD values that were over two-fold lower than those observed in anaerobic cultures of *B. bruxellensis* CBS 2499 (Fig. 7A-B). Supplementation of 32 mg/L of hemin (Fe^{3+}-containing protoporphyrin IX) , which can be imported by *S. cerevisiae* when grown anaerobically [452], did not result in faster anaerobic growth of strain IMS816 on SMD_{NO3} (Fig. 7C). This result indicates that the heme *b* is not the only limiting factor in the tested conditions and time frames [452].

Competition of nitrate-assimilating *S. cerevisiae* and the spoilage yeast *B. bruxellensis* in nitrate-containing media

B. bruxellensis strains are common yeast contaminants in bioethanol plants [422, 453, 454]. Their spoilage phenotype has been related to utilization of nutrients in industrial media that cannot be metabolized by *S. cerevisiae* [455, 456]. Plant biomass-derived substrates, such as the sugarcane juice used in Brazilian bioethanol processes, contains nitrate [421]. In such settings, the ability to (co-)consume nitrate may confer a competitive advantage with *S. cerevisiae* [421]. To evaluate the relative fitness of the engineered nitrate-assimilating *S. cerevisiae* strain IMS816 and *B. bruxellensis* CBS 2499, they were co-cultured in serial aerobic batch cultures on SMD_{NO3}. After inoculation at a *S. cerevisiae*:*B. bruxellensis* ratio of at least 6:4, based on the colonies ratios at time 0, cultures were sequentially transferred to fresh medium at 72 h intervals. As a control, a similar experiment was performed starting with a 9:1 mixture of the nitrate-non-assimilating reference strain *S. cerevisiae* IMX585 and *B. bruxellensis* CBS 2499. At the onset of each cultivation cycle, samples were plated on SMD with either ammonium or nitrate as nitrogen source, using Bromocresol Green for differential staining of the two species (Figure 8A). In the control cultures, the relative abundance of *S. cerevisiae* IMX585 dropped below 10 % after the first cultivation cycle and below detection level after the second transfer. In contrast, in a co-culture of the nitrate-assimilating *S. cerevisiae* strain IMS816 and *B. bruxellensis* CBS 2499, the *S. cerevisiae* strain persisted in the co-culture for about 35 generations (Fig. 8B).

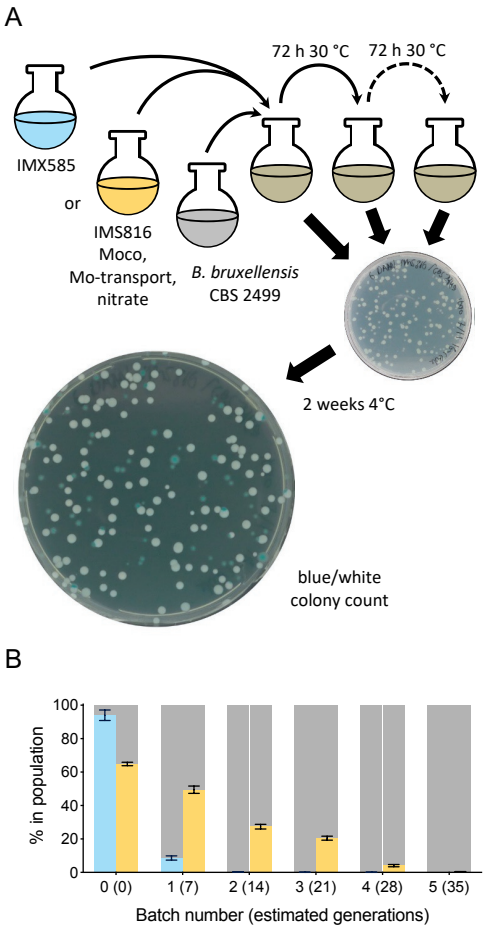


Fig. 8: The engineered nitrogen-assimilating *S. cerevisiae* strain IMS816 shows increased competitiveness in co-cultures with the spoilage yeast *B. bruxellensis*. (A) Schematic representation of the co-cultivation experiment. Either the reference strain *S. cerevisiae* IMX585 (light blue) or strain IMS816 (Evolved IMX1781, x3 Moco, x3 nitrate, x3 Mo-transport, yellow) was co-inoculated with *B. bruxellensis* CBS 2499 (grey) and grown in serial batch cultures. Before each transfer, cells were plated on both $SMD_{NO3\text{-}blue}$ and $SMD_{Amm\text{-}blue}$ agar plates. After 48 h at 30 °C, followed by 2 weeks at 4 °C to selectively stain CBS 2499 colonies, *S. cerevisiae* (white) and *B. bruxellensis* (blue) colonies were counted. (B) Development over time of the percentage of IMX585 (light blue) or IMS816 (Evolved IMX1781, x3 Moco, x3 nitrate, x3 Mo-transport, yellow) *S. cerevisiae* cells relative to total cell count (grey) of co-cultures with *B. bruxellensis* CBS 2499. Error bars represent the standard error of the mean of independent co-cultures (n=3).

Discussion

Over 50 molybdenum-cofactor (Moco) containing enzymes, mostly from prokaryotes, have been characterized and catalyse redox reactions in the global cycles of nitrogen (e.g. nitrate reductase, nitrite oxidase), sulfur (e.g. sulfite oxidase, DMSO-reductase) and carbon (e.g. CO dehydrogenase, aldehyde oxidases, formate dehydrogenase) [265]. More molybdo-enzymes are likely to be discovered as part of the ongoing exploration of microbial diversity. Synthesis of a functional Moco in *S. cerevisiae* represents an essential step towards accessing this diverse group of enzymes for metabolic engineering strategies in this platform organism. In addition to the use of nitrate as a nitrogen source, such strategies could, for example, involve high-k_{cat} molybdoprotein formate dehydrogenase [457] as alternative for the low-k_{cat} native fungal formate dehydrogenases

and molyboprotein furoyl-CoA dehydrogenase, which can contribute to conversion of furanic compounds found in lignocellulosic hydrolysates [458]. Full exploration of these possibilities will require expansion of the range of Moco variants that can be expressed by *S. cerevisiae*.

Recent genome sequence based phylogeny studies showed that fewer than 27 % of the 329 sequenced Saccharomycotina yeast species genomes harbour Moco biosynthesis genes and that only 13 % harbour nitrate-reductase genes [266]. The exact gene complement required for Moco synthesis in yeasts and other organisms has not previously been defined. The gene set identified in this study, based on a combination of mutational analysis in *O. parapolymorpha* and heterologous expression in *S. cerevisiae*, provides a basis for further investigation and engineering of fungal Moco biosynthesis. In this context, subcellular compartmentation of Moco synthesis deserves special attention. In eukaryotes, the molybdopterin biosynthesis intermediate cyclic pyranopterin monophosphate (cPMP) is synthesized in the mitochondrial matrix [412] (Fig. 1) and subsequently translocated to the cytosol for further processing. In *Arabidopsis thaliana*, *At*Atm3, a mitochondrial transporter involved in Fe-S cluster translocation [459], has been proposed to also transport cPMP [444]. *In vitro* and *in vivo* functionality of nitrate reductase in engineered strains suggests that *S. cerevisiae* can export cPMP from mitochondria. By analogy to the situation in *A. thaliana*, the mitochondrial ATP-binding cassette (ABC) transporter *Sc*Atm1, which is involved in transport of iron-sulfur (Fe/S) clusters precursors to the cytosol and essential for aerobic growth [460, 461], is a promising candidate for this role.

Despite the use of strong promoters to drive expression of *O. parapolymorpha* nitrate-assimilation and Moco-biosynthesis genes, growth of engineered *S. cerevisiae* strains on nitrate as sole nitrogen source reproducibly selected for mutants in which single-chromosome disomy or trisomy caused increased copy numbers of these genes. These increased copy numbers coincided with higher abundances of all encoded heterologous proteins (Fig. 4). Consequently, it is not possible to unequivocally identify which protein(s) exerted the strongest control on *in vivo* rates of nitrate reduction. However, only HPODL_02128 (OpNfs1) was not detected prior to gene amplification and may therefore be a priority target in follow-up research. Nfs1 is a cysteine desulfurase involved in iron-sulfur cluster (Fe-S) biogenesis and, in *S. cerevisiae*, is almost exclusively located in the mitochondria [445]. In *O. parapolymorpha*, cytosolic Nfs1 is also required to re-load sulfur on the molybdopterin-synthase adenylyltransferase *Op*Cnx4 (HPODL_00948) via a sulfur mobilization route shared with tRNA thiolation [412, 443, 446].

Earlier reports proposed *S. cerevisiae* as a platform for molybdate import studies due to absence of native high-affinity molybdate transporters [425]. Our results show that, although it does not naturally express molybdoproteins, *S. cerevisiae* can take up MoO_4^{2-} at micromolar concentrations. A hypothesis that molybdate is transported by the sulfate

transporters Sul1 and Sul2 was supported by the observation that expression of the plant sulfate transporter SHST1 enabled high-affinity molybdate import in *S. cerevisiae* [462, 463]. The demonstration that expression of a heterologous high-affinity transporter is required for growth at nM concentrations of molybdate may be relevant for application of Moco-expressing strains in feedstocks that contain extremely low molybdate concentrations.

After an earlier unsuccessful attempt to express a *Nicotiana tabacum* nitrate reductase [464], this study is the first to demonstrate nitrate assimilation by an engineered *S. cerevisiae* strain. In contrast to most naturally nitrate-assimilating fungi [465-467], the engineered nitrate-assimilating *S. cerevisiae* strains did not exhibit ammonium repression of nitrate assimilation and co-consumed both nitrogen sources during fast aerobic growth on an ammonium-nitrate mixture. Hydrolysates of corn, corn stover and switchgrass used as feedstocks for yeast-based bioethanol production contain low but significant amounts of nitrate, whose discharge can have negative environmental consequences [468]. The low rates of nitrate consumption by anaerobic cultures of nitrate-assimilating *S. cerevisiae* strains described in this study may already suffice to eliminate small amounts of nitrate and thereby contribute process sustainability. For more extensive use of nitrate as a nitrogen source in anaerobic bioethanol production processes, for instance to reoxidize cytosolic NADH and thus reduce formation of glycerol as a byproduct [469], further research is needed to improve anaerobic nitrate reduction by engineered strains.

The ability of the spoilage yeast *B. bruxellensis* to assimilate nitrate is frequently cited as explanation for contamination *S. cerevisiae* sugar-cane juice fermentations [421, 422]. Our experiments with laboratory co-cultures demonstrate that, indeed, engineering of nitrate assimilation into *S. cerevisiae* can positively influence competition with *B. bruxellensis*. However, despite a slightly higher growth rate on nitrate of the engineered *S. cerevisiae* strain as compared to *B. bruxellensis* CBS 2499 (0.14 ± 0.01 h^{-1} versus 0.10 ± 0.01 h^{-1}), it was eventually still outcompeted by the spoilage yeast. Although there is a small difference in growth rates when only nitrate is provided, it is worth noticing that the similar evolved isolate IMS817 grew with a growth rate of 0.30 ± 0.01 in SMD_{AN} where both ammoninum and nitrate were present. This value is much higher than what reported for *B. bruxellensis* in a similar medium (0.077 ± 0.004 h^{-1}) [470] and indicates that a nitrate assimilating *S. cerevisiae* strain may persist even longer in a co-culture experiment where a media with both nitrogen sources is used. It was previously reported that dominance of *B. bruxellensis* seems not only related to its ability to assimilate nitrate but also to its higher affinity glucose importers [471, 472]. This suggests that additional engineering is required to further increase *S. cerevisiae* competitiveness.

Further optimization of the kinetics of nitrate uptake and assimilation under the anaerobic conditions prevalent in industrial bioethanol production, combined with growth

experiments on industrial media, are required to assess the full potential of this approach.

Acknowledgments

We thank Maxime den Ridder and Dr. Martin Pabst for helping us with the proteomics sample preparation and Dr. Irina Borodina and Dr. Tune Wulff for helping us with the proteomics sample analysis. This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie action PAcMEN (grant agreement No 722287). JTP. acknowledges support by an Advanced Grant of the European Research Council (grant # 694633). TP, JTP and J-MGD. are inventors on a patent application related to this work (WO2020209718 – Yeast with engineered molybdenum co-factor biosynthesis). The remaining authors have no competing interests to declare.

Chapter 5: Engineering of molybdenum cofactor-dependent nitrate assimilation in *Yarrowia lipolytica*

Thomas Perli, Irina Borodina, Jean-Marc Daran

Abstract

Engineering a new metabolic function in a microbial host can be limited by the availability of the relevant co-factor. For instance, in *Yarrowia lipolytica*, the expression of a functional nitrate reductase is precluded by the absence of molybdenum cofactor (Moco) biosynthesis. In this study, we demonstrated that the core Moco biosynthesis pathway from *Ogataea parapolymorpha* associated with the expression of a high affinity molybdate transporter could lead to the synthesis of Moco in *Y. lipolytica*. This was achieved by coupling Moco biosynthesis to the Moco-dependent nitrate assimilation pathway of the same donor *O. parapolymorpha*. In addition to 11 heterologous genes, fast growth on nitrate required adaptive laboratory evolution which, resulted in up to 100-fold increase in nitrate reductase activity and in up to 4-fold increase in growth rate. Genome sequencing of evolved isolates revealed the presence of a limited number of non-synonymous mutations or small insertions/deletions in annotated coding sequences. This study that builds up on a previous work establishing Moco synthesis in *S. cerevisiae* demonstrated that the Moco pathway could be successfully transferred in very distant yeasts and, potentially, to any other genera, which would enable the expression of new enzyme families and expand the nutrients range used by industrial yeasts.

Introduction

The Dipodascaceae yeast *Yarrowia lipolytica* presents as with *S. cerevisiae* a double interest as a model yeast for dimorphism studies and as an industrial work horse [473]. This strictly aerobic oleaginous Saccharomycetales yeast has been traditionally exploited for its ability to efficiently degrade a wide variety of abundant and cheap hydrophobic substrates such as n-alkanes, fatty acids and oils that was coupled to its remarkable high enzyme secretion capacity and production of organic acids such as citric acid and α-ketoglutarate [474].

The fast development of dedicated molecular tools including the addition of CRISPR technology, enabled to propel *Y. lipolytica* as a potential contender of *S. cerevisiae* for the biosynthesis of commodity and specialities chemicals. *Y. lipolytica* has become a reliable platform for metabolic engineering as illustrated by the synthesis of flavour compounds (e.g. γ-Decalactone and γ-Dodecalactone) [475], insect sex pheromones (e.g. (Z)-hexadec-11-en-1-ol) as pest biocontrol [476] and rebaudioside A, a steviol glycoside used as a non-caloric sweetener [477], as most prominent examples.

Engineering *of Y. lipolytica* with increasingly more complex pathway will require the expression of an even broader range of enzymes as its attractiveness as metabolic engineering platform grows. Many enzyme activities require the presence of one or more essential cofactors [226, 403]. Therefore, the successful expansion of the enzyme repertoire in a microbial host may require the parallel broadening of its cofactor set. Whenever a cofactor requirement cannot be met by media supplementation because either (1) the cofactor is not commercially available, (2) is too unstable, or (3) cannot be imported by the organism, metabolic engineering is required to enable its *de novo* biosynthesis or its transport. This approach was successful in the model yeast *Saccharomyces cerevisiae* as exemplified with the implementation of high affinity Ni^{2+}transport, an inorganic cofactor of Ni-dependent urease [227], or with the engineering of tetrahydrobiopterin pathway, that was instrumental in the implementation of *de novo* biosynthesis of opioids [258, 404]. More recently, the molybdenum co-factor biosynthesis pathway from the methylotrophic yeast *Ogataea parapolymorpha* was introduced in *S. cerevisiae* allowing expression of a functional Moco dependent nitrate reductase that could support growth on media containing nitrate as sole nitrogen source [478]. Although these pioneering studies demonstrated how metabolic landscape could be expanded beyond the natural ability of a microorganism, these approaches have not yet been transposed in other industrially relevant yeast species.

As *S. cerevisiae* and more than 75% of the Saccharomycotina yeast species, *Y. lipolytica* is not able to synthesise molybdopterin cofactor [266] which preclude the possibility to harness enzyme families of biotechnological relevance.

Moco-dependent enzymes encompass more than 30 different catalytic activities, that have been divided in three main families based on the Moco variant they require [414, 418].

In a large majority, Moco-dependent enzymes catalyse oxido/reduction reactions, often involving oxygen, and are implicated in nutrient e.g. carbon, nitrogen and sulfur cycles or in detoxification of growth inhibiting compounds thanks to the redox versatility of the Mo atom [413].

Moco is composed of a molybdate (MoO_4^{2-}) oxyanion coordinated by two sulfur atoms on a tricyclic pterin scaffold called molybdopterin (MPT). Moco cannot be supplemented in the media since the molecule is too unstable due to its oxygen sensitivity [479]. For this reason, Moco is, in the majority of cases, *de novo* synthesized intracellularly. The Moco biosynthetic pathway is very well conserved and it has been extensively studied in both prokaryotes and eukaryotes [405, 408, 409]. The first step, which takes place in the mitochondria of eukaryotic cells, is catalysed by the heterodimer Cnx1/Cnx2, that circularises GTP onto cyclic pyranopterin monophosphate (cPMP). The molecule is then exported to the cytosol through a yet uncharacterized transporter, and it is converted to MPT by action of the MPT synthase complex ($Cnx5_2/Cnx6_2$) which donates two sulfur atoms present on a conserved Cysteine residue in the Cnx5 protein. Cnx5 is then reloaded with sulfur via a sulfur mobilization route that includes the adenyltransferase Cnx4 and a cysteine desulfurase that is typically involved in iron-sulfur cluster biosynthesis and tRNA thiolation, Nfs1 [412]. In the final step, molybdate is inserted in a two-steps reaction catalysed by the multi-domain protein Cnx3 [409]. Previous studies showed that the yeast *S. cerevisiae* lacks a high-affinity molybdate transport system and that it is able to import the oxyanion only with low affinity, unless a high affinity transporter such as *Cr*Mot1 from *Chlamydomonas reinhardtii* is expressed [425, 478]. Moco can be further modified by either sulfuration of the molybdate ion or, in prokaryotes, by the covalent attachment of either GDP or CDP to MPT.

The goals of this study were to investigate whether the heterologous Moco biosynthetic pathway from the nitrate assimilating yeast *O. parapolymorpha* could be functionally engineered in *Y. lipolytica*, together with a Moco-dependent nitrate assimilation pathway. To this end, not fewer than 11 genes encompassing Moco biosynthesis, molybdate transport and nitrate reduction functions were introduced in the oleaginous yeast *Y. lipolytica* using CRISPR/Cas9 gene-editing technology. The engineered strain was then subjected to adaptive laboratory evolution in synthetic media with nitrate as the nitrogen source to evolve a fast-growing population. After that, single cell lines were isolated, phenotyped on nitrate containing medium and characterized by whole-genome resequencing.

Material and Methods

Strains, media and maintenance

All strains used and constructed in this study are shown in Table 1. All *Y. lipolytica* strains were derived from the strain ST6512 (W29, MATa *ku70Δ::Spycas9-EcDsdAMX4*; [475]. Yeast strains were grown on either YP (10 g/L Bacto yeast extract, 20 g/L Bacto peptone) or SM medium [13] with either 5 g/L KNO_3, or 2.3 g/L urea (SM_{NO3} and SM_{urea}, respectively) as sole nitrogen source. In all SM media variants, 6.6 g/L K_2SO_4 was added as a source of sulfate [428]. YP or SM media were autoclaved at 121 °C for 20 min. After sterilization, SM was supplemented with 1 ml/L of filter-sterilized vitamin solution as previously described {Verduyn, 1992 #793}. A concentrated glucose solution was autoclaved at 110 °C for 20 min and then added to the YP and SM medium at a final concentration of 20 g/L, yielding SMD and YPD, respectively. 500-ml Shake flasks containing 100 mL medium and 100-mL shake flasks containing 20 mL medium were incubated at 30 °C and 200 rpm in an Innova Incubator (Brunswick Scientific, Edison, NJ). Solid media were prepared by adding 1.5 % (w/v) Bacto agar and, where indicated, 250 mg/L nourseothricin or 250 mg/L hygromycin B. *Escherichia coli* strains were grown in LB (10 g/L Bacto tryptone, 5 g/L Bacto yeast extract, 5 g/L NaCl) supplemented with 100 mg/L ampicillin. *Y. lipolytica* and *E. coli* cultures were stored at -80 °C after the addition of 30 % v/v glycerol.

Molecular biology techniques

Primers used in this study are shown in Table 2. DNA was amplified using either Phusion Hot Start II High Fidelity Polymerase (Thermo Fisher Scientific, Waltham, MA) or Phusion U (Thermo Fisher Scientific) and desalted or PAGE-purified oligonucleotide primers (Sigma-Aldrich) according to manufacturers' instructions. Diagnostic PCR reactions were performed with DreamTaq polymerase (Thermo Fisher Scientific). PCR products were separated by gel electrophoresis on a 1 % (w/v) agarose gel (Thermo Scientific) in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA; Thermo Scientific) and purified with a Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA). Plasmids were isolated from *E. coli* using a NucleoSpin Plasmid kit (Macherey-Nagel, Düren, Germany), and verified by either restriction digestion or diagnostic PCR. *E. coli* DH5-α (New England BioLabs, Ipswich, MA) was used for transformation [316]. Yeast genomic DNA used for diagnostic PCR reactions was isolated by using the SDS/LiAc protocol [314]. *Y. lipolytica* transformation was performed with the LiAc method as previously described [480]. Four to eight colonies were re-streaked on selective medium to select for single clones and diagnostic PCRs were performed to verify the correct genotypes.

Table 1: Strains used in this study.

Name	Relevant genotype	Parental strain	Reference
ST6512	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2*	W29, Y-63746, A T T C - 20460	[475]
IMX2264	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2*	ST6522	This study
IMX2265	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10*	IMX2264	This study
IMX2266	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2*	IMX2265	This study
IMX2267	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10*	IMX2566	This study
IMX2565	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP2*	IMX2267	This study
IMS1174	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP2* (Evolved on SMD_{NO3} for 21 transfers. Line 2 - Colony 1)	IMX2565	This study

IMS1175	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP2* Evolved on SMD_{NO3} for 21 transfers. Line 2 - Colony 2	IMX2565	This study
IMS1176	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2*E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP2* Evolved on SMD_{NO3} for 21 transfers. Line 2 - Colony 3	IMX2565	This study
IMS1177	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP2* Evolved on SMD_{NO3} for 21 transfers. Line 3 - Colony 1	IMX2565	This study
IMS1178	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 p GPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP2* Evolved on SMD_{NO3} for 21 transfers. Line 3 - Colony 2	IMX2565	This study
IMS1179	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP2* Evolved on SMD_{NO3} for 21 transfers. Line 3 - Colony 3	IMX2565	This study
IMS1180	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP2* Evolved on SMD_{NO3} for 21 transfers. Line 1 - Colony 1	IMX2565	This study

IMS1181	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP* Evolved on SMD_{NO3} for 21 transfers. Line 1 - Colony 2	IMX2565	This study
IMS1182	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP* Evolved on SMD_{NO3} for 21 transfers. Line 1 - Colony 3	IMX2565	This study
IMS1183	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP* Evolved on SMD_{NO3} for 50 transfers. Evolved population line 2	IMX2565	This study
IMS1184	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP* (Evolved on SMD_{NO3} for 50 transfers. Evolved population line 3)	IMX2565	This study
IMS1185	MATa *ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2* E_4Δ::*pTEFin-OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2* C_2Δ::*pTEFin-OpNFS1-tPEX10 pGPD-OpCNX3_tLIP2* E_1Δ::*pTEFin-OpCNX3-tPEX10 pGPD-CrMoT1_tLIP2* C_3Δ::*pTEFin-OpYNR1-tPEX10 pGPD-OpYNT1_tLIP2 pTEFin-OpYNI1-tPEX10* E_3Δ::*pTEFin-OpCNX6-tPEX10 pGPD-OpCNX5_tLIP* Evolved on SMD_{NO3} for 50 transfers. Evolved population line 1	IMX2565	This study

Table 2: Primers used in this study.

Primer number	Primer sequence	Product(s)
22956	AGTACTGCAAAAAGUGCTGGTCGG	PrTEFin-PrGPD_USER_Biobrick
24013	ATCAGTAGCUAGAGACCGGGTTGGCGGCG	PrTEFin-PrGPD_USER_Biobrick
15529	AGCTACTGAUGACGCAGTAGGATGTCCTGCACGG	PrTEFin-PrGPD_USER_Biobrick
15528	ATGACAGAUTGTTGATGTGTGTTTAATTCAAGAATG	PrTEFin-PrGPD_USER_Biobrick
27208	ACACGCGAUAGAGACCGGGTTGGCGG	PrTEFin_USER_Biobrick
22956	AGTACTGCAAAAAGUGCTGGTCGG	PrTEFin_USER_Biobrick
24479	ACTTTTTGCAGTACUAACCGCAGCCCGTGCGACACCTG	*OpCNX1*_USER_Biobrick
24480	CGTGCGAUTTAGCCGCCGATCAGG	*OpCNX1*_USER_Biobrick
24481	ATCTGTCAUGCCACAATGGTGGCCATCCACG	*OpCNX2*_USER_Biobrick
24482	CACGCGAUTTACTTGAAGATGGTAGACAGGTCG	*OpCNX2*_USER_Biobrick
24483	ACTTTTTGCAGTACUAACCGCAGTACCGATTCCGAATTG-GAGC	*OpNFS1*_USER_Biobrick
24484	CGTGCGAUTTAGTGTCCGGCCCACTC	*OpNFS1*_USER_Biobrick
24485	ATCTGTCAUGCCACAATGTCTCTGTCTCTGAACGAGTAC	*OpCNX4*_USER_Biobrick
24486	CACGCGAUTTAGTAGATGGGGAAGTTAGGGTC	*OpCNX4*_USER_Biobrick
24487	ACTTTTTGCAGTACUAACCGCAGTCTATCTTCGTGGACAT-CACC	*OpCNX6*_USER_Biobrick
24488	CGTGCGAUTTAGGTTCGAGACAGCACG	*OpCNX6*_USER_Biobrick
24489	ATCTGTCAUGCCACAATGGTGGCCGTGGC	*OpCNX5*_USER_Biobrick
24490	CACGCGAUTTAGCCAGAGGACACAGGAG	*OpCNX5*_USER_Biobrick
24491	ACTTTTTGCAGTACUAACCGCAGGCCCTGCAGAACGCC	*CrMoT1*_USER_Biobrick
24492	CGTGCGAUTTAGGCTCGGCCGC	*CrMoT1*_USER_Biobrick
24493	ATCTGTCAUGCCACAATGACCGTGGGCATCC	*OpCNX3*_USER_Biobrick
24494	CACGCGAUTTACACGTAGATCTGGTCGATC	*OpCNX3*_USER_Biobrick
24495	ACTTTTTGCAGTACUAACCGCAGGACTCTGTGGTGAC-CGAGGT	*OpYNR1*_USER_Biobrick
24496	CGTGCGAUTTAGAAGTAGACCACGTACTGCTTGTC	*OpYNR1*_USER_Biobrick
24497	ATCTGTCAUGCCACAATGCGACTGTCTACCCTGTG	*OpYNT1*_USER_Biobrick
24498	CACGCGAUTTAGATCTCGGCCTTTCGG	*OpYNT1*_USER_Biobrick
24499	ACTTTTTGCAGTACUAACCGCAGACCTGCTCTGTGCCTC	*OpYNI1*_USER_Biobrick
24500	CGTGCGAUTTACCAGTCGAAAGAGATGGC	*OpYNI1*_USER_Biobrick

24517	TAGATAAATTTACACTCCCTCAGATGCATTCTTGGGCGGT	pCfB9006-7_backbone Gibson_fragment
24518	TCATGGGCCTTCCTTTCACTCAGATGCATTCTTGGGCGGT	pCfB9006-7_2-genes_insert Gibson_fragment
24520	ACCGCCCAAGAATGCATCTGAGGGAGTGTAAATTTATC-TATACAGAGGTAA	pI774_GFPmut3b_spacer Gibson_fragment
24521	ACCGCCCAAGAATGCATCTGAGTGAAAGGAAGGCCCATGA	pI774_GFPmut3b_spacer Gibson_fragmen
24522	TTCATTCATGTTAGTTGCGTTCTGCGTCTGCTGTTTGTGTC	pCfB9006_backbone Gibson_fragment
24523	ACACAAACAGCAGACGCAGAACGCAACTAACATGAAT-GAATACGATATACA	pCfB9006_2-genes_insert Gibson_fragment
24524	TTCATTCATGTTAGTTGCGTGCCATAGCACTATTGTA-GAGTGGCC	pCfB9007_backbone Gibson_fragment
24525	CTCTACAATAGTGCTATGGCACGCAACTAACATGAAT-GAATACGATATACA	pCfB9007_2-genes insert_Gibson_fragment
17887	TCACTTCCCCATCCACACTTTTAGGTTCGAGACAGCACGT	pUDI264_insert Gibson_fragment
17888	AGGTTGATTCCGAACAGAAGTTAGCCAGAGGACACAGGAG	pUDI264_insert Gibson_fragment
17889	CTCCTGTGTCCTCTGGCTAACTTCTGTTCGGAATCAACCTC	pUDI264_backbone Gibson_fragment
17890	ACGTGCTGTCTCGAACCTAAAAGTGTGGATGGGGAAGTGA	pUDI264_backbone Gibson_fragment

Plasmid construction

Plasmids used in this study are shown in Table 3. Gene sequences coding for proteins involved in Moco biosynthesis (*OpCNX1*, *OpCNX2, OpCNX3*, *OpCNX4*, *OpCNX5*, *OpCNX6*, and *OpNFS1*) nitrate assimilation (*OpYNT1*, *OpYNR1*, *OpYNI1*) were retrieved from *O. parapolymorpha* DL-1 genome sequence [426, 478] BioProject PRJNA60503). A gene coding for a previously characterized high-affinity molybdenum transporter, *CrMoT1,* from *Chlamydomonas reinhardtii* was also included in the gene-set [425]. Each gene was codon-optimized for expression in *Y. lipolytica* using the GeneOptimizer tool (Thermo Fisher Scientific) and ordered as synthetic DNA from GeneArt (Thermo Fisher Scientific) resulting in plasmid pUD1057 (*OpCNX6*), pUD1058 (*OpCNX4*), pUD1059 (*OpCNX5*), pUD1060 (*OpNFS1*), pUD1061 (*OpCNX1*), pUD1062 (*OpCNX2*), pUD1063 (*CrMoT1*) pUD1064 (*OpCNX3*), pUD1065 (*OpYNR1*), pUD1066 (*OpYNI1*), and pUD1067 (*OpYNT1*). Single-gene Biobricks compatible with USER cloning (New England BioLabs) were amplified from pUD1057, pUD1058, pUD1059, pUD1060,

Table 3: Plasmids used in this study.

Name	Characteristics	Reference
pCfB6371	*bla* ColE1 ^{NotI}C_3-3'homology *tPEX20-tLIP2* C_3-5'homology^{NotI}	[482]
pCfB6677	*bla* ColE1 ^{NotI}E_1-3'homology *tPEX20-tLIP2* E_1-5'homology^{NotI}	[482]
pCfB6679	*bla* ColE1 ^{NotI}E_4-3'homology *tPEX20-tLIP2* E_4-5'homology^{NotI}	[482]
pCfB6681	*bla* ColE1 ^{NotI}E_3-3'homology *tPEX20-tLIP2* E_3-5'homology^{NotI}	[482]
pCfB6682	*bla* ColE1 ^{NotI}C_2-3'homology *tPEX20-tLIP2* C_2-5'homology^{NotI}	[482]
pCfB6684	*bla* ColE1 ^{NotI}D_1-3'homology *tPEX20-tLIP2* D_1-5'homology^{NotI}	[482]
pCfB6627	*bla* ColE1 *NAT* gRNA_C_2	[482]
pCfB6630	*bla* ColE1 *NAT* gRNA_C_3	[482]
pCfB6631	*bla* ColE1 *NAT* gRNA_D_1	[482]
pCfB6633	*bla* ColE1 *NAT* gRNA_E_1	[482]
pCfB6637	*bla* ColE1 *NAT* gRNA_E_3	[482]
pCfB6638	*bla* ColE1 *NAT* gRNA_E_4	[482]
pI774	*bla* ColE1 *Gfpmut3*	Unpublished
pUD1057	*bla* ColE1 *OpCNX6**	GeneArt
pUD1058	*bla* ColE1 *OpCNX4**	GeneArt
pUD1059	*bla* ColE1 *OpCNX5**	GeneArt
pUD1060	*bla* ColE1 *OpNFS1**	GeneArt
pUD1061	*bla* ColE1 *OpCNX1**	GeneArt
pUD1062	*bla* ColE1 *OpCNX2**	GeneArt
pUD1063	*bla* ColE1 *CrMoT1**	GeneArt
pUD1064	*bla* ColE1 OpCNX3*	GeneArt
pUD1065	*bla* ColE1 *OpYNR1**	GeneArt
pUD1066	*bla* ColE1 *OpYNI1**	GeneArt
pUD1067	*bla* ColE1 *OpYNT1**	GeneArt
pCfB8966	*bla* ColE1 ^{NotI}E_4-3'homology *tPEX20-OpCNX1*-pTEF1in pGPD-OpCNX2*-tLIP2* E_4-5'homology^{NotI}	This study
pCfB8967	*bla* ColE1 ^{NotI}C_2-3'homology *tPEX20-OpNFS1*-pTEF1in pGPD-OpCNX4*-tLIP2* C_2-5'homology^{NotI}	This study
pCfB8968	*bla* ColE1 ^{NotI}E_1-3'homology *tPEX20-OpCNX6*-pTEF1in pGPD-OpCNX5*-tLIP2* E_1-5'homology^{NotI}	This study
pCfB8969	*bla* ColE1 ^{NotI}E_3-3'homology *tPEX20-CrMoT1*-pTEF1in pGPD-OpCNX3*-tLIP2* E_3-5'homology^{NotI}	This study
pCfB8970	*bla* ColE1 ^{NotI}C_3-3'homology *tPEX20-OpYNR1*-pTEF1in pGPD-OpYNT1*-tLIP2* C_3-5'homology^{NotI}	This study
pCfB8971	*bla* ColE1 ^{NotI}D_1-3'homology *tPEX20-OpYNI1*-pTEF1in-tLIP2* D_1-5'homology^{NotI}	This study

pCfB9006	*bla* ColE1 ^{NotI}E_1-3'homology *tPEX20-OpCNX6*-pTEF1in pGPD-OpCNX5*-tLIP2 Gfpmut3 tPEX20-CrMoT1*-pTEF1in pGPD-OpCNX3*-tLIP2* E_1-5'homology^{NotI}	This study
pCfB9007	*bla* ColE1 ^{NotI}C_3-3'homology *tPEX20-OpYNR1*-pTEF1in pGPD-OpYNT1*-tLIP2 Gfpmut3 tPEX20-OpYNI1*-pTEF1in- tLIP2* C_3-5'homology^{NotI}	This study
pUDI264	*bla* ColE1 ^{NotI}E_3-3'homology *tPEX20-OpCNX6*-pTEF1in pGPD-OpCNX5*-tLIP2* E_3-5'homology^{NotI}	This study

*Codon optimized for expression in *Y. lipolytica*.

pUD1061, pUD1062, pUD1063, pUD1064, pUD1065, pUD1066, and pUD1067 using primer pair 24487/24488, 24485/24486, 24489/24490, 24483/24484, 24479/24480, 24481/24482, 24491/24492, 24493/24494, 24495/24496, 24499/24500, and 24497/24498, respectively. Single promoter Biobricks PrTEFin and PrGPD were amplified from ST6512 genomic DNA and primer pairs 22956/24013 and 15529/15528, respectively. Then, a Biobrick carrying the back-to-back promoter pair PrTEFin-PrGPD was cloned by USER cloning [481] fusion of the two single-promoter Biobricks and primer pair 22956/15528. A single promoter PrTEFin Biobrick was amplified using ST6512 genomic DNA and primer pair 27208/22956. Backbone Biobricks for the integration at the E_4, C_2, E_1, E_3, C_3, and D_1 integration sites were prepared by digestion and nicking of plasmids pCfB6679, pCfB6682, pCfB6677, pCfB6681, pCfB6371, and pCfB6684, respectively, using endonuclease FastDigest AsiSI (Life Technologies, Carlsbad, CA) and Nb.BsmI (New England BioLabs) followed by gel purification as previously described [482]. A backbone E_4 Biobrick was combined with biobricks PrTEFin-PrGPD, *OpCNX1*, and *OpCNX2* in a USER cloning reactions as previously described [482] to yield plasmid CfB8966. Similarly, plasmids pCfB8967, pCfB8968, pCfB8969, and pCfB8970 were cloned by combining a PrTEFin-PrGPD promoter biobrick with a C_2, E_1, E_3, or C_3 backbone, respectively, and gene Biobricks *OpNFS1*/*OPCNX4*, *OpCNX6*/*OpCNX5*, *CrMOT1*/*OpCNX3*, and *OpYNR1*/*OpYNT1*, respectively.

Plasmid pCfB8971 was cloned by combining a D_1 backbone Biobrick with a PrTEFin promoter and *OpYNI1* gene Biobricks. Gibson assembly [395] was used to construct the integration plasmid pCfB9006 that carries overexpression cassettes for *OpCNX3*, *CrMoT1*, *OpCNX5*, and *OpCNX6*. First, the plasmid pCfB8968 carrying the *OpCNX5-OpCNX6* cassette was linearized with primers 24522/24517 and the *OpCNX3-CrMoT1* cassette was amplified using primers 24518/24523 and pCfB8969 as template. Secondly, the GFPmut3 spacer cassette [483] was amplified from plasmid pI774 (Unpublished, doi: 10.4121/14230238) with primer pairs 24520/24521. Each fragment was gel-purified and combined in equimolar amounts in a Gibson reaction following manufacturer's instructions to form pCfB9006. In a similar way, the integration plasmid pCfB9007, carrying nitrate

assimilation pathway was cloned by combining a *OpYNT1-OpYNR1* cassette with an *OpYNI1* cassette and a GFPmut3 spacer. Plasmid pCfB8970 was linearized with primers 24524/24517 while the *OpYNI1* cassette was amplified with primers 24525/24518 and pCfB8971 as template. Plasmid pUDI264 for the integration of *OpCNX5-OpCNX6* at the E_3 integration site was cloned by Gibson assembly by combining a backbone fragment amplified with primers 17889/17890 and pCfB6681 as template and a fragment carrying the *OpCNX5-OpCNX6* cassette amplified with primers 17887/17888 and pCfB8968 as template. Correct plasmid assembly was verified by Sanger sequencing.

Strain construction

CRISPR/Cas9-mediated marker-free gene integration in *Y. lipolytica* was performed following the EASYcloneYali method as previously described [482]. In brief, 1 µg of NotI (Thermo Fisher Scientific) digested integrative plasmid carrying two or more genes flanked by long-homology arms to the integration *locus* was co-transformed together with 500 ng of a plasmid for the expression of a gRNA targeting the *locus* of interest. After selection and correct genotyping of transformants via diagnostic PCR, one colony of correct clone was inoculated in a 50 ml Greiner tube containing 20 ml YPD and incubated overnight at 30 °C, 200 rpm to allow the gRNA plasmid loss. The next day, cells were streaked to single colonies on a YPD plate and then, after incubation at 30 °C overnight, single colonies were patched on both selective and non-selective media. One plasmid-free colony was then picked and grown overnight in a YPD flask prior stocking at -80°C.
First, strain ST6512 was transformed with NotI digested pCfB8966 and the E_4 targeting gRNA expression plasmid pCfB6638 to yield IMX2264. Then, IMX2264 was transformed with NotI digested pCfB8967 and the C_2 targeting gRNA expression plasmid pCfB6628 to yield IMX2265. IMX2266 was obtained by transforming IMX2265 with NotI digested pCfB9006 and the E_1 targeting gRNA expression plasmid pCfB6633. IMX2264 was transformed with NotI digested pCfB9007 and the C_3 targeting gRNA expression plasmid pCfB6630 to yield IMX2267. Whole genome re-sequencing of IMX2267 revealed that *OpCNX5* and *OpCNX5* were not integrated but instead *CrMoT1* and *OpCNX3* were present twice (PRJNA704845). To correct this absence, IMX2267 was subsequently transformed with NotI digested pUDI264 together with the E_3 targeting gRNA expression plasmid pCfB6637, to yield the final strain IMX2565. The presence of all integrated genes in IMX2565 was confirmed by whole genome re-sequencing (PRJNA704845).

Adaptive laboratory evolution

To evolve IMX2565 for fast growth in nitrate-containing media, the strain was inoculated in triplicate in 100 ml flasks containing 20 ml SMDNO3. Flasks were incubated at 30 °C, 200 rpm until OD660 reached a value above 5. Then, 0.2 ml of each culture was

transferred in a new shake flask containing the same medium and incubated again. This process was repeated for 50 times, corresponding to approximately 335 generations, after which the evolved population were stocked and named IMS1183 (line 1), IMS1184 (line 2), and IMS1185 (line 3). Glycerol stocks for each evolution line were also prepared at intermediate steps after 3, 6, 9, 12, 21, 27, 38 and 50 transfers. Before transferring the cultures from batch 21, single colonies were isolated by restreaking each culture on SMDNO3 agar plates. After incubation of the plates at 30 °C in a static incubator for 48 h, three single colonies for each evolution line were picked and inoculated in 100 ml shake flasks containing 20 ml SMD_{NO3}. Flasks were incubated for 48 h at 30 °C, 200 rpm and the grown biomass was then stocked at -80 °C by the addition of 30% v/v glycerol and named IMS1174 (line 2, colony 1), IMS1175 (line 2, colony2), IMS1176 (line 2, colony 3), IMS1177 (line 3, colony 1), IMS1178 (line 3, colony 2), IMS1179 (line 3, colony 3), IMS1180 (line 1, colony 1), IMS1181 (line 1, colony 2), and IMS1182 (line 1, colony 3).

High-throughput strain cultivation and growth rate estimation

The growth rate of strains IMS1174, IMS1175, IMS1176, IMS1177, IMS1178, IMS1179, IMS1180, IMS1181, and IMS1182, together with the three evolved populations evolved populations IMX1183, IMS1184, and IMS1185 at batch number 3, 6, 9, 15, 21, 38, and 50, were estimated by cultivation in 96 deep-well plates in a Growth Profiler 960 instrument (Enzyscreen, Heemstede, The Netherlands). A glycerol stock for each strain was inoculated in a 100 ml shake flask containing 20 ml SMD_{NO3} and incubated overnight at 30 °C, 200 rpm. The next day, each culture was centrifuged at 3000 *g* for 5 minutes, supernatant was discarded, and cell pellets were resuspended in SMD_{NO3} to and OD_{660} of 5. Then, two 96 deep-well plates were filled with 250 µl of SMD_{NO3} medium and each well was inoculated with 5 µl of one of the tested strains. Each strain was inoculated in triplicate with the exception of IMS1174, IMS1176, IMS1178, IMS1179, IMS1181, and IMS1182 that were inoculated in duplicate. Plates were incubated at 30 °C, 250 rpm until the cultures reached stationary phase. To convert the measured "green" cell density values into OD_{660} equivalent, a calibration curve was prepared by correlating the "green" value of a IMS2565 cultures in SMD_{Urea} at eight different OD_{660} values that were measured externally with a 7200 Jenway Spectrophotometer (Jenway, Stone, United Kingdom). Moreover, values measured for each position in the plate were normalized by a factor that was calculated by measuring the green value of a IMS2565 culture in SMD_{Urea} of $OD_{660} = 5$ and by dividing that value by the average value measured across the whole plate (position normalization). After normalizing each green value time point by its own position normalization factor, OD_{660} equivalent values were calculated by fitting values with the calibration curve. Growth rate of each culture was calculated by fitting the exponential growth function with points of OD_{660} equivalent values between 0.5 and 2.

Aerobic cultivation in shake flasks

For the determination of the growth rate of IMS1174, IMS1175, IMS1176, IMS1177, IMS1178, IMS1179, IMS1180, IMS1181, and IMS1182, IMX1183, IMS1184, and IMS1185, frozen stock cultures were used to inoculate 20 mL SMD_{NO3} starter cultures. These were subsequently used to inoculate 100 mL SMD_{NO3} flask cultures to initial OD_{660} values between 0.1 and 0.2. Growth of these cultures was monitored with a 7200 Jenway Spectrophotometer (Jenway). Specific growth rates were calculated from at least five time points in the exponential growth phase of each culture. At each time point, 2 ml of liquid culture was centrifuged for 5 minutes at 14000 *g*, and supernatant was collected for HPLC and nitrate, nitrite, and ammonia analysis.

Whole-genome sequencing

Genomic DNA of strains ST6512, IMX2267, IMX2565, IMS1175, IMS1177, and IMS1180 was isolated with a Blood & Cell Culture DNA Kit with 100/G Genomics-tips (QIAGEN, Hilden, Germany) following manufacturer's instructions. Illumina-based paired-end sequencing with 150-bp reads was performed on 550-bp TruSeq DNA PCR-free insert libraries with a minimum resulting coverage of 50 x (Macrogen-Europe, Amsterdam, The Netherlands). Data mapping was performed using bwa 0.7.15-r1142-dirty against the *Y. lipolytica* CLIB122 genome [484, 485] to which 5 extra contigs containing the relevant integration cassettes had been previously added. Data processing and chromosome copy number variation determinations were done as previously described [318, 439] except for VCF file intersection and annotation that was performed with VCFtools_0.1.13 (vcf-isec command) and SnpEff, respectively [486, 487].

In vitro nitrate reductase activity measurements from cell extract

Nitrate reductase activity was measured from cell extract of strain ST6512, IMX2565, IMS1175 IMS1177, IMS1180, and IMX2565 evolution lines 1, 2, and 3. Cell extract preparation and activity measurements were performed as previously described [478]. In brief, frozen stock cultures of each tested strain and evolved populations were used to inoculate 20 mL starter cultures, which were then used to inoculate 100 mL shake flask cultures on the same medium, to an initial OD_{660} of 0.2. Shake flasks were incubated until the OD_{660} exceeded 40. All strains were grown on SMD_{urea} with the exception of the evolution lines 1, 2, and 3 that were grown on SMD_{NO3}. Cultures were then centrifuged at 3000 *g* for 5 min, supernatant was discarded, and cell pellets were resuspended in 2 ml lysis buffer (100 mM KPO_4 pH 7 supplemented with complete ULTRA EDTA-free protease inhibitor cocktail, Roche, Basel, Switzerland). Cell resuspensions were aliquoted in 1.5 ml bead-beating tubes along with 0.75 g of 400-600 µm acid-washed glass beads (Sigma Aldrich) per tube. Cells were disrupted by six 1-min cycles at 5 m/s speed in a

Fast-Prep 24 cell homogenizer (MP Biomedicals, Santa Ana, CA), with 5-min cooling on ice between cycles. Glass beads were separated from the cell extract by centrifuging the tubes for 10 min at 4 °C and 15000 *g* on a tabletop centrifuge. Supernatant was recovered in a fresh tube and clarified by centrifuging for 1 h at 4 °C, 20000 rpm. Clarified cell extracts were recovered and diluted 10 times in lysis buffer and kept on ice prior to analysis. Nitrate reductase activity was measured by monitoring substrate dependent NADPH consumption at 340 nm using a spectrophotometer (Jasco, Easton, MA). Reactions were performed in 1 ml final volume of 100 mM KPO_4 buffer pH 7 supplemented with 200 µM NADPH, 20 FAD, 1 mM KNO_3 as substrate, and either 50 or 25 µl of clarified cell extract. Protein contents of cell extracts were quantified with a Quick Start Bradford Assay (Bio-Rad Laboratories, Hercules, CA) following manufacturer's instructions. Specific activities of nitrate reductase in cell extracts were expressed in µmol $NADP^+$ min^{-1} (mg protein)$^{-1}$.

In silico mitochondrial targeting prediction

The likelihood of mitochondrial targeting for the wild-type and mutated *Op*Cnx1 and *Op*Cnx2 sequences were calculated using five different web-based tools: TargetP 2.0 [488], DeepLoc 1.0 [489], MitoFates [490], Predotar [491], and PredSL [492]. When possible, non-plant or fungal database was selected as option.

Analytical methods

Metabolite concentrations in culture supernatants were analysed by high-performance liquid chromatography (HPLC) on an Agilent 1260 HPLC (Agilent Technologies, Santa Clara, CA) fitted with a Bio-Rad HPX 87 H column (Bio-Rad). The flow rate was set at 0.6 mL min^{-1}, 0.5 g L^{-1} H_2SO_4 was used as eluent and the column temperature was set at 65 °C. An Agilent refractive-index detector and an Agilent 1260 VWD detector were used for metabolite quantification [309]. Nitrate, nitrite and ammonium concentrations culture supernatants were measured with a Hach DR3900 spectrophotometer and Hach kits LCK 339, LCK 341, and LCK 304 (Hach Lange, Düsseldorf, Germany), according to the manufacturer's instructions.

Statistical analysis

Statistical significance of differences between measurements from replicate samples were calculated by using a two-tailed t-test assuming unequal variances (Welch's correction).

Data availability

All measurement data and calculations used for each figure in the manuscript are available at the 4TU.Centre for research data repository (https://researchdata.4tu.nl/) under doi: 10.4121/14230238. DNA sequencing data of *Yarrowia lipolytica* strains ST6512, IMX2267, IMX2565, IMS1175, IMS11777, and IMS1180 were deposited at NCBI (https://www.ncbi.nlm.nih.gov/) under BioProject accession number PRJNA704845.

Results

Design and engineering of Moco biosynthesis and nitrate assimilation in *Yarrowia lipolytica*

The absence of molybdenum-dependent enzymes in *Yarrowia* metabolism strongly suggested that the engineering of Moco biosynthesis may not only require functional expression of Moco biosynthesis genes, but also of a high-affinity Molybdate transporter. (Figure 1). Based on a previous work in *S. cerevisiae* [478], it is not fewer than 11 genes from the yeast *O. parapolymorpha* and the algae *Chlamydomonas reinhardtii* that would be required to introduce Moco biosynthesis and nitrate assimilation in *Y. lipolytica*. The gene-set comprises seven genes coding for Moco biosynthesis proteins (*OpCNX1*, *OpCNX2*, *OpCNX3*, *OpCNX4*, *OpCNX5*, *OpCNX6*, and *OpNFS1*) and three genes coding for the nitrate assimilation pathway (*OpYNT1*, *OpYNR1*, and *OpYNI1*) from *O. parapolymorpha*, one gene coding for the high-affinity molybdate transporter (*CrMoT1*) from *C. reinhardtii* (Figure 1). The codon optimised genes were integrated in the chromosome of the *Yarrowia* strain ST6512 (W29, *ku70Δ::pTEF1-Cas9-tTEF12::pGPD-DsdA-tLIP2*) by using CRISPR/Cas9 gene-editing and the EASYcloneYali promoter parts and integrative plasmids [482]. Genes were sequentially integrated in five different integration sites (E_4, C_2, E_1, C_3, and E_3) that were previously tested for heterologous gene expression [482]. At each transformation, two genes were integrated, with the exception of the last transformation in which all three genes encoding for nitrate assimilation pathway were all integrated in one step. After five consecutive transformation rounds, the successful construction of the final strain IMX2565 was confirmed by Illumina short-read sequencing (PRJNA704845) (Figure 2).

Adaptive laboratory evolution on nitrate containing media

Similarly to what was observed in *S. cerevisiae*, adaptation was required to observe growth of the *Y. lipolytica* strain IMX2565 on nitrate containing medium [478]. We inoculated strain IMX2565 in SMD_{NO3} and incubated the flasks at 30 °C in triplicate. Reproducibly, after two weeks the *Yarrowia* strain showed full growth (OD_{660} above 20) in all flasks. To improve the strain growth rate, IMX2565 was subjected to adaptive laboratory evolution

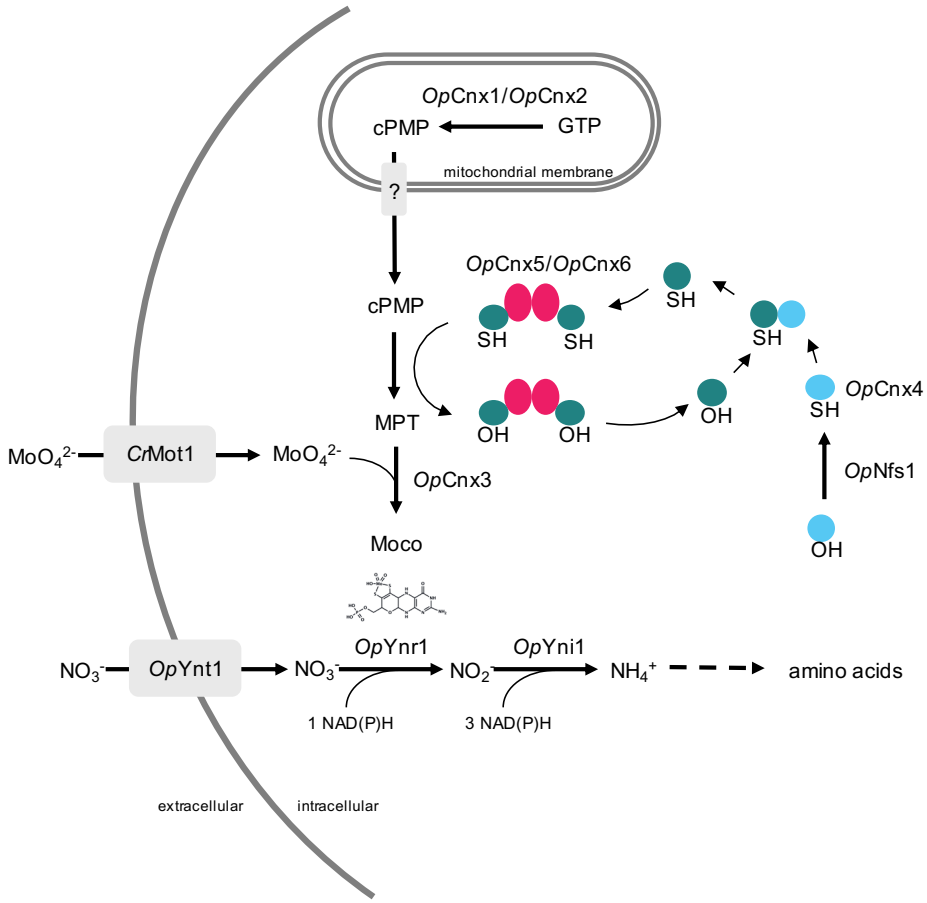


Figure 1: Schematic representation of the Moco biosynthesis pathway coupled to the nitrate assimilation pathway. GTP is converted to cyclic pyranopterin phosphate (cPMP) in the yeast mitochondria by *Op*Cnx1 and *Op*Cnx2. In the cytosol, cPMP is converted to molybdopterin (MPT) by *Op*Cnx5 and *Op*Cnx6. The sulfur moiety on *Op*Cnx5 is restored by *Op*Cnx4 that transfers the sulfur atom obtained by action of the cysteine desulfurase *Op*Nfs1. Molybdate (MoO_4^{2-}) is imported through the high affinity transporter *Cr*Mot1 and is inserted in MPT by *Op*Cnx3 to form Moco. Nitrate is imported via *Op*Ynt1 and reduced by the Moco-dependent nitrate reductase *Op*Ynr1 to nitrite. *Op*Yni1 converts nitrite to ammonia that finally enters the native nitrogen assimilation pathway. *Op*Cnx4, *Op*Cnx5, and *Op*Cnx6 are shown in light blue, teal and magenta, respectively. A question mark indicates a yet unknown cPMP transporter and a dashed line indicates multiple enzymatic steps.

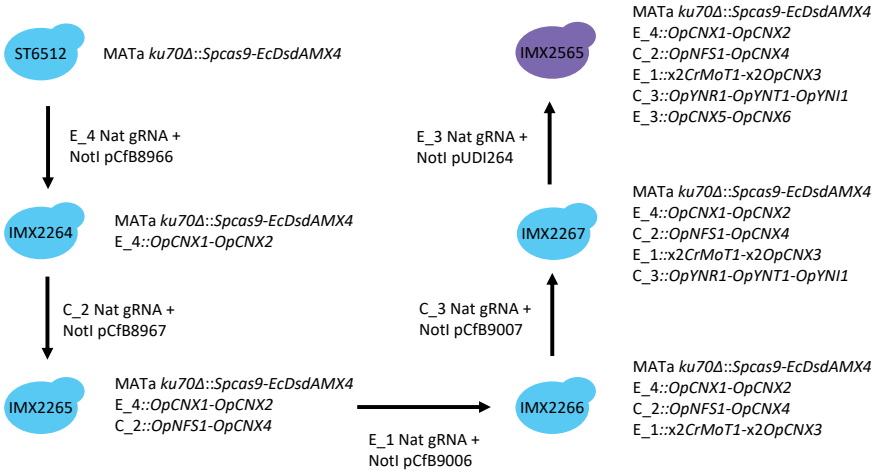


Figure 2: Schematic representation of strain construction. Strain ST6512, which constitutively expresses a Spcas9, was sequentially transformed 5 times with a gRNA expression plasmid and a NotI-linearized plasmid carrying the integration cassettes as repair fragment, yielding IMX2565. Intermediate strains and the final strain IMX2565 are shown in light-blue and purple, respectively.

(ALE) by sequential transfers in flasks containing SMD_{NO3} for 335 generations (50 consecutive batches). Evolving cultures were periodically stocked and used to monitor the growth rate of the evolving populations (Figure 3A). At the onset of the ALE experiment, the initial batches exhibited growth rates ranged from 0.04 to 0.05 h^{-1}. Throughout the first 21st transfers the specific growth rate of the yeast populations increased and levelled off to 0.11-0.12 h^{-1}. At that time-point, corresponding to a about 140 generations, three single colonies were isolated from each evolving culture and named IMS1180 (line 1, colony 1), IMS1181 (line 1, colony 2), and IMS1182 (line 1, colony 3), IMS1174 (line 2, colony 1), IMS1175 (line 2, colony2), IMS1176 (line 2, colony 3), IMS1177 (line 3, colony 1), IMS1178 (line 3, colony 2), IMS1179 (line 3, colony 3). The specific growth rates of the single colony isolates on SMD_{NO3} were ranged from 0.03 to 0.09 h^{-1} (Figure 3B).
The evolution experiment was prolonged for 29 additional sequential batches, summing up to a total of 50 batches (335 generations) to further probe evolvability of the phenotype. However, the growth rate of evolving populations stabilized to a value of about 0.13 h^{-1} and did not further increase (Figure 3A).

Whole-genome sequencing of evolved strains and mutations identification

To identify mutations responsible for the increase in growth rate, the genomes of clones derived from each evolution line (IMS1175, IMS1177, and IMS1180) and well as those of parental strains (ST6512 and IMX2565) were re-sequenced by Illumina short-

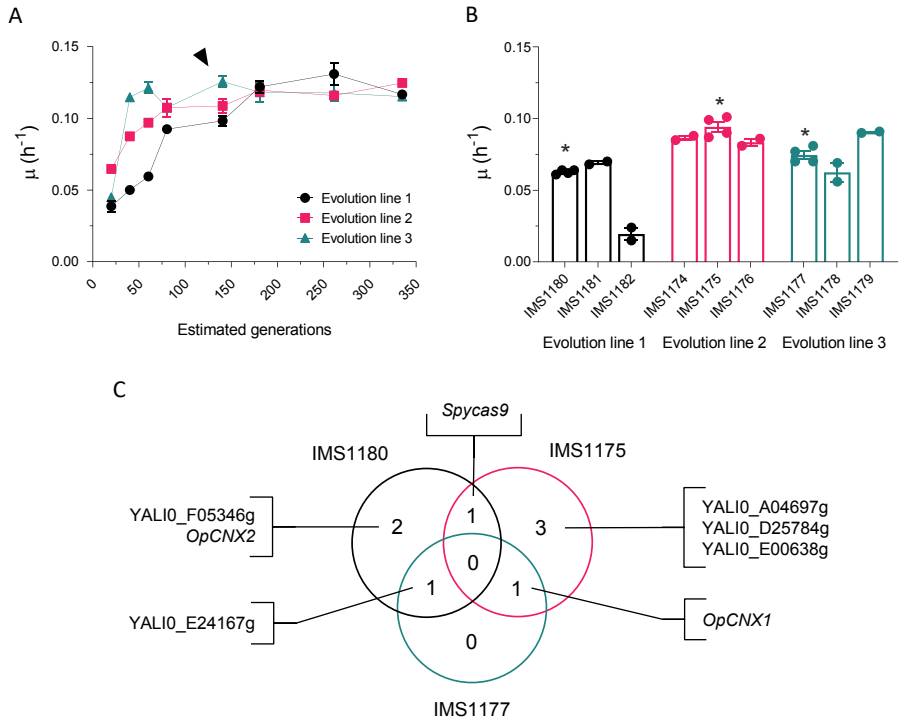


Figure 3: Overview of IMX2565 evolution on nitrate containing media. Growth rate values of IMX2565 throughout the adaptive laboratory evolution experiment (A), growth rates of single colony isolates (B) on SMD_{NO3} and Venn diagram highlighting genes affected by non-synonymous mutations and/or INDELs in independently evolved isolates IMS1175, IMS1177, and IMS1180 (C). An arrow indicates the time-point when single colonies were isolated from the three independent evolving populations while an asterisk indicates single colony isolates selected for whole-genome re-sequencing and further characterization. Evolution line 1, 2 and 3 are shown in black, magenta, and teal, respectively. Error bars represent the standard error of the mean for replicate cultures (n=4 with the exception of IMS1181, IMS1182, IMS1174, IMS1176, IMS1178, and IMS1179 where n=2). The Venn diagrams shows genes that acquired one or more non-synonymous mutations or INDELs in multiple independent evolution experiments as well as genes that were affected in a single replicate. Apparent mutations also found in the genome of the parent strain ST6512 and/or IMX2565 were subtracted and not shown.

read technology. After aligning the reads to the *Y. lipolytica* W29 reference genome (PRJNA601425) [484, 485], mapped data were analysed for the presence of either copy number variations (CNVs), single nucleotide variations (SNVs), and/or insertions/deletions (INDELs) that occurred in the annotated coding sequences. As opposed to what happened in *S. cerevisiae* [478], no gene or chromosome CNVs were observed between the parental strain IMX2565 and evolved isolates IMS1175, IMS1177, and IMS1180. SNV and INDELs analysis was systematically performed and data from the three sequenced colony isolates

were then compared. To minimize the number of miscalls caused by mapping artefacts, SNV and INDELs that were also detected in the two sequenced parental strains ST6512 and IMX2565, mapped to the same reference sequence, were systematically removed. After curation of SNVs and INDELs miscalls, a total of five, two, and four mutations were found in evolved isolates IMS1175, IMS1177, and IMS1180, respectively (Figure 3C, Table 4). While no genes were found mutated in all three independently evolved isolates, three genes (*Spycas9*, *OpCNX1* and YALI0_E24167g) were found differently mutated in two sequenced isolates. The mutations identified in *OpCNX1* that were found in IMS1175 and IMS1177, were identical and located in the 5' end of the coding sequence (G22A). Since the evolution lines were started independently and that the mutation was not present in the engineered strain IMX2565, the recurrence of this mutation in OpCNX1 might be critical in the acquisition of faster growth of *Y. lipolytica* on nitrate as N-source. The second gene (YALI0_E24167g) mutated in two different isolates IMS1180 and IMS1177 encoded a putative sulphite transporter that shared similarity to the *S. cerevisiae* sulfite efflux pump Ssu1 [493]. The nature of the mutations found in YALI0_E24167g would suggest a loss of function. In IMS1177 the mutation G743A resulted in the introduction of a stop at position 248. The mutated protein was truncated of 45% of the original sequence which should prevent functionality.

The last gene found mutated in two different isolates was *Spycas9*. Although different these two mutations occurred in a section of the gene encoding amino acids located in the same functional domain. The *Spycas9* mutation in IMS1175 (G2959A) and in IMS1180 (C3020A) led to amino acids change in the RuvC-III domain of *Spy*Cas9 [494]. Whereas IMS1175 and IMS1177 both harboured mutations in *OpCNX1*, interestingly IMS1180 had a mutation in *OpCNX2*. As observed for *OpCNX1*, the *OpCNX2* mutation was located in the 5' end the gene (G4A). On top of that, four more mutations were identified in only one isolates, it included YALI0_D25784g encoding a hypothetical protein, YALI0_A04697g a gene encoding a putative serine/threonine protein kinase, YALI0_E00638g a gene encoding a putative methyl citrate synthase in IMS1175 and YALI0_F05346g a gene encoding a protein exhibiting similarity with a cutinase from *Fusarium solani* cutinase.

Adaptive Laboratory Evolution for fast growth on nitrate was associated with an increased Moco-dependent nitrate reductase activity.

To investigate the effects of the adaptive laboratory evolution experiment, NADPH Moco-dependent nitrate reductase activity was assayed from cell-extracts of single colony isolates IMS1180, IMS1175, and IMS1177, evolved populations 1-2-3 after 50 transfers in SMD_{NO3}, the parental un-evolved engineered strain IMX2565 and the Cas9-expressing parental strain ST6512 (Figure 4). While as expected, cell extract from ST6512 which does not carry any of the heterologous genes for Moco biosynthesis and nitrate assimilation,

Table 4: SNVs and INDELs found in single colony isolates IMS1775, IMS1177, and IMS1180 obtained from the serial transfer evolution experiment of strain IMX2565 in SMD_{NO3}.

Mutated gene	Mutation type	Base change	Amino acid change	Gene annotation
IMS1175				
YALI0_A04697g	INDEL	T970TG	Gln324Frameshift	similar to uniprot|O42626 *Neurospora crassa* Serine/threonine-protein kinase nrc-2 (Nonrepressible conidiation protein 2)
YALI0_D25784g	SNV	T457C	Ser151Pro	weakly similar to uniprot|O74782 *Schizosaccharomyces pombe* Hypothetical protein
YALI0_E00638g	SNV	G868T	Gly290Cys	similar to uniprot|Q9TEM3 *Emericella nidulans* MCSA Methylcitrate synthase precursor
OpCNX1	SNV	G22A	Glu8Lys	GTP 3',8-cyclase
Spycas9	SNV	G2959A	Ala987Thr	CRISPR-associated endonuclease Cas9
IMS1177				
YALI0_E24167g	SNV	G743A	Trp248Stop	weakly similar to uniprot|Q2VQ77 *Saccharomyces cerevisiae* YPL092w *SSU1* Plasma membrane sulfite pump and required for efficient sulfite efflux
OpCNX1	SNV	G22A	Glu8Lys	GTP 3',8-cyclase
IMS1180				
YALI0_E24167g	SNV	A440G	His147Arg	weakly similar to uniprot|Q2VQ77 *Saccharomyces cerevisiae* YPL092w *SSU1* Plasma membrane sulfite pump and required for efficient sulfite efflux
YALI0_F05346g	SNV	G337A	Ala113Thr	weakly similar to uniprot|Q00858 *Fusarium solani* cutinase gene palindrome-binding protein
OpCNX2	SNV	G4A	Val2Met	cyclic pyranopterin monophosphate synthase
Spycas9	SNV	C3020A	Ala1007Asp	CRISPR-associated endonuclease Cas9

showed no significant nitrate reductase activity, the engineered strain IMX2565 showed an activity of 0.004 ± 0.001 μM $NADP^+$ min^{-1} mg total $protein^{-1}$. Although significant, this activity was up to 55-fold lower than the one observed for the evolved single colony isolates IMS1180, IMS1175, and IMS1177 that reached up to 0.22 ± 0.01 μM $NADP^+$ min^{-1} mg total $protein^{-1}$. Finally, cell extracts from the evolved populations showed activity values that were comparable for the 335 generations old, prolonged evolution line 1 and within 2-fold for the 335 generations old, prolonged evolution 2 and 3 with those measured for the single colony isolates, reaching up to 0.41 ± 0.03 μM $NADP^+$ min^{-1} mg total $protein^{-1}$.

To see whether the increase in nitrate reductase activity was associated with improved growth performance on nitrate, the strains ST6512, IMS11780, IMS1175, IMS1177, as well as the prolonged evolved populations at T50 1-2-3 were cultivated in chemically defined medium with nitrate as N-source (SMD_{NO3}^-). The strains were grown in aerobic shake flasks and OD_{660}, glucose, nitrate, nitrite, and ammonium levels were monitored over time (Figure 5). As expected, the parental strain ST6512 did not show any growth and/or glucose/nitrate consumption (Figure 5B).

All the other strains and evolved populations showed a growth rate ranging between 0.07 ± 0.01 and 0.16 ± 0.01 h^{-1}. Although slightly different in absolute values when compared to the growth rates measured in 96-wells format (Figure 3B), the ones measured in 100 ml shake flasks followed the same trend with IMS1180 and IMS1175 being the slowest and being the fastest growing isolates, respectively. Growth rate of the evolved populations, that were kept evolving for about 195 generations after single colonies were isolated, was significantly higher when compared to the respective single clones, with up to 2-folds increase (Figure 5A). Notably, although moderated both evolved isolates and further evolved populations excreted the intermediates nitrite and ammonia during the exponential phase of growth, suggesting that nitrate assimilation was not growth limiting (Figure 5CDEFGH). Similarly to what happened to engineered nitrate assimilating *S. cerevisiae* strains [478], ammonia and nitrite were also excreted toward the end of the batch fermentation, resulting from continuous conversion or possible cell lysis.

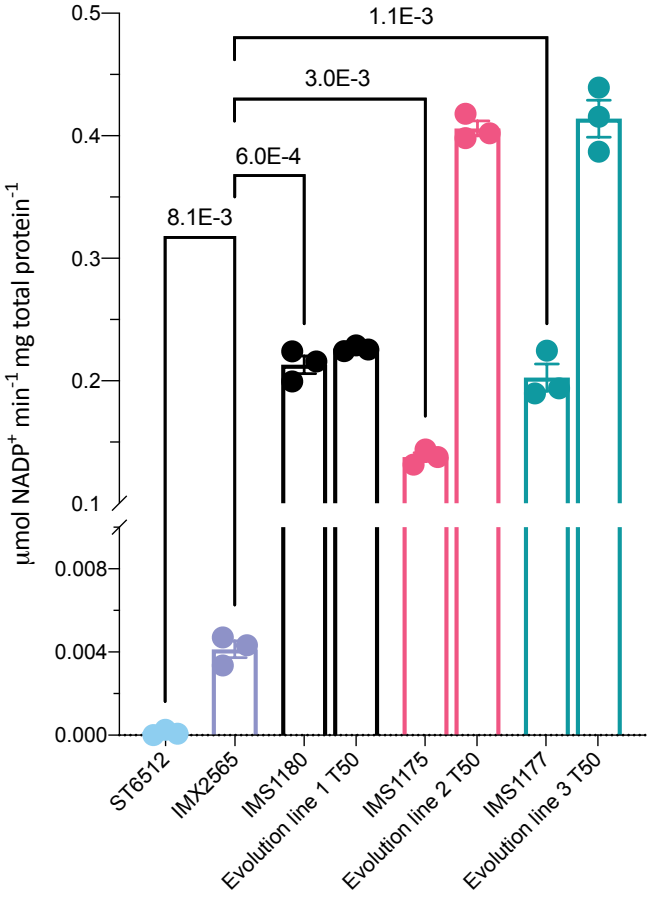
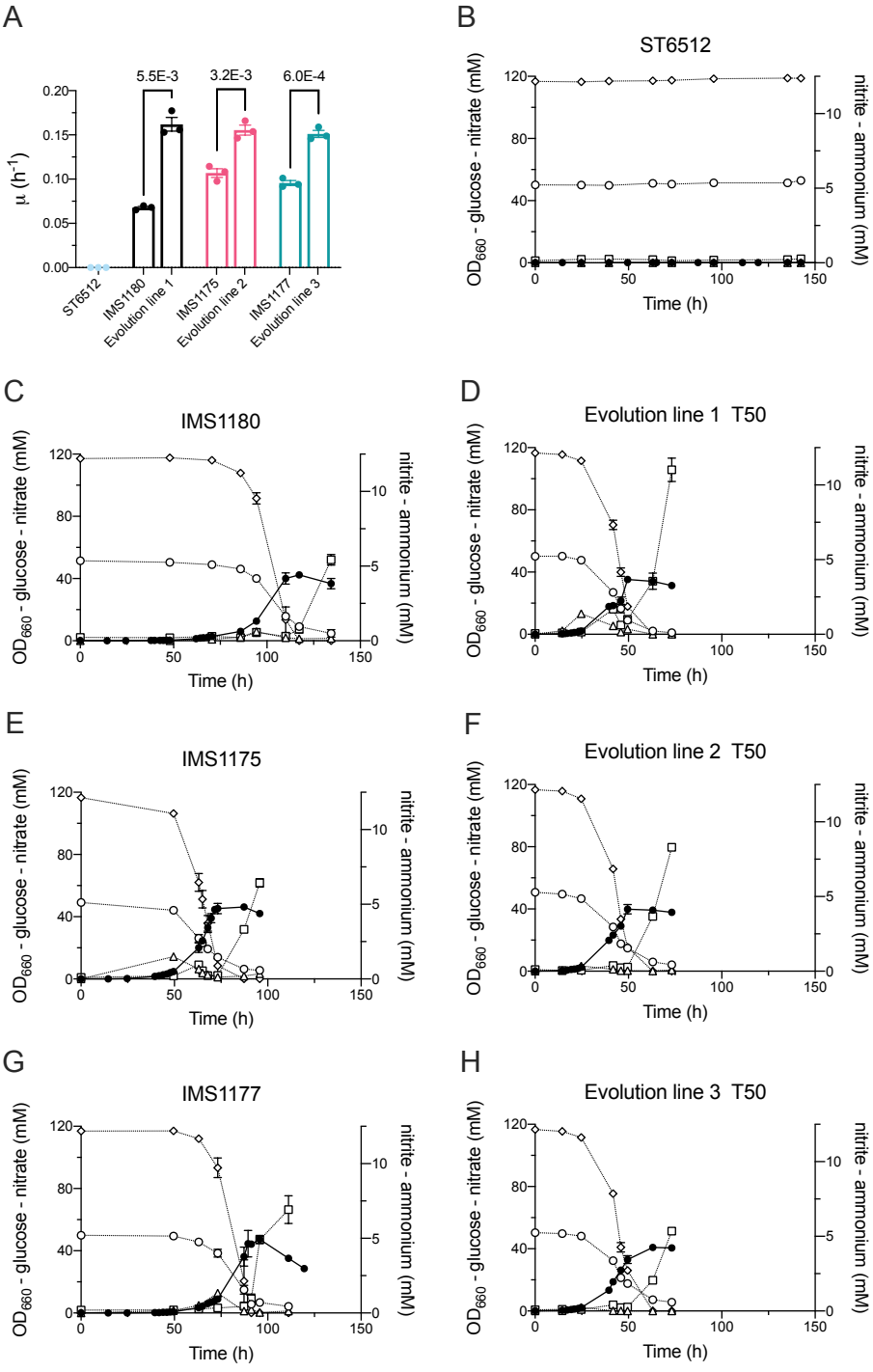


Figure 4: NADPH-dependent nitrate reductase activity measured from cell extracts. ST6512, IMX2565, evolution line 1 and isolate IMS1180, evolution line 2 and isolate IMS1175, and evolution line 3 and isolate IMS1177 are shown in light blue, purple, black, magenta, and teal, respectively. Cell extract was prepared from 100 ml of stationary phase cell cultures in SMD_{urea} for ST6512, IMX2565, IMS1180, IMS1175, and 1177 and in SMD_{NO3} for evolution lines T50 1, 2, and 3. P-values for two-tailed Welch's t-test are shown above the tested pairs. Error bars represent the standard error of the mean of technical replicates (n=3).

Figure 5: Growth rates (A) and growth curves of ST6512 (B), evolution line 1 T50 (C), IMS1180 (D), evolution line 2 T50 (E), IMS1175 (F), evolution line 3 T50 (G), and IMS1177 (H) on SMD_{NO3}. Symbols indicate biomass (black circle), glucose (white diamond), nitrate (white circle), nitrite (white square) and ammonium (white triangle). Statistical analysis was based on a two-tailed Welch's t-test and P-values are reported for tested pairs. Error bars represent the standard error of the mean for replicate cultures (n=3).


A
5.5E-3
3.2E-3
6.0E-4
μ (h-1)
0.20
0.15
0.10
0.05
0.00
ST6512
IMS1180
Evolution line 1
IMS1175
Evolution line 2
IMS1177
Evolution line 3
B
ST6512
OD660 - glucose - nitrate (mM)
nitrite - ammonium (mM)
Time (h)
C
IMS1180
D
Evolution line 1 T50
E
IMS1175
F
Evolution line 2 T50
G
IMS1177
H
Evolution line 3 T50

Discussion

Moco-dependent enzymes catalyse many redox reactions involved in the carbon, nitrogen and sulfur cycles [265] and might be harnessed to expand substrates range for microbial growth. This study firmly established that although Moco pathway is predominantly absent in Saccharomycotina yeasts, this function can be implemented by metabolic engineering. This study is, after introduction in *S. cerevisiae*, the second successful engineering of the pathway in yeasts. The large phylogenetic distance between the two tested hosts would also suggest that this pathway engineering could be extrapolated to an even larger set of budding yeasts [266]. While in both biological systems functionality through growth coupling with nitrate assimilation required an evolution step, the resulting mutations identified were different. In the precedent example in *S. cerevisiae*, adaptation to fast utilization of nitrate involved segmental aneuploidy of the whole heterologous cluster, accompanied with a strong increase of nitrate reductase activity. However, since the increase in copy number affected both genes of the Moco biosynthesis and the nitrate assimilation pathways it was not really possible to discriminate which metabolic branch was the most limiting. In *Yarrowia lipolytica* instead, the results would indicate that Moco synthesis is most limiting. This was supported by the absence of gene dosage but instead by recurrent mutations in the mitochondrial enzymes *Op*Cnx1 and*Op*Cnx2 that catalyse the formation of cPMP in the very first step of Moco biosynthesis (Figure 1). Evolved isolates IMS1175 and IMS1177 carried a mutation in the *Op*Cnx1 whereas the third isolate IMS1180 carried a mutation in *Op*Cnx2. But more specifically, the non-synonymous mutations led to amino acid change in N-terminus of both proteins changing the sequence of the mitochondrial targeting signal. Although theoretical, the presence of the mutation systematically improved the *in silico* prediction of mitochondrial targeting likelihood using five different algorithms [488-492]; Table 5). These mutations could contribute to improve OpCnx1 and 2 translocations into the organelle and optimise the supply of cPMP for Moco biosynthesis. The higher Moco availability would then be at the origin of the massive increase by 1 to 2 order of magnitude of nitrate reductase activity (Figure 4). These results seem to indicate that although the *Ogataea parapolymorpha* biosynthetic genes are sufficient to implement the synthesis of the new cofactor in other yeast species, the adaptation required to tune its supply optimally might be species dependent.

Among the three genes that were mutated in at least two isolates we found YALI0_E24167g (mutated in IMS1177 and IMS1180), a protein containing a predicted SLAC1 domain characteristic of voltage-dependent anion transporters such as nitrate and sulfite [495] and has similarity to the *S. cerevisiae* sulfite transporter Ssu1 which has been previously shown to be able to also export nitrate and nitrite [496]. While the IMS1180 mutation in YALI0_E24167g led to an His to Arg change, the mutation found in strain IMS1177 introduced a stop codon (Trp248Stop), creating a truncation of the translation product

Table 5: Likelihood of mitochondrial targeting in WT and mutated *Op*Cnx1-*Op*Cnx2 protein sequences across different prediction tools.

Prediction tool	*Op*Cnx1		*Op*Cnx2	
	WT	Mutated	WT	Mutated
TargetP 2.0	0.000208	0.000649	0.000525	0.001879
DeepLoc 1.0	0.0875	0.1946	0.1252	0.1261
MitoFates	0.080	0.280	0.000	0.000
Predotar	0.01	0.14	0.00	0.00
PredSL	0.006826	0.097417	0.004088	0.004091

by 180 amino acids, representing 42% of the protein length, that included not fewer than four transmembrane domains that undoubtedly caused a loss of function. Assuming that YALI0_E24167g shared function with its *S. cerevisiae* ortholog, the loss of function might represent a mechanism to avoid nitrogen loss through excessive nitrite export; it is worth noticing that this transport system would not be the only one as moderate extracellular nitrite concentration could still be measured. Maintenance of a higher intracellular nitrite could also partially explain the increase in nitrate reductase activity and growth rate after strain evolution in nitrate containing medium.

Interestingly, the third gene that was affected by non-synonymous mutations in at least two isolates (IMS1175 and IMS1180), was *Spycas9*. Although resulting in amino acid changes, it is not obvious to predict the impact of the mutations on the endonuclease activity. However, both mutations (Ala987Thr and Ala1007Asp) were found in the RuvC-III nuclease domain of Cas9 protein that extends from 925 to 1101. Several other mutations in RuvC-III at position 982, 983 and 986 have been described [497] and all yielded in either a decrease or an alteration of the endonuclease activity resulting in mutants able to only cleave one strand instead of two [494, 498]. This was confirmed experimentally, attempts further engineered the evolved isolates systematically failed. We cannot exclude that recurrent mutagenesis of Cas9 might take place to counteract potential endonuclease toxicity.

The Moco platform strain constructed and characterized in this study represents a steppingstone towards the exploitation of a new class of enzymes that might contribute to expand the metabolic capabilities of yeast microbial cell factories by enabling new metabolic engineering strategies.

Acknowledgments

We thank Marcel van den Broek for his support with the bioinformatic analysis and Dr. Jonathan Dahlin and Jonathan Asmund Arnesen for the technical support throughout the strain construction process. We also thank Prof. Jack T. Pronk for his support and contribution in the scientific discussion. This work was supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie action PAcMEN (grant agreement No 722287). T.P. and J.-M.G.D. are inventors on a patent application related to this work (WO2020209718 – Yeast with engineered molybdenum co-factor biosynthesis). The remaining authors have no competing interests to declare.

Author contributions

TP and JMD designed the experiments and wrote a first version of the manuscript. All authors critically read this version, provided input and approved the final version. TP performed the experiments and analysed the genome sequence data.

Outlook

Immense progress has been made in the last decade in the industrial biotechnology field, also thanks to the parallel development of key technologies.

Design and build

The advent of CRISPR/Cas9 gene-editing technology enabled precise, high-efficiency, and multiplexed DNA modifications, not only in model microorganisms, but also in less-conventional and less genetically amenable microbes [499, 500]. Chapter 4 of this thesis showcases how standardised CRISPR/Cas9-based protocols [313] contribute to improving the efficiency of pathway engineering in baker's yeast. The application of these tools allows for targeted and marker-free integration of entire pathways of up to 11 genes in one single transformation step. Chapters 4 and 5 illustrate the critical impact of these genome editing tools on functional analysis in *Ogataea parapolymorpha* and on engineering new metabolic pathways in *Yarrowia lipolytica* [423, 482]. Democratization of these editing tools will allow the future exploitation of microbial hosts with complex, multigene phenotypic characteristics that are typically hard to engineer (e.g. thermotolerance, low-pH resistance). While Cas9 technology already contributed to a significant reduction of the time required for strain construction, it may also help to accelerate strain optimisation process through application of deactivated Cas9 which, once fused to an activating or repressing domain, can be used to collectively deregulate sets of target genes [501-503]. Such technologies could be used in future work to design synthetic feedback regulation mechanisms to uncouple growth from production formation, to optimally control vitamin supply in vitamin prototrophic strains (chapters 2 and 3) to adapt synthesis to anabolic needs.

Costs of DNA reading and writing have exponentially decreased and will probably continue to decrease in the next decade [504]. Next generation genome sequencing technologies are becoming a routine in metabolic and evolutionary engineering studies. Mainly motivated by the cost, short read sequencing methods are almost standard, not only for detecting gene dosage and mutations in evolutionary engineering studies, but also for validating genotypes of heavily genetically engineered strains derived from multiple rounds of transformation and selection (chapters 2, 3, 4). Reuse of non-coding sequences such as promoters and terminators in such strains complicates genome data analysis. This problem is partly solved by the emergence of single-molecule sequencing technologies, which generate sequencing reads with lengths of up to dozens of kilobases. Among them, Oxford Nanopore technology and the associated pocket size MinION device will further

drive implementation of genome sequencing in metabolic engineering cycles [505-507]. Long-read sequencing technologies will contribute to accurately assemble complex and repetitive genomic regions and allow researchers to explore the genetic information until the very end of each chromosome [225, 508]. Nanopore sequencing is not restricted to long-read DNA sequencing but can also be applied to direct RNA sequencing, thus allowing RNA detection in field applications. The next challenge is the sequencing of peptides. Development of accurate techniques for this purpose could have a significant impact on metabolic engineering, as they will provide a direct access to protein abundance, a cellular measurement that today is only accessible through mass spectrometry. Cheap and easy access to protein concentration therefore has the potential to transform modelling and metabolic engineering approaches in the future [509-511].

Enzymatic DNA synthesis will probably be one of the hot topics in the coming years [512, 513]. Although the cost of DNA synthesis drastically decreased due to the advent of silicon-powered and semiconductor-based DNA manufacturing [514] , it remains expensive. The current costs of DNA synthesis prohibit direct synthesis of numerous coding sequences for screening purposes for academic research. Although the pathway for Moco biosynthesis and nitrate assimilation presented in chapters 4 and 5 comprised the same protein set, the coding genes were different. The genes used in *Y. lipolytica* were specifically codon optimised for this yeast, whose genome is characterised by higher GC content that that of *S. cerevisiae* (49% vs 39%, respectively). While codon optimization algorithms are designed with the objective to optimise the translation of the encoded mRNAs, different algorithms yield different nucleotide sequences without accurately predicting translation efficiency [515]. Further cost reduction of DNA synthesis could contribute to screen gene sequences derived from multiple optimisation algorithms. This will be even more critical with increased pathway complexity. For instance, further expansion of the yeast co-factor set by expression of a *de novo* B_{12} synthesis pathway, which would require expression of over 30 genes [516], would require an even higher investment in synthetic DNA, especially if multiple pathway configuration implicating multiple homologs for each catalytic step are to be considered. Enzymatic DNA synthesis is a technology that promises to change this scenario, allowing cheaper and rapid DNA synthesis, maybe even on a small benchtop instrument that could stand right next to the nanopore sequencer [517]. Such a technology could lower the DNA synthesis costs to a point where reverse engineering of mutant alleles, as performed in chapter 2, could be achieved by directly synthesizing the mutated gene, allowing for multiplexed reverse engineering of multiple mutant alleles in a single step.

Cost reductions in DNA reading and writing translates into more accessible high-throughput methodologies in the strain construction process. As result, process automation started to take place in different academic environments, with a number of

biofoundries opening their services to both academic and private laboratories [518], and with the increase popularity of open-source [519, 520] and alternative liquid handling solutions [521].

Test

While it is obvious that our capacity to construct strains faster and combinatorial cloning techniques allow for multiparallel strain construction, the throughput of strain testing has to improve proportionally. Traditional high-throughput screening strategies have relied on micro-titer-plate-based cultivation with a spectroscopic readout for product titer. However, this approach has limitations in terms of library size and in terms of the range of products that can be adequately quantified (e.g., because they contain conjugated double-bonds or are fluorescent). Advances on both fronts have been made in the last decade but challenges still need to be overcome. Cultivation in pico-liter sized droplets using microfluidic devices, enabled ultra-high throughput (~10^8 samples/day) screening [522, 523]. However, improvements observed in such miniaturized experiments are usually hard to directly translate to production scale and development of cost-models is key to maximize the correlation to at-scale strain performance [524-526]. Although a more expensive option, micro-fermentation systems can be a good bridge between ability to control cultivation parameters (e.g. substrate feed, pH etc.) and throughput [527]. Such high-throughput cultivation methods are key to enable faster strain optimization through combinatorial approach. Coupling a selection pressure to study complex characters or to improve strain performance can be a very powerful tool to unravel yet uncharacterized cellular mechanisms [268, 269]. However, adapted laboratory evolution experiments are typically time consuming and low throughput. Automated control of conditions during high-throughput cultivation by the use of setups like the eVOLVER, a cheap and scalable do-it-yourself framework which can be configured to carry out high-throughput growth experiments, can be a very powerful and efficient approach to study the strain adaptation across different conditions [528]. For instance, evolution for efficient nitrate assimilation by the engineered *Y. lipolytica* strain presented in chapter 5 could be performed using different carbon sources or by varying the concentration of nitrate and/or Mo in the media, while monitoring how these different contexts affect evolutionary adaptation.

Product detection techniques have also been greatly improved over the last years. One prominent example concerns the recently developed modular bio-sensors for coupling product titer to a fluorescent readout, enabling ultra-high-throughput screening pipelines for a wide range of products [529]. Moreover, mass spectrometry-based analytical assays, which allow highly-applicable, quantitative, and selective molecule detection, have gone high-throughput thanks to the development of platforms such as the RapidFire (~350 samples/day) [530] and the acoustic MIST ionization mass spectrometer (~10^6 samples/

day) [531], which enable fast detection while keeping a good linear response and signal-to-noise ratio.

Learn

The learning process represents the final step of the microbial strain engineering cycle. Machine learning approaches are taking a prominent role in different areas of strain and enzyme engineering, together with genome scale and kinetic metabolic models, they enable data-driven decision making in the strain metabolic engineering cycle [532]. Artificial intelligence and machine learning techniques have already been applied to solve complex problems and to predict biological behaviour such as protein folding and dynamics, and the properties and strength of promoter sequences [533-535]. The development of complex predictive models may in the future be applied to predict the effect of mutations found in evolutionary engineering studies and thereby help in prioritizing which of these mutations should be reverse engineered first by generating ranking tables of potentially causal mutations. High-throughput screening methods with high signal to noise ratio and accuracy coupled to data standardization are of crucial importance to successfully train such algorithms.

Future applications for yeasts with engineered vitamin and cofactor synthesis pathways

While efforts in strain engineering in industry mainly focus on the optimization stages and scale-up of established production hosts, academic laboratories in the industrial microbiology field have more freedom to explore and test new proof of concepts and provide innovative solutions. In the presented thesis, the work was focused at studying different approaches that ultimately yielded improved chassis yeast strains able to grow in the absence of vitamins and able to functionally express a new enzyme family that requires a non-native cofactor. Minimization costs of bioprocesses while achieving a minimal environmental impact is key for a transition to a viable bioeconomy. Simplification of industrial media by eliminating the need for addition of vitamin mixtures, thanks to the availability of a vitamin prototrophic production strain, can contribute to reducing the process costs. Additionally, it can result in increased medium shelf life and reproducibility, for example, in processes where feedstock preparation requires heating and/or acid-treatment steps [270, 536]. Another fraction of process cost is typically allocated to sterilization of the fermentation equipment. In fact, keeping the processes aseptic, especially at industrial scale, is a huge challenge and open and non-aseptic operation of fermentation processes should, when possible, be preferred in order to make biobased alternatives more economically attractive [537]. However, the successful implementation of such processes requires the use of microbial strains and cultivation conditions that are

more robust and that minimize the likelihood of contaminations within the process time frames. Open bioethanol fermentation processes are often found to be contaminated by the spoilage yeast *Brettanomyces bruxellensis*. These contaminations have been proposed to be partially caused by this yeast's ability to utilize nitrate as nitrogen source [421]. Equipping *S. cerevisiae* with a molybdenum-cofactor-dependent nitrate assimilation pathway as described in chapter 4, resulted in a strain with increased competitiveness against such spoilage yeast, showcasing how engineering of new cofactor biosynthesis can also result in strains that can improve the overall process robustness. Moreover, enabling the use of an engineered yeast that is able to grow in vitamin-free media can also minimize the chance of contaminations by, for instance, vitamin auxotrophic bacteria. On top of that, a vitamin producing strain could also be applied in the future as a food supplement, providing a cheap and low impact source of vitamins for human consumption. For such application, engineering of a heterologous vitamin biosynthetic pathway (*e.g.,* vitamin B_{12}) in yeast holds great potential to further expand the yeast vitamin-set. In the presented work in chapter 2 and 3, we provide solutions for removal of many of the vitamins that typically present in the chemically defined media for yeast, but more work may be needed to combine these solutions in one single strain, therefore enabling fast growth of yeasts in a vitamin-free mineral media.

Another advantage of engineering heterologous cofactor biosynthetic pathway in yeast is connected to the expansion of the host metabolic capabilities, therefore enabling the functional expression of new enzyme families. Such a technological advantage is exemplified by studies on *de novo* biosynthesis of opioids in yeast, which required the parallel engineering of a tetrahydrobiopterin cofactor biosynthetic pathway [258, 404]. In chapter 4 and 5, we describe the expansion of the enzymatic repertoire of two industrially relevant yeast species by enabling the functional expression of molybdenum-cofactor-dependent enzymes. The presented results can be considered an essential step toward enabling expressing of the full set of over 35 Mo-cofactor-dependent enzymes described to date [265]. In addition to the use of nitrate as a nitrogen source, such strategies could, for example, enable the expression of high-k_{cat} molybdoprotein formate dehydrogenase [457] as alternative for the low-k_{cat} native fungal formate dehydrogenases, and of the molybdoprotein furoyl-CoA dehydrogenase, which can contribute to conversion of toxic furanic compounds found in lignocellulosic hydrolysates [458].

Overall, the presented work contributes to our understanding of the yeast cofactor metabolism, provides new insights and a solid basis for further research in this as yet underexplored field of application-inspired yeast research.

Bibliography

1. Gowland Hopkins F. The analyst and the medical man. Analyst. 1906;31:385b-404.
2. Funk C. The vitamines. 2nd ed. E. DH, editor. Baltimore: Williams & Wilkins Company; 1922.
3. Combs GF. The vitamins : fundamental aspects in nutrition and health. 3 ed: Elsevier Academic Press; 2008. 583 p.
4. Potvin J, Fonchy E, Conway J, Champagne CP. An automatic turbidimetric method to screen yeast extracts as fermentation nutrient ingredients. J Microbiol Meth. 1997;29:153-60.
5. Zhang J, Reddy J, Buckland B, Greasham R. Toward consistent and productive complex media for industrial fermentations: studies on yeast extract for a recombinant yeast fermentation process. Biotechnology and Bioengineering. 2003;82:640-52.
6. Dahod SK, Greasham R, Kennedy M. Raw materials selection and medium development for industrial fermentation processes. Manual of Industrial Microbiology and Biotechnology, Third Edition. 2010:659-68.
7. Burkholder PR, McVeigh I, Moyer D. Studies on some growth factors of yeasts. J Bacteriol. 1944;48:385-91.
8. Cutts NS, Rainbow C. Studies of a yeast exacting towards *para*-aminobenzoic acid. Journal of General Microbiology. 1950;4:150-5.
9. Wickerham LJ. Taxonomy of yeasts. Technical Bulletin US Dept Agriculture, Washington DC. 1951;1029.
10. Sherman F, Fink GR, Hicks J. Methods in yeast genetics: A laboratory course manual. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 1987.
11. Wickerham LJ. A critical evaluation of the nitrogen assimilation tests as commonly used in the classification of the yeasts. J Bacteriol. 1946;51:6-.
12. Bruinenberg PM, van Dijken JP, Scheffers WA. An enzymic analysis of NADPH production and consumption in *Candida utilis*. Journal of General Microbiology. 1983;129:965-71.
13. Verduyn C, Postma E, Scheffers WA, Van Dijken JP. 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.
14. Guthrie C, Fink GR. Getting started with yeast. Methods in Enzymology. 1991;194:3-21.
15. Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, et al. Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast. 1998;14:115-32.
16. Hanscho M, Ruckerbauer DE, Chauhan N, Hofbauer HF, Krahulec S, Nidetzky B, et al. Nutritional requirements of the BY series of *Saccharomyces cerevisiae* strains for optimum growth. Fems Yeast Res. 2012;12:796-808.
17. Emery WB, McLeod N, Robinson FA. Comparative microbiological assays of members of the vitamin B complex in yeast and liver extracts. Biochemical Journal. 1946;40:426-32.
18. Pirner HM, Stolz J. Biotin sensing in *Saccharomyces cerevisiae* is mediated by a conserved DNA element and requires the activity of biotin-protein ligase. J Biol Chem. 2006;281:12381-9.
19. de Kock SH, du Preez JC, Kilian SG. The effect of vitamins and amino acids on glucose uptake in aerobic chemostat cultures of three *Saccharomyces cerevisiae* strains. Systematic and Applied Microbiology. 2000;23:41-6.
20. Rogosa M. Vitamin requirements of lactose-fermenting and certain other yeasts. J Bacteriol. 1944;47:159-70.
21. Burkholder PR. Vitamin deficiencies in yeasts. American Journal of Botany. 1943;30:206-11.
22. Andreasen AA, Stier TJ. Anaerobic nutrition of *Saccharomyces cerevisiae.* II. Unsaturated fatty acid requirement for growth in a defined medium. Journal of Cellular and Comparative Physiology. 1954;43:271-81.
23. Andreasen AA, Stier TJ. Anaerobic nutrition of *Saccharomyces cerevisiae*. I. Ergosterol requirement for growth in a defined medium. J Cell Comp Physiol. 1953;41:23-36.
24. Snoek IS, Steensma HY. Why does *Kluyveromyces lactis* not grow under anaerobic conditions? Comparison of essential anaerobic genes of *Saccharomyces cerevisiae* with the *Kluyveromyces lactis* genome. Fems Yeast Res. 2006;6:393-403.
25. Rosenfeld E, Beauvoit B. Role of the non-respiratory pathways in the utilization of molecular oxygen by *Saccharomyces cerevisiae*. Yeast. 2003;20:1115-44.

26. Pronk JT. Auxotrophic yeast strains in fundamental and applied research. Appl Environ Microb. 2002;68:2095-100.
27. Keresztesy JC, Stevens JR. Vitamin B6. Journal of the American Chemical Society. 1938;60:1267-8.
28. Lepkovsky S. Crystalline Factor I. Science. 1938;87:169-70.
29. Harris SA, Folkers K. Synthesis of vitamin B6. Journal of the American Chemical Society. 1939;61:1245-7.
30. Schultz AS, Atkin L, Frey CN. Vitamin B6 a growth promoting factor for yeast. Journal of the American Chemical Society. 1939;61:1931-.
31. Andersen G, Andersen B, Dobritzsch D, Schnackerz KD, Piskur J. A gene duplication led to specialized g-aminobutyrate and b-alanine aminotransferase in yeast. Febs J. 2007;274:1804-17.
32. Volland C, Felix F. Isolation and properties of 5-aminolevulinate synthase from the yeast *Saccharomyces cerevisiae*. Eur J Biochem. 1984;142:551-7.
33. Hilger F, Mortimer RK. Genetic mapping of *ARG1* and *ARG8* in *Saccharomyces cerevisiae* by trisomic analysis combined with interallelic complementation. J Bacteriol. 1980;141:270-4.
34. Phalip V, Kuhn I, Lemoine Y, Jeltsch JM. 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.
35. Takada Y, Noguchi T. Characteristics of alanine: glyoxylate aminotransferase from *Saccharomyces cerevisiae*, a regulatory enzyme in the glyoxylate pathway of glycine and serine biosynthesis from tricarboxylic acid-cycle intermediates. Biochemical Journal. 1985;231:157-63.
36. Dai YN, Chi CB, Zhou K, Cheng W, Jiang YL, Ren YM, et al. Structure and catalytic mechanism of yeast 4-amino-4-deoxychorismate lyase. J Biol Chem. 2013;288:22985-92.
37. Zhang H, Yang Z, Shen Y, Tong L. Crystal structure of the carboxyltransferase domain of acetyl-coenzyme A carboxylase. Science. 2003;299:2064-7.
38. Karsten WE, Reyes ZL, Bobyk KD, Cook PF, Chooback L. Mechanism of the aromatic aminotransferase encoded by the *ARO8* gene from *Saccharomyces cerevisiae*. Archives in Biochemistry and Biophysics. 2011;516:67-74.
39. Yagi T, Kagamiyama H. Aspartate: 2-oxoglutarate aminotransferase from bakers' yeast: crystallization and characterization. Journal of Biochemistry. 1982;92:35-43.
40. Morin PJ, Subramanian GS, Gilmore TD. *AAT1*, a gene encoding a mitochondrial aspartate aminotransferase in *Saccharomyces cerevisiae*. Biochimica and Biophysica Acta. 1992;1171:211-4.
41. Eden A, Simchen G, Benvenisty N. Two yeast homologs of *ECA39*, a target for c-Myc regulation, code for cytosolic and mitochondrial branched-chain amino acid aminotransferases. J Biol Chem. 1996;271:20242-5.
42. Bornaes C, Petersen JG, Holmberg S. Serine and threonine catabolism in *Saccharomyces cerevisiae*: the *CHA1* polypeptide is homologous with other serine and threonine dehydratases. Genetics. 1992;131:531-9.
43. Holt S, Cordente AG, Williams SJ, Capone DL, Jitjaroen W, Menz IR, et al. Engineering *Saccharomyces cerevisiae* to release 3-Mercaptohexan-1-ol during fermentation through overexpression of an *S. cerevisiae* gene, *STR3*, for improvement of wine aroma. Appl Environ Microb. 2011;77:3626-32.
44. Jhee KH, McPhie P, Miles EW. Domain architecture of the heme-independent yeast cystathionine b-synthase provides insights into mechanisms of catalysis and regulation. Biochemistry. 2000;39:10548-56.
45. Messerschmidt A, Worbs M, Steegborn C, Wahl MC, Huber R, Laber B, et al. Determinants of enzymatic specificity in the Cys-Met-metabolism PLP-dependent enzymes family: crystal structure of cystathionine g-lyase from yeast and intrafamiliar structure comparison. Biological Chemistry. 2003;384:373-86.
46. Hansen J, Johannesen PF. Cysteine is essential for transcriptional regulation of the sulfur assimilation genes in *Saccharomyces cerevisiae*. Mol Gen Genet. 2000;263:535-42.
47. Pandey A, Golla R, Yoon H, Dancis A, Pain D. Persulfide formation on mitochondrial cysteine desulfurase: enzyme activation by a eukaryote-specific interacting protein and Fe-S cluster synthesis. Biochemical Journal. 2012;448:171-87.
48. Ito T, Hemmi H, Kataoka K, Mukai Y, Yoshimura T. A novel zinc-dependent D-serine dehydratase from *Saccharomyces cerevisiae*. Biochemical Journal. 2008;409:399-406.
49. Coleman ST, Fang TK, Rovinsky SA, Turano FJ, Moye-Rowley WS. Expression of a glutamate decarboxylase homologue is required for normal oxidative stress tolerance in *Saccharomyces cerevisiae*. J Biol Chem. 2001;276:244-50.
50. Sinclair DA, Hong SP, Dawes IW. Specific induction by glycine of the gene for the P-subunit of glycine decarboxylase from *Saccharomyces cerevisiae*. Mol Microbiol. 1996;19:611-23.

51. Becker JU, Wingender-Drissen R, Schiltz E. Purification and properties of phosphorylase from baker's yeast. Archives in Biochemistry and Biophysics. 1983;225:667-78.
52. Alifano P, Fani R, Lio P, Lazcano A, Bazzicalupo M, Carlomagno MS, et al. Histidine biosynthetic pathway and genes: structure, regulation, and evolution. Microbiology Reviews. 1996;60:44-69.
53. Yamagata S, Takeshima K. O-acetylserine and O-acetylhomoserine sulfhydrylase of yeast - further purification and characterization as a pyridoxal enzyme. Journal of Biochemistry. 1976;80:777-85.
54. Panozzo C, Nawara M, Suski C, Kucharczyka R, Skoneczny M, Becam AM, et al. Aerobic and anaerobic NAD^+ metabolism in *Saccharomyces cerevisiae*. FEBS Letters. 2002;517:97-102.
55. Wada M, Nakamori S, Takagi H. Serine racemase homologue of *Saccharomyces cerevisiae* has L-threo-3-hydroxyaspartate dehydratase activity. Fems Microbiol Lett. 2003;225:189-93.
56. Liu JQ, Nagata S, Dairi T, Misono H, Shimizu S, Yamada H. The *GLY1* gene of *Saccharomyces cerevisiae* encodes a low-specific L-threonine aldolase that catalyzes cleavage of L-allo-threonine and L-threonine to glycine - expression of the gene in *Escherichia coli* and purification and characterization of the enzyme. Eur J Biochem. 1997;245:289-93.
57. Degols G, Jauniaux JC, Wiame JM. Molecular characterization of transposable-element-associated mutations that lead to constitutive L-ornithine aminotransferase expression in *Saccharomyces cerevisiae*. Eur J Biochem. 1987;165:289-96.
58. Tyagi AK, Tabor CW, Tabor H. Ornithine decarboxylase from *Saccharomyces cerevisiae*. Purification, properties, and regulation of activity. J Biol Chem. 1981;256:12156-63.
59. Melcher K, Rose M, Kunzler M, Braus GH, Entian KD. Molecular analysis of the yeast *SER1* gene encoding 3-phosphoserine aminotransferase: regulation by general control and serine repression. Current Genetics. 1995;27:501-8.
60. Garcia-Campusano F, Anaya VH, Robledo-Arratia L, Quezada H, Hernandez H, Riego L, et al. *ALT1*-encoded alanine aminotransferase plays a central role in the metabolism of alanine in *Saccharomyces cerevisiae*. Canadian Journal of Microbiology. 2009;55:368-74.
61. Duff SM, Rydel TJ, McClerren AL, Zhang W, Li JY, Sturman EJ, et al. The enzymology of alanine aminotransferase (AlaAT) isoforms from *Hordeum vulgare* and other organisms, and the HvAlaAT crystal structure. Arch Biochem Biophys. 2012;528:90-101.
62. Wogulis M, Chew ER, Donohoue PD, Wilson DK. Identification of formyl kynurenine formamidase and kynurenine aminotransferase from *Saccharomyces cerevisiae* using crystallographic, bioinformatic and biochemical evidence. Biochemistry. 2008;47:1608-21.
63. Zhang N, Merlotti C, Wu J, Ismail T, El-Moghazy AN, Khan SA, et al. Functional Analysis of six novel ORFs on the left arm of Chromosome XII of *Saccharomyces cerevisiae* reveals three of them responding to S-starvation. Yeast. 2001;18:325-34.
64. Hughes AL, Hughes CE, Henderson KA, Yazvenko N, Gottschling DE. Selective sorting and destruction of mitochondrial membrane proteins in aged yeast. Elife. 2016;5.
65. Nakamura KD, Trewyn RW, Parks LW. Purification and characterization of serine transhydroxy-methylase from *Saccharomyces cerevisiae*. Biochimica and Biophysica Acta. 1973;327:328-35.
66. Perry DK. Serine palmitoyltransferase: role in apoptotic *de novo* ceramide synthesis and other stress responses. Biochimica and Biophysica Acta. 2002;1585:146-52.
67. Saba JD, Nara F, Bielawska A, Garrett S, Hannun YA. The *BST1* gene of *Saccharomyces cerevisiae* is the sphingosine-1-phosphate lyase. J Biol Chem. 1997;272:26087-90.
68. Karassevitch Y, Robichon-Szulmajster H. Reversible dissociation of threonine deaminase in an *ilvl* mutant of *Saccharomyces cerevisiae*. Mol Gen Genet. 1972;117:113-23.
69. Parsot C. Evolution of biosynthetic pathways: a common ancestor for threonine synthase, threonine dehydratase and D-serine dehydratase. Embo J. 1986;5:3013-9.
70. Bartholmes P, Boker H, Jaenicke R. Purification of tryptophan synthase from *Saccharomyces cerevisiae* and partial activity of its nicked subunits. Eur J Biochem. 1979;102:167-72.
71. Arlt H, Perz A, Ungermann C. An overexpression screen in *Saccharomyces cerevisiae* identifies novel genes that affect endocytic protein trafficking. Traffic. 2011;12:1592-603.
72. Coquille S, Roux C, Fitzpatrick TB, Thore S. The last piece in the vitamin B1 biosynthesis puzzle: structural and functional insight into yeast 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate (HMP-P) synthase. J Biol Chem.

2012;287:42333-43.

73. Repetto B, Tzagoloff A. Structure and regulation of *KGD1*, the structural gene for yeast a-ketoglutarate dehydrogenase. Molecular and Cellular Biology. 1989;9:2695-705.
74. Falco SC, Dumas KS, Livak KJ. Nucleotide sequence of the yeast *ILV2* gene which encodes acetolactate synthase. Nucleic Acids Research. 1985;13:4011-27.
75. Notzel C, Lingner T, Klingenberg H, Thoms S. Identification of new fungal peroxisomal matrix proteins and revision of the *PTS1* consensus. Traffic. 2016;17:1110-24.
76. Schmitt HD, Ciriacy M, Zimmermann FK. The synthesis of yeast pyruvate decarboxylase is regulated by large variations in the messenger RNA level. Mol Gen Genet. 1983;192:247-52.
77. Hohmann S, Cederberg H. Autoregulation may control the expression of yeast pyruvate decarboxylase structural genes *PDC1* and *PDC5*. Eur J Biochem. 1990;188:615-21.
78. Hohmann S. Characterization of *PDC6*, a third structural gene for pyruvate decarboxylase in *Saccharomyces cerevisiae*. J Bacteriol. 1991;173:7963-9.
79. Zeeman AM, Luttik MA, Thiele C, van Dijken JP, Pronk JT, Steensma HY. Inactivation of the *Kluyveromyces lactis KlPDA1* gene leads to loss of pyruvate dehydrogenase activity, impairs growth on glucose and triggers aerobic alcoholic fermentation. Microbiology. 1998;144:3437-46.
80. Miran SG, Lawson JE, Reed LJ. Characterization of *PDH* b1, the structural gene for the pyruvate dehydrogenase b-subunit from *Saccharomyces cerevisiae*. Proceedings of the national Academy of Sciences, USA. 1993;90:1252-6.
81. Nishimura H, Kawasaki Y, Kaneko Y, Nosaka K, Iwashima A. A positive regulatory gene, *THI3*, is required for thiamine metabolism in *Saccharomyces cerevisiae*. J Bacteriol. 1992;174:4701-6.
82. Vuralhan Z, Morais MA, Tai SL, Piper MD, Pronk JT. Identification and characterization of phenylpyruvate decarboxylase genes in *Saccharomyces cerevisiae*. Appl Environ Microb. 2003;69:4534-41.
83. Kochetov G, Sevostyanova IA. Binding of the coenzyme and formation of the transketolase active center. IUBMB Life. 2005;57:491-7.
84. Schaaff-Gerstenschlager I, Mannhaupt G, Vetter I, Zimmermann FK, Feldmann H. *TKL2*, a second transketolase gene of *Saccharomyces cerevisiae*. Cloning, sequence and deletion analysis of the gene. Eur J Biochem. 1993;217:487-92.
85. Hoja U, Marthol S, Hofmann J, Stegner S, Schulz R, Meier S, et al. *HFA1* encoding an organelle-specific acetyl-CoA carboxylase controls mitochondrial fatty acid synthesis in *Saccharomyces cerevisiae*. J Biol Chem. 2004;279:21779-86.
86. Cronan JE, Jr., Wallace JC. The gene encoding the biotin-apoprotein ligase of *Saccharomyces cerevisiae*. Fems Microbiol Lett. 1995;130:221-9.
87. Walker ME, Val DL, Rohde M, Devenish RJ, Wallace JC. Yeast pyruvate carboxylase: identification of two genes encoding isoenzymes. Biochemical and Biophysical Research Communications. 1991;176:1210-7.
88. Genbauffe FS, Cooper TG. The urea amidolyase (*DUR1,2*) gene of *Saccharomyces cerevisiae*. DNA Sequence. 1991;2:19-32.
89. di Salvo ML, Contestabile R, Safo MK. Vitamin B6 salvage enzymes: mechanism, structure and regulation. Biochimica and Biophysica Acta. 2011;1814:1597-608.
90. Stolz J, Vielreicher M. Tpn1p, the plasma membrane vitamin B6 transporter of *Saccharomyces cerevisiae*. J Biol Chem. 2003;278:18990-6.
91. Loubbardi A, Marcireau C, Karst F, Guilloton M. Sterol uptake induced by an impairment of pyridoxal phosphate synthesis in *Saccharomyces cerevisiae*: cloning and sequencing of the *PDX3* gene encoding pyridoxine (pyridoxamine) phosphate oxidase. J Bacteriol. 1995;177:1817-23.
92. Dong YX, Sueda S, Nikawa J, Kondo H. Characterization of the products of the genes *SNO1* and *SNZ1* involved in pyridoxine synthesis in *Saccharomyces cerevisiae*. Eur J Biochem. 2004;271:745-52.
93. Raschle T, Amrhein N, Fitzpatrick TB. On the two components of pyridoxal 5'-phosphate synthase from *Bacillus subtilis*. J Biol Chem. 2005;280:32291-300.
94. Bauer JA, Bennett EM, Begley TP, Ealick SE. Three-dimensional structure of YaaE from *Bacillus subtilis*, a glutaminase implicated in pyridoxal-5'-phosphate biosynthesis. J Biol Chem. 2004;279:2704-11.
95. Hanes JW, Burns KE, Hilmey DG, Chatterjee A, Dorrestein PC, Begley TP. Mechanistic studies on pyridoxal phosphate synthase: the reaction pathway leading to a chromophoric intermediate. Journal of American Chemical Society.

2008;130:3043-52.
96. Zhang X, Teng YB, Liu JP, He YX, Zhou K, Chen Y, et al. Structural insights into the catalytic mechanism of the yeast pyridoxal 5-phosphate synthase Snz1. Biochemical Journal. 2010;432:445-50.
97. Padilla PA, Fuge EK, Crawford ME, Errett A, Werner-Washburne M. The highly conserved, coregulated *SNO* and *SNZ* gene families in *Saccharomyces cerevisiae* respond to nutrient limitation. J Bacteriol. 1998;180:5718-26.
98. Natarajan K, Meyer MR, Jackson BM, Slade D, Roberts C, Hinnebusch AG, et al. Transcriptional profiling shows that Gcn4p is a master regulator of gene expression during amino acid starvation in yeast. Molecular and Cellular Biology. 2001;21:4347-68.
99. Tice-Baldwin K, Fink GR, Arndt KT. *BAS1* has a Myb motif and activates *HIS4* transcription only in combination with *BAS2*. Science. 1989;246:931-5.
100. Daignan-Fornier B, Fink GR. Coregulation of purine and histidine biosynthesis by the transcriptional activators *BAS1* and *BAS2*. Proceedings of the National Academy of Sciences, USA. 1992;89:6746-50.
101. Mieczkowski PA, Dominska M, Buck MJ, Gerton JL, Lieb JD, Petes TD. Global analysis of the relationship between the binding of the Bas1p transcription factor and meiosis-specific double-strand DNA breaks in *Saccharomyces cerevisiae*. Molecular and Cellular Biology. 2006;26:1014-27.
102. Subramanian M, Qiao WB, Khanam N, Wilkins O, Der SD, Lalich JD, et al. Transcriptional regulation of the one-carbon metabolism regulon in *Saccharomyces cerevisiae* by Bas1p. Mol Microbiol. 2005;57:53-69.
103. Paxhia MD, Downs DM. *SNZ3* encodes a PLP synthase involved in thiamine synthesis in *Saccharomyces cerevisiae*. Genes | Genomes | Genetics (Bethesda). 2019;9:335-44.
104. Rodriguez-Navarro S, Llorente B, Rodriguez-Manzaneque MT, Ramne A, Uber G, Marchesan D, et al. Functional analysis of yeast gene families involved in metabolism of vitamins B1 and B6. Yeast. 2002;19:1261-76.
105. Jansen BCP, Donath WF. On the isolation of the anti-beri-beri vitamin. P K Akad Wet-Amsterd. 1926;29:1390-400.
106. Williams RR, Waterman RE, Keresztesy JC. Larger yields of crystalline antineuritic vitamin. Journal of the American Chemical Society. 1934;56:1187-91.
107. Weiss S, Wilkins RW. The nature of the cardiovascular disturbances in nutritional deficiency states (beriberi). Ann Intern Med. 1937;11:104-48.
108. Nemeria NS, Chakraborty S, Balakrishnan A, Jordan F. Reaction mechanisms of thiamin diphosphate enzymes: defining states of ionization and tautomerization of the cofactor at individual steps. Febs J. 2009;276:2432-46.
109. Wolak N, Kowalska E, Kozik A, Rapala-Kozik M. Thiamine increases the resistance of baker's yeast *Saccharomyces cerevisiae* against oxidative, osmotic and thermal stress, through mechanisms partly independent of thiamine diphosphate-bound enzymes. Fems Yeast Res. 2014;14:1249-62.
110. Muller IB, Bergmann B, Groves MR, Couto I, Amaral L, Begley TP, et al. The vitamin B1 metabolism of *Staphylococcus aureus* is controlled at enzymatic and transcriptional levels. Plos One. 2009;4:e7656.
111. Enjo F, Nosaka K, Ogata M, Iwashima A, Nishimura H. Isolation and characterization of a thiamin transport gene, *THI10*, from *Saccharomyces cerevisiae*. J Biol Chem. 1997;272:19165-70.
112. Nosaka K, Kaneko Y, Nishimura H, Iwashima A. A possible role for acid phosphatase with thiamin-binding activity encoded by *PHO3* in yeast. Fems Microbiol Lett. 1989;51:55-9.
113. Lai RY, Huang S, Fenwick MK, Hazra A, Zhang Y, Rajashankar K, et al. Thiamin pyrimidine biosynthesis in *Candida albicans*: a remarkable reaction between histidine and pyridoxal phosphate. Journal of American Chemical Society. 2012;134:9157-9.
114. Wightman R, Meacock PA. The *THI5* gene family of *Saccharomyces cerevisiae*: distribution of homologues among the hemiascomycetes and functional redundancy in the aerobic biosynthesis of thiamin from pyridoxine. Microbiology. 2003;149:1447-60.
115. Kawasaki Y, Onozuka M, Mizote T, Nosaka K. Biosynthesis of hydroxymethylpyrimidine pyrophosphate in *Saccharomyces cerevisiae*. Current Genetics. 2005;47:156-62.
116. Haas AL, Laun NP, Begley TP. Thi20, a remarkable enzyme from *Saccharomyces cerevisiae* with dual thiamin biosynthetic and degradation activities. Bioorganic Chemistry. 2005;33:338-44.
117. Tanaka K, Tazuya K, Yamada K, Kumaoka H. Biosynthesis of thiamin under anaerobic conditions in *Saccharomyces cerevisiae*. Biol Pharm Bull. 2000;23:108-11.

118. Praekelt UM, Byrne KL, Meacock PA. Regulation of *THI4* (*MOL1*), a thiamine-biosynthetic gene of *Saccharomyces cerevisiae*. Yeast. 1994;10:481-90.
119. Chatterjee A, Jurgenson CT, Schroeder FC, Ealick SE, Begley TP. Biosynthesis of thiamin thiazole in eukaryotes: conversion of NAD to an advanced intermediate. Journal of American Chemical Society. 2007;129:2914-22.
120. Chatterjee A, Schroeder FC, Jurgenson CT, Ealick SE, Begley TP. Biosynthesis of the thiamin-thiazole in eukaryotes: identification of a thiazole tautomer intermediate. Journal of American Chemical Society. 2008;130:11394-8.
121. Chatterjee A, Abeydeera ND, Bale S, Pai PJ, Dorrestein PC, Russell DH, et al. *Saccharomyces cerevisiae* THI4p is a suicide thiamine thiazole synthase. Nature. 2011;478:542-6.
122. Muller EH, Richards EJ, Norbeck J, Byrne KL, Karlsson KA, Pretorius GH, et al. Thiamine repression and pyruvate decarboxylase autoregulation independently control the expression of the *Saccharomyces cerevisiae PDC5* gene. FEBS Letters. 1999;449:245-50.
123. Machado CR, Praekelt UM, de Oliveira RC, Barbosa AC, Byrne KL, Meacock PA, et al. Dual role for the yeast *THI4* gene in thiamine biosynthesis and DNA damage tolerance. Journal of Molecular Biology. 1997;273:114-21.
124. Hohmann S, Meacock PA. Thiamin metabolism and thiamin diphosphate-dependent enzymes in the yeast *Saccharomyces cerevisiae*: genetic regulation. Biochimica and Biophysica Acta. 1998;1385:201-19.
125. Nosaka K. Recent progress in understanding thiamin biosynthesis and its genetic regulation in *Saccharomyces cerevisiae*. Appl Microbiol Biot. 2006;72:30-40.
126. Nishimura H, Kawasaki Y, Nosaka K, Kaneko Y, Iwashima A. A constitutive thiamine metabolism mutation, thi80, causing reduced thiamine pyrophosphokinase activity in *Saccharomyces cerevisiae*. J Bacteriol. 1991;173:2716-9.
127. Nishimura H, Kawasaki Y, Kaneko Y, Nosaka K, Iwashima A. Cloning and characteristics of a positive regulatory gene, *THI2* (*PHO6*), of thiamin biosynthesis in *Saccharomyces cerevisiae*. FEBS Letters. 1992;297:155-8.
128. Hohmann S. Characterisation of *PDC2*, a gene necessary for high level expression of pyruvate decarboxylase structural genes in *Saccharomyces cerevisiae*. Mol Gen Genet. 1993;241:657-66.
129. Nosaka K, Onozuka M, Konno H, Kawasaki Y, Nishimura H, Sano M, et al. Genetic regulation mediated by thiamin pyrophosphate-binding motif in *Saccharomyces cerevisiae*. Mol Microbiol. 2005;58:467-79.
130. Dickinson JR, Lanterman MM, Danner DJ, Pearson BM, Sanz P, Harrison SJ, et al. A ^{13}C nuclear magnetic resonance investigation of the metabolism of leucine to isoamyl alcohol in *Saccharomyces cerevisiae*. J Biol Chem. 1997;272:26871-8.
131. Romagnoli G, Luttik MA, Kotter P, Pronk JT, Daran JM. Substrate specificity of thiamine pyrophosphate-dependent 2-oxo-acid decarboxylases in *Saccharomyces cerevisiae*. Appl Environ Microb. 2012;78:7538-48.
132. Vuralhan Z, Luttik MA, Tai SL, Boer VM, Morais MA, Schipper D, et al. Physiological characterization of the *ARO10*-dependent, broad-substrate-specificity 2-oxo acid decarboxylase activity of *Saccharomyces cerevisiae*. Appl Environ Microb. 2005;71:3276-84.
133. Hazelwood LA, Daran JM, van Maris AJ, Pronk JT, Dickinson JR. The Ehrlich pathway for fusel alcohol production: a century of research on *Saccharomyces cerevisiae* metabolism. Appl Environ Microb. 2008;74:2259-66.
134. Mojzita D, Hohmann S. Pdc2 coordinates expression of the *THI* regulon in the yeast *Saccharomyces cerevisiae*. Molecular Genetics and Genomics. 2006;276:147-61.
135. Li M, Petteys BJ, McClure JM, Valsakumar V, Bekiranov S, Frank EL, et al. Thiamine biosynthesis in *Saccharomyces cerevisiae* is regulated by the NAD^+-dependent histone deacetylase Hst1. Molecular and Cellular Biology. 2010;30:3329-41.
136. Kogl F, Kostermans DGFB. The consitutition-specificity of the hetero-auxins. 16. Announcement on the plant growth materials. H-S Z Physiol Chem. 1935;235:201-16.
137. Gyorgy P, Melville DB, Burk D, V DUV. The possible identity of vitamin H with biotin and coenzyme R. Science. 1940;91:243-5.
138. Lardy HA, Potter RL, Harris RH. Metabolic functions of biotin; the role of biotin in bicarbonate utilization by *Lactobacillus arabinosus* studied with ^{14}C. J Biol Chem. 1949;179:721-31.
139. Wakil SJ, Titchener EB, Gibson DM. Evidence for the participation of biotin in the enzymic synthesis of fatty acids. Biochimica and Biophysica Acta. 1958;29:225-6.
140. Morris CP, Lim F, Wallace JC. Yeast pyruvate carboxylase: gene isolation. Biochemical and Biophysical Research Com-

munications. 1987;145:390-6.
141. Roon RJ, Hampshire J, Levenberg B. Urea amidolyase. The involvement of biotin in urea cleavage. J Biol Chem. 1972;247:7539-45.
142. Kim HS, Hoja U, Stolz J, Sauer G, Schweizer E. 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.
143. Mishina M, Roggenkamp R, Schweizer E. Yeast mutants defective in acetyl-coenzyme A carboxylase and biotin: apocarboxylase ligase. Eur J Biochem. 1980;111:79-87.
144. Hoja U, Wellein C, Greiner E, Schweizer E. Pleiotropic phenotype of acetyl-CoA-carboxylase-defective yeast cells - viability of a *BPL1*-amber mutation depending on its readthrough by normal tRNA(Gln)(CAG). Eur J Biochem. 1998;254:520-6.
145. Suomi F, Menger KE, Monteuuis G, Naumann U, Kursu VA, Shvetsova A, et al. Expression and evolution of the non-canonically translated yeast mitochondrial acetyl-CoA carboxylase Hfa1p. Plos One. 2014;9:e114738.
146. Stolz J, Hoja U, Meier S, Sauer N, Schweizer E. 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.
147. Berkovitch F, Nicolet Y, Wan JT, Jarrett JT, Drennan CL. Crystal structure of biotin synthase, an S-adenosylmethionine-dependent radical enzyme. Science. 2004;303:76-9.
148. Jarrett JT. The novel structure and chemistry of iron-sulfur clusters in the adenosylmethionine-dependent radical enzyme biotin synthase. Archives in Biochemistry and Biophysics. 2005;433:312-21.
149. 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.
150. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, et al. Life with 6000 genes. Science. 1996;274:546, 63-7.
151. Nijkamp JF, van den Broek M, Datema E, de Kok S, Bosman L, Luttik MA, 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. Microbial Cell Factories. 2012;11:36.
152. Wu H, Ito K, Shimoi H. Identification and characterization of a novel biotin biosynthesis gene in *Saccharomyces cerevisiae*. Appl Environ Microb. 2005;71:6845-55.
153. Barbosa R, Pontes A, Santos RO, Montandon GG, de Ponzzes-Gomes CM, Morais PB, et al. Multiple rounds of artificial selection promote microbe secondary domestication - The case of cachaça yeasts. Genome Biology & Evolution. 2018;10:1939-55.
154. Bracher JM, de Hulster E, Koster CC, van den Broek M, Daran JG, van Maris AJA, et al. Laboratory evolution of a biotin-requiring *Saccharomyces cerevisiae* strain for full biotin prototrophy and identification of causal mutations. Appl Environ Microbiol. 2017;83.
155. Manandhar M, Cronan JE. Pimelic acid, the first precursor of the *Bacillus subtilis* biotin synthesis pathway, exists as the free acid and is assembled by fatty acid synthesis. Mol Microbiol. 2017;104:595-607.
156. Ohsugi M, Imanishi Y. Microbiological activity of biotin-vitamers. J Nutr Sci Vitaminol (Tokyo). 1985;31:563-72.
157. Wronska AK, Haak MP, Geraats E, Slot EB, van den Broek M, Pronk JT, et al. Exploiting the diversity of Saccharomycotina yeasts to engineer biotin-independent growth of *Saccharomyces cerevisiae*. Appl Environ Microbiol. 2020;86:e00270-20.
158. Weider M, Machnik A, Klebl F, Sauer N. Vhr1p, a new transcription factor from budding yeast, regulates biotin-dependent expression of *VHT1* and *BIO5*. J Biol Chem. 2006;281:13513-24.
159. Shakoury-Elizeh M, Tiedeman J, Rashford J, Ferea T, Demeter J, Garcia E, et al. Transcriptional remodeling in response to iron deprivation in *Saccharomyces cerevisiae*. Molecular Biology of the Cell. 2004;15:1233-43.
160. Williams RJ, Lyman CM, Goodyear GH, Truesdail JH, Holaday D. "Pantothenic acid", a growth determinant of universal biological occurrence. Journal of the American Chemical Society. 1933;55:2912-27.
161. Sugama S. Sake brewery yeast (II): requirement of pantothenic acid in a medium containing inorganic nitrogen source. Journal of the Brewing Society of Japan 1965;60:453-6.
162. White WH, Gunyuzlu PL, Toyn JH. *Saccharomyces cerevisiae* is capable of *de novo* pantothenic acid biosynthesis involving a novel pathway of β-alanine production from spermine. J Biol Chem. 2001;276:10794-800.

163. Schadeweg V, Boles E. Increasing n-butanol production with *Saccharomyces cerevisiae* by optimizing acetyl-CoA synthesis, NADH levels and trans-2-enoyl-CoA reductase expression. Biotechnology for Biofuels. 2016;9:257.
164. Lussier M, White AM, Sheraton J, di Paolo T, Treadwell J, Southard SB, et al. Large scale identification of genes involved in cell surface biosynthesis and architecture in *Saccharomyces cerevisiae*. Genetics. 1997;147:435-50.
165. Patil KR, Nielsen J. Uncovering transcriptional regulation of metabolism by using metabolic network topology. Proceedings of the national Academy of Sciences, USA. 2005;102:2685-9.
166. Knijnenburg TA, de Winde JH, Daran JM, Daran-Lapujade P, Pronk JT, Reinders MJ, et al. Exploiting combinatorial cultivation conditions to infer transcriptional regulation. BMC Genomics. 2007;8:25.
167. Knijnenburg TA, Daran JM, van den Broek MA, Daran-Lapujade PA, de Winde JH, Pronk JT, et al. Combinatorial effects of environmental parameters on transcriptional regulation in *Saccharomyces cerevisiae*: a quantitative analysis of a compendium of chemostat-based transcriptome data. BMC Genomics. 2009;10:53.
168. Olzhausen J, Schubbe S, Schuller HJ. Genetic analysis of coenzyme A biosynthesis in the yeast *Saccharomyces cerevisiae*: identification of a conditional mutation in the pantothenate kinase gene *CAB1*. Current Genetics. 2009;55:163-73.
169. Behaghel O, Rothman S, Schultze W. Beziehungen zwischen selektiver Ultraviolett-Absorption mit chemischer Konstitution. Strahlentherapie. 1928;28:110-4.
170. Rothman S. Die Beeinflussung der Lichentzündung und der Pigmentierung durch Novokain-Einspritzungen. Strahlentherapie. 1926;22:729-35.
171. Mackie BS, Mackie LE. The PABA story. Australasian Journal of Dermatology. 1999;40:51-3.
172. Rothman S, Rubin J. Sunburn and *para*-aminobenzoic acid. Journal of investigation in dermatology. 1942;5:445-57.
173. Brown GM, Weisman RA, Molnar DA. The biosynthesis of folic acid: I. Substrate and cofactor requirements for enzymatic synthesis by cell-free extracts of *Escherichia coli*. J Biol Chem. 1961;236:2534-43.
174. Pierrel F, Hamelin O, Douki T, Kieffer-Jaquinod S, Muhlenhoff U, Ozeir M, et al. Involvement of mitochondrial ferredoxin and *para*-aminobenzoic acid in yeast coenzyme Q biosynthesis. Chemical Biology. 2010;17:449-59.
175. Marbois B, Xie LX, Choi S, Hirano K, Hyman K, Clarke CF. *para*-Aminobenzoic acid is a precursor in coenzyme Q6 biosynthesis in *Saccharomyces cerevisiae*. J Biol Chem. 2010;285:27827-38.
176. Braus GH. Aromatic amino acid biosynthesis in the yeast *Saccharomyces cerevisiae*: a model system for the regulation of a eukaryotic biosynthetic pathway. Microbiological Reviews. 1991;55:349-70.
177. Luttik MA, Vuralhan Z, Suir E, Braus GH, Pronk JT, Daran JM. Alleviation of feedback inhibition in *Saccharomyces cerevisiae* aromatic amino acid biosynthesis: quantification of metabolic impact. Metab Eng. 2008;10:141-53.
178. Ambroset C, Petit M, Brion C, Sanchez I, Delobel P, Guerin C, et al. Deciphering the molecular basis of wine yeast fermentation traits using a combined genetic and genomic approach. G3 (Bethesda). 2011;1:263-81.
179. Steyer D, Ambroset C, Brion C, Claudel P, Delobel P, Sanchez I, et al. QTL mapping of the production of wine aroma compounds by yeast. BMC Genomics. 2012;13:573.
180. Averesch NJ, Winter G, Kromer JO. Production of *para*-aminobenzoic acid from different carbon-sources in engineered *Saccharomyces cerevisiae*. Microbial Cell Factories. 2016;15:89.
181. Elvehjem CA, Madden RJ, Strong FM, Woolley DW. The isolation and identification of the anti-black tongue factor. J Biol Chem. 1938;123:137-49.
182. Bedalov A, Hirao M, Posakony J, Nelson M, Simon JA. NAD^+-dependent deacetylase Hst1p controls biosynthesis and cellular NAD^+ levels in *Saccharomyces cerevisiae*. Molecular and Cellular Biology. 2003;23:7044-54.
183. Lin SJ, Guarente L. Nicotinamide adenine dinucleotide, a metabolic regulator of transcription, longevity and disease. Current Opinion in Cell Biology. 2003;15:241-6.
184. Preiss J, Handler P. Biosynthesis of diphosphopyridine nucleotide. I. Identification of intermediates. J Biol Chem. 1958;233:488-92.
185. Preiss J, Handler P. Enzymatic synthesis of nicotinamide mononucleotide. J Biol Chem. 1957;225:759-70.
186. Klebl F, Zillig M, Sauer N. Transcription of the yeast *TNA1* gene is not only regulated by nicotinate but also by *p*-aminobenzoate. FEBS Letters. 2000;481:86-7.
187. Llorente B, Dujon B. Transcriptional regulation of the *Saccharomyces cerevisiae DAL5* gene family and identification of the high affinity nicotinic acid permease *TNA1* (YGR260w). FEBS Lett. 2000;475:237-41.
188. Bieganowski P, Brenner C. Discoveries of nicotinamide riboside as a nutrient and conserved *NRK* genes establish a

Preiss-Handler independent route to NAD^+ in fungi and humans. Cell. 2004;117:495-502.
189. Belenky P, Racette FG, Bogan KL, McClure JM, Smith JS, Brenner C. Nicotinamide riboside promotes Sir2 silencing and extends lifespan via Nrk and Urh1/Pnp1/Meu1 pathways to NAD^+. Cell. 2007;129:473-84.
190. Tempel W, Rabeh WM, Bogan KL, Belenky P, Wojcik M, Seidle HF, et al. Nicotinamide riboside kinase structures reveal new pathways to NAD^+. PLoS Biology. 2007;5:e263.
191. Voet D, Voet JG, Pratt CW. Fundamentals of biochemistry life at the molecular level: John Wiley and Sons; 2006.
192. Wierman MB, Smith JS. Yeast sirtuins and the regulation of aging. Fems Yeast Res. 2014;14:73-88.
193. Culver GM, McCraith SM, Consaul SA, Stanford DR, Phizicky EM. A 2'-phosphotransferase implicated in tRNA splicing is essential in *Saccharomyces cerevisiae*. J Biol Chem. 1997;272:13203-10.
194. Rusche LN, Kirchmaier AL, Rine J. The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae*. Annual Review of Biochemistry. 2003;72:481-516.
195. Bürkle A. Poly (ADP-ribose). The most elaborate metabolite of NAD^+. Febs J. 2005;272:4576-89.
196. Chini EN. CD38 as a regulator of cellular NAD: a novel potential pharmacological target for metabolic conditions. Current Pharmaceutical Design. 2009;15:57-63.
197. Kato M, Lin SJ. Regulation of NAD^+ metabolism, signaling and compartmentalization in the yeast *Saccharomyces cerevisiae*. DNA Repair. 2014;23:49-58.
198. Kawai S, Suzuki S, Mori S, Murata K. Molecular cloning and identification of *UTR1* of a yeast *Saccharomyces cerevisiae* as a gene encoding an NAD kinase. Fems Microbiol Lett. 2001;200:181-4.
199. Laurenson P, Rine J. *SUM1-1*: a suppressor of silencing defects in *Saccharomyces cerevisiae*. Genetics. 1991;129:685-96.
200. McCord R, Pierce M, Xie J, Wonkatal S, Mickel C, Vershon AK. Rfm1, a novel tethering factor required to recruit the Hst1 histone deacetylase for repression of middle sporulation genes. Molecular and Cellular Biology. 2003;23:2009-16.
201. James Theoga Raj C, Croft T, Venkatakrishnan P, Groth B, Dhugga G, Cater T, et al. The copper-sensing transcription factor Mac1, the histone deacetylase Hst1, and nicotinic acid regulate de novo NAD^+ biosynthesis in budding yeast. J Biol Chem. 2019;294:5562-75.
202. Lu SP, Lin SJ. Regulation of yeast sirtuins by NAD^+ metabolism and calorie restriction. Biochimica and Biophysica Acta. 2010;1804:1567-75.
203. Pinson B, Ceschin J, Saint-Marc C, Daignan-Fornier B. Dual control of NAD^+ synthesis by purine metabolites in yeast. Elife. 2019;8.
204. Lu SP, Lin SJ. Phosphate-responsive signaling pathway is a novel component of NAD^+ metabolism in *Saccharomyces cerevisiae*. J Biol Chem. 2011;286:14271-81.
205. Boswell-Casteel RC, Johnson JM, Duggan KD, Roe-Zurz Z, Schmitz H, Burleson C, et al. *FUN26* (function unknown now 26) protein from *Saccharomyces cerevisiae* is a broad selectivity, high affinity, nucleoside and nucleobase transporter. J Biol Chem. 2014;289:24440-51.
206. Eastcott EV. Wildiers' bios - The isolation and identification of "bios I". J Phys Chem-Us. 1928;32:1094-111.
207. White MJ, Lopes JM, Henry SA. Inositol metabolism in yeasts. Advances in Microbial Physiology. 1991;32:1-51.
208. Yoko-o T, Matsui Y, Yagisawa H, Nojima H, Uno I, Toh-e A. The putative phosphoinositide-specific phospholipase C gene, *PLC1*, of the yeast *Saccharomyces cerevisiae* is important for cell growth. Proceedings of the national Academy of Sciences, USA. 1993;90:1804-8.
209. Pittet M, Conzelmann A. Biosynthesis and function of GPI proteins in the yeast *Saccharomyces cerevisiae*. Biochimica and Biophysica Acta. 2007;1771:405-20.
210. Culbertson MR, Donahue TF, Henry SA. Control of inositol biosynthesis in *Saccharomyces cerevisiae*; inositol-phosphate synthetase mutants. J Bacteriol. 1976;126:243-50.
211. Donahue TF, Henry SA. myo-Inositol-1-phosphate synthase. Characteristics of the enzyme and identification of its structural gene in yeast. J Biol Chem. 1981;256:7077-85.
212. Murray M, Greenberg ML. Expression of yeast *INM1* encoding inositol monophosphatase is regulated by inositol, carbon source and growth stage and is decreased by lithium and valproate. Mol Microbiol. 2000;36:651-61.
213. Henry SA, Gaspar ML, Jesch SA. The response to inositol: regulation of glycerolipid metabolism and stress response signaling in yeast. Chemistry and Physics of Lipids. 2014;180:23-43.
214. Henry SA, Kohlwein SD, Carman GM. Metabolism and regulation of glycerolipids in the yeast *Saccharomyces cerevisi-*

ae. Genetics. 2012;190:317-49.
215. Greenberg ML, Reiner B, Henry SA. Regulatory mutations of inositol biosynthesis in yeast: isolation of inositol-excreting mutants. Genetics. 1982;100:19-33.
216. Hancock LC, Behta RP, Lopes JM. Genomic analysis of the Opi$^-$ phenotype. Genetics. 2006;173:621-34.
217. Jelier R, Semple JI, Garcia-Verdugo R, Lehner B. Predicting phenotypic variation in yeast from individual genome sequences. Nature Genetics. 2011;43:1270-4.
218. Eddy SR. Accelerated Profile HMM Searches. PLoS Computational Biology. 2011;7:e1002195.
219. Winzeler EA, Castillo-Davis CI, Oshiro G, Liang D, Richards DR, Zhou YY, et al. Genetic diversity in yeast assessed with whole-genome oligonucleotide arrays. Genetics. 2003;163:79-89.
220. Baker E, Wang B, Bellora N, Peris D, Hulfachor AB, Koshalek JA, et al. The genome sequence of *Saccharomyces eubayanus* and the domestication of lager-brewing yeasts. Mol Biol Evol. 2015;32:2818-31.
221. Scannell DR, Zill OA, Rokas A, Payen C, Dunham MJ, Eisen MB, et al. The awesome power of yeast evolutionary genetics: new genome sequences and strain resources for the *Saccharomyces sensu stricto* genus. Genes | Genomes | Genetics (Bethesda). 2011;1:11-25.
222. Liti G, Nguyen Ba AN, Blythe M, Muller CA, Bergstrom A, Cubillos FA, et al. High quality *de novo* sequencing and assembly of the *Saccharomyces arboricolus* genome. BMC Genomics. 2013;14:69.
223. Naseeb S, James SA, Alsammar H, Michaels CJ, Gini B, Nueno-Palop C, et al. *Saccharomyces jurei* sp. nov., isolation and genetic identification of a novel yeast species from *Quercus robur*. International Journal of Systematic and Evolutionary Microbiology. 2017;67:2046-52.
224. Yue JX, Li J, Aigrain L, Hallin J, Persson K, Oliver K, et al. Contrasting evolutionary genome dynamics between domesticated and wild yeasts. Nature Genetics. 2017;49:913-24.
225. Salazar AN, Gorter de Vries AR, van den Broek M, Wijsman M, de la Torre Cortes P, Brickwedde A, et al. Nanopore sequencing enables near-complete *de novo* assembly of *Saccharomyces cerevisiae* reference strain CEN.PK113-7D. FEMS Yeast Res. 2017;17.
226. Broderick JB. Coenzymes and Cofactors. eLS2001.
227. Milne N, Luttik MAH, Cueto Rojas HF, Wahl A, van Maris AJA, Pronk JT, et al. Functional expression of a heterologous nickel-dependent, ATP-independent urease in *Saccharomyces cerevisiae*. Metab Eng. 2015;30:130-40.
228. Steinman HM. The amino acid sequence of copper-zinc superoxide dismutase from bakers' yeast. J Biol Chem. 1980;255:6758-65.
229. Berger L, Slein MW, Colowick SP, Cori CF. Isolation of Hexokinase from baker's Yeast. J Gen Physiol. 1946;29:379-91.
230. Green SM, Ginsburg A, Lewis MS, Hensley P. Roles of metal ions in the maintenance of the tertiary and quaternary structure of arginase from *Saccharomyces cerevisiae*. J Biol Chem. 1991;266:21474-81.
231. Vallee BL, Hoch FL. Zinc, a component of yeast alcohol dehydrogenase. Proc Natl Acad Sci U S A. 1955;41:327-38.
232. Zhang S, Sanyal I, Bulboaca GH, Rich A, Flint DH. The gene for biotin synthase from *Saccharomyces cerevisiae*: cloning, sequencing, and complementation of *Escherichia coli* strains lacking biotin synthase. Arch Biochem Biophys. 1994;309:29-35.
233. Buren S, Pratt K, Jiang X, Guo Y, Jimenez-Vicente E, Echavarri-Erasun C, et al. Biosynthesis of the nitrogenase active-site cofactor precursor NifB-co in *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A. 2019;116:25078-86.
234. Nogae I, Johnston M. Isolation and characterization of the *ZWF1* gene of *Saccharomyces cerevisiae*, encoding glucose-6-phosphate dehydrogenase. Gene. 1990;96:161-9.
235. Fang H, Kang J, Zhang D. Microbial production of vitamin B12: a review and future perspectives. Microb Cell Fact. 2017;16:15.
236. Lomakin IB, Xiong Y, Steitz TA. The crystal structure of yeast fatty acid synthase, a cellular machine with eight active sites working together. Cell. 2007;129:319-32.
237. Kastanos EK, Woldman YY, Appling DR. Role of mitochondrial and cytoplasmic serine hydroxymethyltransferase isozymes in *de novo* purine synthesis in *Saccharomyces cerevisiae*. Biochemistry. 1997;36:14956-64.
238. Akbari S, Rasouli-Ghahroudi AA. Vitamin K and bone metabolism: a review of the latest evidence in preclinical studies. Biomed Res Int. 2018;2018:4629383.
239. Branduardi P, Fossati T, Sauer M, Pagani R, Mattanovich D, Porro D. Biosynthesis of vitamin C by yeast leads to in-

creased stress resistance. Plos One. 2007;2:e1092.
240. Satoh T, Horie M, Watanabe H, Tsuchiya Y, Kamei T. Enzymatic properties of squalene epoxidase from *Saccharomyces cerevisiae*. Biol Pharm Bull. 1993;16:349-52.
241. Nagy M, Lacroute F, Thomas D. Divergent evolution of pyrimidine biosynthesis between anaerobic and aerobic yeasts. Proc Natl Acad Sci U S A. 1992;89:8966-70.
242. Shima S, Warkentin E, Grabarse W, Sordel M, Wicke M, Thauer RK, et al. Structure of coenzyme F(420) dependent methylenetetrahydromethanopterin reductase from two methanogenic archaea. J Mol Biol. 2000;300:935-50.
243. Lobo Z, Maitra PK. Physiological role of glucose-phosphorylating enzymes in *Saccharomyces cerevisiae*. Arch Biochem Biophys. 1977;182:639-45.
244. Thomas D, Becker A, Surdin-Kerjan Y. Reverse methionine biosynthesis from S-adenosylmethionine in eukaryotic cells. J Biol Chem. 2000;275:40718-24.
245. Ellermann J, Hedderich R, Bocher R, Thauer RK. The final step in methane formation. Investigations with highly purified methyl-CoM reductase (component C) from *Methanobacterium thermoautotrophicum* (strain Marburg). Eur J Biochem. 1988;172:669-77.
246. Marres CA, de Vries S, Grivell LA. Isolation and inactivation of the nuclear gene encoding the rotenone-insensitive internal NADH:ubiquinone oxidoreductase of mitochondria from *Saccharomyces cerevisiae*. Eur J Biochem. 1991;195:857-62.
247. Han GS, O'Hara L, Siniossoglou S, Carman GM. Characterization of the yeast *DGK1*-encoded CTP-dependent diacylglycerol kinase. J Biol Chem. 2008;283:20443-53.
248. Gan ZR, Polokoff MA, Jacobs JW, Sardana MK. Complete amino acid sequence of yeast thioltransferase (glutaredoxin). Biochem Biophys Res Commun. 1990;168:944-51.
249. Rytka J, Sledziewski A, Lukaszkiewicz J, Bilinski T. Haemoprotein formation in yeast. III. The role of carbon catabolite repression in the regulation of catalase A and T formation. Mol Gen Genet. 1978;160:51-7.
250. Hirabayashi T, Harada T. Isolation and properties of alpha-ketoglutarate dehydrogenase complex from baker's yeast (Saccharomyces cerevisiae). Biochemical and Biophysical Research Communications. 1971;45:1369-75.
251. Schmitz RA, Albracht SP, Thauer RK. A molybdenum and a tungsten isoenzyme of formylmethanofuran dehydrogenase in the thermophilic archaeon *Methanobacterium wolfei*. Eur J Biochem. 1992;209:1013-8.
252. Linder T. A genomic survey of nitrogen assimilation pathways in budding yeasts (sub-phylum Saccharomycotina). Yeast. 2019;36:259-73.
253. Neumann M, Mittelstadt G, Iobbi-Nivol C, Saggu M, Lendzian F, Hildebrandt P, et al. A periplasmic aldehyde oxidoreductase represents the first molybdopterin cytosine dinucleotide cofactor containing molybdo-flavoenzyme from *Escherichia coli*. FEBS J. 2009;276:2762-74.
254. Hartmann T, Schwanhold N, Leimkuhler S. Assembly and catalysis of molybdenum or tungsten-containing formate dehydrogenases from bacteria. Bba-Proteins Proteom. 2015;1854:1090-100.
255. Gopal PK, Ballou CE. Regulation of the protein glycosylation pathway in yeast: structural control of N-linked oligosaccharide elongation. Proc Natl Acad Sci U S A. 1987;84:8824-8.
256. Leyh TS, Cook I, Wang T. Structure, dynamics and selectivity in the sulfotransferase family. Drug Metab Rev. 2013;45:423-30.
257. Oubrie A, Rozeboom HJ, Kalk KH, Olsthoorn AJ, Duine JA, Dijkstra BW. Structure and mechanism of soluble quinoprotein glucose dehydrogenase. EMBO J. 1999;18:5187-94.
258. Galanie S, Thodey K, Trenchard IJ, Filsinger Interrante M, Smolke CD. Complete biosynthesis of opioids in yeast. Science. 2015;349:1095-100.
259. Martinez-Gomez NC, Nguyen S, Lidstrom ME. Elucidation of the role of the methylene-tetrahydromethanopterin dehydrogenase MtdA in the tetrahydromethanopterin-dependent oxidation pathway in *Methylobacterium extorquens* AM1. J Bacteriol. 2013;195:2359-67.
260. Acevedo-Rocha CG, Gronenberg LS, Mack M, Commichau FM, Genee HJ. Microbial cell factories for the sustainable manufacturing of B vitamins. Curr Opin Biotechnol. 2019;56:18-29.
261. Chen R, Yang S, Zhang L, Zhou YJ. Advanced strategies for production of natural products in yeast. iScience. 2020;23:100879.

262. Zhang C, Wu D, Ren H. Economical production of vitamin K2 using crude glycerol from the by-product of biodiesel. Sci Rep. 2020;10:5959.
263. Taylor M, Scott C, Grogan G. F420-dependent enzymes - potential for applications in biotechnology. Trends Biotechnol. 2013;31:63-4.
264. Shah MV, Antoney J, Kang SW, Warden AC, Hartley CJ, Nazem-Bokaee H, et al. Cofactor F-420-dependent enzymes: an under-explored resource for asymmetric redox biocatalysis. Catalysts. 2019;9.
265. Leimkuhler S, Iobbi-Nivol C. Bacterial molybdoenzymes: old enzymes for new purposes. FEMS Microbiol Rev. 2016;40:1-18.
266. Shen XX, Opulente DA, Kominek J, Zhou X, Steenwyk JL, Buh KV, et al. Tempo and mode of genome evolution in the budding yeast subphylum. Cell. 2018;175:1533-45 e20.
267. Perli T, Wronska AK, Ortiz-Merino RA, Pronk JT, Daran JM. Vitamin requirements and biosynthesis in *Saccharomyces cerevisiae*. Yeast. 2020;37:283-304.
268. Mans R, Daran JG, Pronk JT. Under pressure: evolutionary engineering of yeast strains for improved performance in fuels and chemicals production. Curr Opin Biotechnol. 2018;50:47-56.
269. Sandberg TE, Salazar MJ, Weng LL, Palsson BO, Feist AM. The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. Metab Eng. 2019;56:1-16.
270. Jansen MLA, Bracher JM, Papapetridis I, Verhoeven MD, de Bruijn H, de Waal PP, et al. *Saccharomyces cerevisiae* strains for second-generation ethanol production: from academic exploration to industrial implementation. FEMS Yeast Res. 2017;17.
271. Lynd LR, Elamder RT, Wyman CE. Likely features and costs of mature biomass ethanol technology. Applied Biochemistry and Biotechnology. 1996;57:741-61.
272. Skinner KA, Leathers TD. Bacterial contaminants of fuel ethanol production. J Ind Microbiol Biotechnol. 2004;31:401-8.
273. Entian KD, Kötter P. 25 yeast genetic strain and plasmid collections. Methods in Microbiology. 2007;36:629-66.
274. van Dijken JP, Bauer J, Brambilla L, Duboc P, Francois JM, Gancedo C, et al. An interlaboratory comparison of physiological and genetic properties of four Saccharomyces cerevisiae strains. Enzyme Microb Technol. 2000;26:706-14.
275. Hassing EJ, de Groot PA, Marquenie VR, Pronk JT, Daran JG. Connecting central carbon and aromatic amino acid metabolisms to improve *de novo* 2-phenylethanol production in *Saccharomyces cerevisiae*. Metab Eng. 2019;56:165-80.
276. Papapetridis I, van Dijk M, van Maris AJA, Pronk JT. Metabolic engineering strategies for optimizing acetate reduction, ethanol yield and osmotolerance in Saccharomyces cerevisiae. Biotechnol Biofuels. 2017;10:107.
277. Kuijpers NG, Solis-Escalante D, Luttik MA, Bisschops MM, Boonekamp FJ, van den Broek M, et al. Pathway swapping: Toward modular engineering of essential cellular processes. Proc Natl Acad Sci U S A. 2016;113:15060-5.
278. Solis-Escalante D, Kuijpers NG, Barrajon-Simancas N, van den Broek M, Pronk JT, Daran JM, et al. A Minimal Set of Glycolytic Genes Reveals Strong Redundancies in Saccharomyces cerevisiae Central Metabolism. Eukaryot Cell. 2015;14:804-16.
279. de Kok S, Nijkamp JF, Oud B, Roque FC, de Ridder D, Daran JM, 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.
280. Ball SG, Wickner RB, Cottarel G, Schaus M, Tirtiaux C. Molecular-Cloning and Characterization of Aro7-Osm2, a Single Yeast Gene Necessary for Chorismate Mutase Activity and Growth in Hypertonic Medium. Mol Gen Genet. 1986;205:326-30.
281. Edman JC, Goldstein AL, Erbe JG. *Para*-aminobenzoate synthase gene of *Saccharomyces cerevisiae* encodes a bifunctional enzyme. Yeast. 1993;9:669-75.
282. Rudolph HK, Antebi A, Fink GR, Buckley CM, Dorman TE, LeVitre J, et al. The yeast secretory pathway is perturbed by mutations in *PMR1*, a member of a Ca^{2+} ATPase family. Cell. 1989;58:133-45.
283. Ton VK, Mandal D, Vahadji C, Rao R. Functional expression in yeast of the human secretory pathway Ca^{2+}, Mn^{2+}-ATPase defective in Hailey-Hailey disease. J Biol Chem. 2002;277:6422-7.
284. Cyert MS. Calcineurin signaling in *Saccharomyces cerevisiae*: how yeast go crazy in response to stress. Biochem Biophys Res Commun. 2003;311:1143-50.
285. Lesuisse E, Blaiseau PL, Dancis A, Camadro JM. Siderophore uptake and use by the yeast *Saccharomyces cerevisiae*.

Microbiology. 2001;147:289-98.
286. Yun CW, Tiedeman JS, Moore RE, Philpott CC. Siderophore-iron uptake in *Saccharomyces cerevisiae*. Identification of ferrichrome and fusarinine transporters. J Biol Chem. 2000;275:16354-9.
287. Zhang L, Guarente L. Evidence that *TUP1/SSN6* has a positive effect on the activity of the yeast activator *HAP1*. Genetics. 1994;136:813-7.
288. Han SJ, Lee JS, Kang JS, Kim YJ. Med9/Cse2 and Gal11 modules are required for transcriptional repression of distinct group of genes. J Biol Chem. 2001;276:37020-6.
289. Fazzio TG, Gelbart ME, Tsukiyama T. Two distinct mechanisms of chromatin interaction by the Isw2 chromatin remodeling complex in vivo. Mol Cell Biol. 2005;25:9165-74.
290. Whitney PA, Morris DR. Polyamine auxotrophs of *Saccharomyces cerevisiae*. J Bacteriol. 1978;134:214-20.
291. Joets J, Pousset D, Marcireau C, Karst F. Characterization of the *Saccharomyces cerevisiae FMS1* gene related to *Candida albicans* corticosteroid-binding protein 1. Curr Genet. 1996;30:115-20.
292. Flagfeldt DB, Siewers V, Huang L, Nielsen J. Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. Yeast. 2009;26:545-51.
293. Teixeira MC, Monteiro PT, Palma M, Costa C, Godinho CP, Pais P, et al. YEASTRACT: an upgraded database for the analysis of transcription regulatory networks in *Saccharomyces cerevisiae*. Nucleic Acids Res. 2018;46:D348-D53.
294. Fendt SM, Oliveira AP, Christen S, Picotti P, Dechant RC, Sauer U. Unraveling condition-dependent networks of transcription factors that control metabolic pathway activity in yeast. Mol Syst Biol. 2010;6:432.
295. Fordyce PM, Gerber D, Tran D, Zheng J, Li H, DeRisi JL, et al. De novo identification and biophysical characterization of transcription-factor binding sites with microfluidic affinity analysis. Nat Biotechnol. 2010;28:970-5.
296. Papamichos-Chronakis M, Conlan RS, Gounalaki N, Copf T, Tzamarias D. Hrs1/Med3 is a Cyc8-Tup1 corepressor target in the RNA polymerase II holoenzyme. J Biol Chem. 2000;275:8397-403.
297. Komachi K, Johnson AD. Residues in the WD repeats of Tup1 required for interaction with alpha2. Mol Cell Biol. 1997;17:6023-8.
298. Proft M, Pascual-Ahuir A, de Nadal E, Arino J, Serrano R, Posas F. Regulation of the Sko1 transcriptional repressor by the Hog1 MAP kinase in response to osmotic stress. EMBO J. 2001;20:1123-33.
299. Cooper JP, Roth SY, Simpson RT. The global transcriptional regulators, *SSN6* and *TUP1*, play distinct roles in the establishment of a repressive chromatin structure. Gene Dev. 1994;8:1400-10.
300. Mizuno T, Harashima S. Gal11 is a general activator of basal transcription, whose activity is regulated by the general repressor Sin4 in yeast. Mol Genet Genomics. 2003;269:68-77.
301. Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, et al. Functional profiling of the *Saccharomyces cerevisiae* genome. Nature. 2002;418:387-91.
302. Jedidi I, Zhang F, Qiu H, Stahl SJ, Palmer I, Kaufman JD, et al. Activator Gcn4 employs multiple segments of Med15/Gal11, including the KIX domain, to recruit mediator to target genes *in vivo*. J Biol Chem. 2010;285:2438-55.
303. Tsukiyama T, Palmer J, Landel CC, Shiloach J, Wu C. Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. Gene Dev. 1999;13:686-97.
304. Georgatsou E, Alexandraki D. Two distinctly regulated genes are required for ferric reduction, the first step of iron uptake in Saccharomyces cerevisiae. Mol Cell Biol. 1994;14:3065-73.
305. Antebi A, Fink GR. The yeast Ca-2+-Atpase homolog, Pmr1, is required for normal golgi function and localizes in a novel golgi-like distribution. Mol Biol Cell. 1992;3:633-54.
306. Nosaka K, Nishimura H, Kawasaki Y, Tsujihara T, Iwashima A. Isolation and characterization of the *THI6* gene encoding a bifunctional thiamin-phosphate pyrophosphorylase/hydroxyethylthiazole kinase from *Saccharomyces cerevisiae*. J Biol Chem. 1994;269:30510-6.
307. Llorente B, Fairhead C, Dujon B. Genetic redundancy and gene fusion in the genome of the Baker's yeast *Saccharomyces cerevisiae*: functional characterization of a three-member gene family involved in the thiamine biosynthetic pathway. Mol Microbiol. 1999;32:1140-52.
308. Steynparve EP. Partial purification and properties of thiaminokinase from yeast. Biochim Biophys Acta. 1952;8:310-24.
309. Verhoeven MD, Lee M, Kamoen L, van den Broek M, Janssen DB, Daran JMG, et al. Mutations in *PMR1* stimulate xylose isomerase activity and anaerobic growth on xylose of engineered *Saccharomyces cerevisiae* by influencing manganese

homeostasis. Sci Rep-Uk. 2017;7:46155.
310. Alper H, Moxley J, Nevoigt E, Fink GR, Stephanopoulos G. Engineering yeast transcription machinery for improved ethanol tolerance and production. Science. 2006;314:1565-8.
311. Bachmann H, Bruggeman FJ, Molenaar D, Branco Dos Santos F, Teusink B. Public goods and metabolic strategies. Curr Opin Microbiol. 2016;31:109-15.
312. Nilsson A, Nielsen J. Metabolic Trade-offs in Yeast are Caused by F1F0-ATP synthase. Sci Rep-UK. 2016;6:22264.
313. Mans R, van Rossum HM, Wijsman M, Backx A, Kuijpers NG, van den Broek 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.
314. Looke M, Kristjuhan K, Kristjuhan A. Extraction of genomic DNA from yeasts for PCR-based applications. Biotechniques. 2011;50:325-8.
315. Gietz RD, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzymol. 2002;350:87-96.
316. Inoue H, Nojima H, Okayama H. High efficiency transformation of Escherichia coli with plasmids. Gene. 1990;96:23-8.
317. Lee ME, DeLoache WC, Cervantes B, Dueber JE. A highly characterized yeast toolkit for modular, multipart assembly. ACS Synth Biol. 2015;4:975-86.
318. Nijkamp JF, van den Broek MA, Geertman JM, Reinders MJ, Daran JM, de Ridder D. *De novo* detection of copy number variation by co-assembly. Bioinformatics. 2012;28:3195-202.
319. Mans R, Wijsman M, Daran-Lapujade P, Daran JM. A protocol for introduction of multiple genetic modifications in *Saccharomyces cerevisiae* using CRISPR/Cas9. FEMS Yeast Res. 2018;18.
320. Gruninger RJ, Puniya AK, Callaghan TM, Edwards JE, Youssef N, Dagar SS, et al. Anaerobic fungi (phylum Neocallimastigomycota): advances in understanding their taxonomy, life cycle, ecology, role and biotechnological potential. FEMS Microbiol Ecol. 2014;90:1-17.
321. Wilken SE, Seppala S, Lankiewicz TS, Saxena M, Henske JK, Salamov AA, et al. Genomic and proteomic biases inform metabolic engineering strategies for anaerobic fungi. Metab Eng Commun. 2020;10:e00107.
322. Kemp P, Lander DJ, Orpin CG. The lipids of the rumen fungus *Piromonas communis*. J Gen Microbiol. 1984;130:27-37.
323. Garcia-Vallve S, Romeu A, Palau J. Horizontal gene transfer of glycosyl hydrolases of the rumen fungi. Mol Biol Evol. 2000;17:352-61.
324. Harhangi HR, Akhmanova AS, Emmens R, van der Drift C, de Laat WT, van Dijken JP, et al. Xylose metabolism in the anaerobic fungus *Piromyces sp.* strain E2 follows the bacterial pathway. Arch Microbiol. 2003;180:134-41.
325. Murphy CL, Youssef NH, Hanafy RA, Couger MB, Stajich JE, Wang Y, et al. Horizontal gene transfer as an indispensable driver for evolution of Neocallimastigomycota into a distinct gut-dwelling fungal lineage. Appl Environ Microbiol. 2019;85:e00988-19.
326. Youssef NH, Couger MB, Struchtemeyer CG, Liggenstoffer AS, Prade RA, Najar FZ, et al. The genome of the anaerobic fungus *Orpinomyces* sp. strain C1A reveals the unique evolutionary history of a remarkable plant biomass degrader. Appl Environ Microbiol. 2013;79:4620-34.
327. Wang Y, Youssef NH, Couger MB, Hanafy RA, Elshahed MS, Stajich JE. Molecular dating of the emergence of anaerobic rumen fungi and the impact of laterally acquired genes. mSystems. 2019;4, 10.1128/mSystems.00247-19.
328. Weete JD, Abril M, Blackwell M. Phylogenetic distribution of fungal sterols. Plos One. 2010;5:e10899.
329. Summons RE, Bradley AS, Jahnke LL, Waldbauer JR. Steroids, triterpenoids and molecular oxygen. Philos T R Soc B. 2006;361:951-68.
330. Ourisson G, Rohmer M, Poralla K. Prokaryotic hopanoids and other polyterpenoid sterol surrogates. Annu Rev Microbiol. 1987;41:301-33.
331. Takishita K, Chikaraishi Y, Leger MM, Kim E, Yabuki A, Ohkouchi N, et al. Lateral transfer of tetrahymanol-synthesizing genes has allowed multiple diverse eukaryote lineages to independently adapt to environments without oxygen. Biol Direct. 2012;7:5.
332. Wiersma SJ, Mooiman C, Giera M, Pronk JT. Expression of a squalene-tetrahymanol cyclase enables sterol-independent growth of *Saccharomyces cerevisiae*. Appl Environ Microbiol. 2020.
333. Landry J, Sternglanz R. Yeast Fms1 is a FAD-utilizing polyamine oxidase. Biochem Bioph Res Co. 2003;303:771-6.

334. White WH, Skatrud PL, Xue ZX, Toyn JH. Specialization of function among aldehyde dehydrogenases: The *ALD2* and *ALD3* genes are required for β-alanine biosynthesis in *Saccharomyces cerevisiae*. Genetics. 2003;163:69-77.
335. Orpin CG, Greenwood Y. Nutritional and germination requirements of the rumen chytridiomycete *Neocallimastix patriciarum*. T Brit Mycol Soc. 1986;86:103-9.
336. Hao JF, Petriacq P, de Bont L, Hodges M, Gakiere B. Characterization of L-aspartate oxidase from *Arabidopsis thaliana*. Plant Sci. 2018;271:133-42.
337. Tedeschi G, Negri A, Mortarino M, Ceciliani F, Simonic T, Faotto L, et al. L-aspartate oxidase from *Escherichia coli* . II. Interaction with C4 dicarboxylic acids and identification of a novel L-aspartate:fumarate oxidoreductase activity. Eur J Biochem. 1996;239:427-33.
338. Katoh A, Uenohara K, Akita M, Hashimoto T. Early steps in the biosynthesis of NAD in *Arabidopsis* start with aspartate and occur in the plastid. Plant Physiol. 2006;141:851-7.
339. Arakane Y, Lomakin J, Beeman RW, Muthukrishnan S, Gehrke SH, Kanost MR, et al. Molecular and functional analyses of amino acid decarboxylases involved in cuticle tanning in *Tribolium castaneum*. J Biol Chem. 2009;284:16584-94.
340. Tomita H, Yokooji Y, Ishibashi T, Imanaka T, Atomi H. An archaeal glutamate decarboxylase homolog functions as an aspartate decarboxylase and Is involved in β-alanine and coenzyme A biosynthesis. J Bacteriol. 2014;196:1222-30.
341. Yutin N, Galperin MY. A genomic update on clostridial phylogeny: Gram-negative spore formers and other misplaced clostridia. Environ Microbiol. 2013;15:2631-41.
342. Seshadri R, Joseph SW, Chopra AK, Sha J, Shaw J, Graf J, et al. Genome sequence of Aeromonas hydrophila ATCC 7966T: jack of all trades. J Bacteriol. 2006;188:8272-82.
343. Petitdemange E, Caillet F, Giallo J, Gaudin C. Clostridium cellulolyticum sp. nov., a Cellulolytic, Mesophilic: Species from Decayed Grass. International Journal of Systematic and Evolutionary Microbiology. 1984;34:155-9.
344. McInerney MJ, Rohlin L, Mouttaki H, Kim U, Krupp RS, Rios-Hernandez L, et al. The genome of Syntrophus aciditrophicus: life at the thermodynamic limit of microbial growth. Proc Natl Acad Sci U S A. 2007;104:7600-5.
345. Kuever J. The Family Desulfobacteraceae. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The Prokaryotes: Deltaproteobacteria and Epsilonproteobacteria. Berlin, Heidelberg: Springer Berlin Heidelberg; 2014. p. 45-73.
346. Naranjo-Ortiz MA, Brock M, Brunke S, Hube B, Marcet-Houben M, Gabaldon T. Widespread Inter- and Intra-Domain Horizontal Gene Transfer of d-Amino Acid Metabolism Enzymes in Eukaryotes. Front Microbiol. 2016;7:2001.
347. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics. 2019;35:4453-5.
348. Tavaré S, Miura R. Some mathematical questions in biology: DNA sequence analysis lectures on mathematics in the life sciences. Stat Med. 1986;4:523-4.
349. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci. 1992;8:275-82.
350. Le SQ, Gascuel O. An improved general amino acid replacement matrix. Mol Biol Evol. 2008;25:1307-20.
351. Marcet-Houben M, Gabaldon T. Acquisition of prokaryotic genes by fungal genomes. Trends Genet. 2010;26:5-8.
352. Vermeersch L, Perez-Samper G, Cerulus B, Jariani A, Gallone B, Voordeckers K, et al. On the duration of the microbial lag phase. Curr Genet. 2019;65:721-7.
353. Borodina I, Kildegaard KR, Jensen NB, Blicher TH, Maury J, Sherstyk S, et al. Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in *Saccharomyces cerevisiae* via β-alanine. Metab Eng. 2015;27:57-64.
354. Papapetridis I, van Dijk M, Dobbe APA, Metz B, Pronk JT, van Maris AJA. Improving ethanol yield in acetate-reducing *Saccharomyces cerevisiae* by cofactor engineering of 6-phosphogluconate dehydrogenase and deletion of *ALD6*. Microbial Cell Factories. 2016;15:67.
355. Visser W, Scheffers WA, Batenburg-van der Vegte WH, van Dijken JP. Oxygen requirements of yeasts. Appl Environ Microbiol. 1990;56:3785-92.
356. da Costa BLV, Basso TO, Raghavendran V, Gombert AK. Anaerobiosis revisited: growth of *Saccharomyces cerevisiae* under extremely low oxygen availability. Appl Microbiol Biotechnol. 2018;102:2101-16.
357. Dekker WJC, Wiersma SJ, Bouwknegt J, Mooiman C, Pronk JT. Anaerobic growth of *Saccharomyces cerevisiae* CEN.

PK113-7D does not depend on synthesis or supplementation of unsaturated fatty acids. FEMS Yeast Res. 2019;19.

358. Sandmeier E, Hale TI, Christen P. Multiple evolutionary origin of pyridoxal-5'-phosphate-dependent amino acid decarboxylases. Eur J Biochem. 1994;221:997-1002.
359. Salzmann D, Christen P, Mehta PK, Sandmeier E. Rates of evolution of pyridoxal-5'-phosphate-dependent enzymes. Biochem Biophys Res Commun. 2000;270:576-80.
360. Li Y, Steenwyk JL, Chang Y, Wang Y, James TY, Stajich JE, et al. A genome-scale phylogeny of the kingdom Fungi. Curr Biol. 2021.
361. Schlegel M, Munsterkotter M, Guldener U, Bruggmann R, Duo A, Hainaut M, et al. Globally distributed root endophyte Phialocephala subalpina links pathogenic and saprophytic lifestyles. BMC Genomics. 2016;17:1015.
362. Ramjee MK, Genschel U, Abell C, Smith AG. *Escherichia coli* L-aspartate-α-decarboxylase: preprotein processing and observation of reaction intermediates by electrospray mass spectrometry. Biochem J. 1997;323 (Pt 3):661-9.
363. Tamaki N, Aoyama H, Kubo K, Ikeda T, Hama T. Purification and properties of β-alanine aminotransferase from rabbit liver. J Biochem. 1982;92:1009-17.
364. Hayaishi O, Nishizuka Y, Tatibana M, Takeshita M, Kuno S. Enzymatic studies on the metabolism of β-alanine. J Biol Chem. 1961;236:781-90.
365. Schnackerz KD, Dobritzsch D. Amidohydrolases of the reductive pyrimidine catabolic pathway purification, characterization, structure, reaction mechanisms and enzyme deficiency. Biochim Biophys Acta. 2008;1784:431-44.
366. Dalluge JJ, Liao H, Gokarn R, Jessen H. Discovery of enzymatic activity using stable isotope metabolite labeling and liquid chromatography-mass spectrometry. Analytical Chemistry. 2005;77:6737-40.
367. Gojkovic Z, Sandrini MP, Piskur J. Eukaryotic β-alanine synthases are functionally related but have a high degree of structural diversity. Genetics. 2001;158:999-1011.
368. Ahmad F, Moat AG. Nicotinic acid biosynthesis in prototrophs and tryptophan auxotrophs of *Saccharomyces cerevisiae*. J Biol Chem. 1966;241:775-80.
369. Suomalainen H, Nurminen T, Vihervaara K, Oura E. Effect of aeration on the synthesis of nicotinic acid and nicotinamide adenine dinucleotide by baker's yeast. J Inst Brew. 1965;71:227-31.
370. Kampers LFC, van Heck RGA, Donati S, Saccenti E, Volkers RJM, Schaap PJ, et al. In silico-guided engineering of *Pseudomonas putida* towards growth under micro-oxic conditions. Microb Cell Fact. 2019;18:179.
371. Rousset C, Fontecave M, Ollagnier de Choudens S. The [4Fe-4S] cluster of quinolinate synthase from *Escherichia coli*: investigation of cluster ligands. FEBS Lett. 2008;582:2937-44.
372. Saunders AH, Griffiths AE, Lee KH, Cicchillo RM, Tu L, Stromberg JA, et al. Characterization of quinolinate synthases from *Escherichia coli*, *Mycobacterium tuberculosis*, and *Pyrococcus horikoshii* indicates that [4Fe-4S] clusters are common cofactors throughout this class of enzymes. Biochemistry. 2008;47:10999-1012.
373. Zou R, Zhou K, Stephanopoulos G, Too HP. Combinatorial engineering of 1-deoxy-D-xylulose 5-phosphate pathway using cross-lapping in vitro assembly (CLIVA) method. Plos One. 2013;8:e79557.
374. Carlsen S, Ajikumar PK, Formenti LR, Zhou K, Phon TH, Nielsen ML, et al. Heterologous expression and characterization of bacterial 2-C-methyl-D-erythritol-4-phosphate pathway in *Saccharomyces cerevisiae*. Appl Microbiol Biotechnol. 2013;97:5753-69.
375. Partow S, Siewers V, Daviet L, Schalk M, Nielsen J. Reconstruction and evaluation of the synthetic bacterial MEP pathway in *Saccharomyces cerevisiae*. Plos One. 2012;7:e52498.
376. Benisch F, Boles E. The bacterial Entner-Doudoroff pathway does not replace glycolysis in *Saccharomyces cerevisiae* due to the lack of activity of iron-sulfur cluster enzyme 6-phosphogluconate dehydratase. J Biotechnol. 2014;171:45-55.
377. Waks Z, Silver PA. Engineering a synthetic dual-organism system for hydrogen production. Appl Environ Microbiol. 2009;75:1867-75.
378. Kozak BU, van Rossum HM, Benjamin KR, Wu L, Daran JM, Pronk JT, et al. Replacement of the Saccharomyces cerevisiae acetyl-CoA synthetases by alternative pathways for cytosolic acetyl-CoA synthesis. Metab Eng. 2014;21:46-59.
379. M NM, Ollagnier-de-Choudens S, Sanakis Y, Abdel-Ghany SE, Rousset C, Ye H, et al. Characterization of *Arabidopsis thaliana* SufE2 and SufE3: functions in chloroplast iron-sulfur cluster assembly and NAD synthesis. J Biol Chem. 2007;282:18254-64.
380. Perli T, Moonen DPI, van den Broek M, Pronk JT, Daran JM. Adaptive laboratory evolution and reverse engineering of

single-vitamin prototrophies in Saccharomyces cerevisiae. Appl Environ Microbiol. 2020;86:e00388-20.
381. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421.
382. Solomon KV, Haitjema CH, Henske JK, Gilmore SP, Borges-Rivera D, Lipzen A, et al. Early-branching gut fungi possess a large, comprehensive array of biomass-degrading enzymes. Science. 2016;351:1192-5.
383. Dobin A, Gingeras TR. Mapping RNA-seq Reads with STAR. Curr Protoc Bioinformatics. 2015;51:11 4 1- 4 9.
384. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078-9.
385. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics. 2012;28:464-9.
386. Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 2013;41:e121.
387. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772-80.
388. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972-3.
389. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019;47:W256-W9.
390. Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics. 2011;12:124.
391. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.
392. Luttik MA, Kotter P, Salomons FA, van der Klei IJ, van Dijken JP, Pronk JT. The *Saccharomyces cerevisiae ICL2* gene encodes a mitochondrial 2-methylisocitrate lyase involved in propionyl-coenzyme A metabolism. J Bacteriol. 2000;182:7007-13.
393. Solis-Escalante D, Kuijpers NG, Bongaerts N, Bolat I, Bosman L, Pronk JT, et al. amdSYM, a new dominant recyclable marker cassette for Saccharomyces cerevisiae. FEMS Yeast Res. 2013;13:126-39.
394. Perli T, Moonen DPI, van den Broek M, Pronk JT, Daran JM. Adaptive laboratory evolution and reverse engineering of single-vitamin prototrophies in *Saccharomyces cerevisiae*. Appl Environ Microbiol. 2020;86.
395. Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods. 2009;6:343-5.
396. Papapetridis I, Goudriaan M, Vazquez Vitali M, de Keijzer NA, van den Broek M, van Maris AJA, et al. Optimizing anaerobic growth rate and fermentation kinetics in *Saccharomyces cerevisiae* strains expressing Calvin-cycle enzymes for improved ethanol yield. Biotechnol Biofuels. 2018;11:17.
397. Lee ME, DeLoache WC, Cervantes B, Dueber JE. A highly characterized yeast toolkit for modular, multipart assembly. Acs Synth Biol. 2015;4:975-86.
398. DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. Nucleic Acids Res. 2013;41:4336-43.
399. Medina VG, Almering MJH, van Maris AJA, Pronk JT. Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. Appl Environ Microb. 2010;76:190-5.
400. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754-60.
401. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. Plos One. 2014;9:e112963.
402. Alberts B. Molecular biology of the cell2018.
403. Champe PC, Harvey RA, Ferrier DR. Biochemistry: Lippincott Williams & Wilkins; 2005.
404. Li Y, Smolke CD. Engineering biosynthesis of the anticancer alkaloid noscapine in yeast. Nat Commun. 2016;7:12137.
405. Rajagopalan KV, Johnson JL. The pterin molybdenum cofactors. J Biol Chem. 1992;267:10199-202.
406. Rubio LM, Ludden PW. Biosynthesis of the iron-molybdenum cofactor of nitrogenase. Annu Rev Microbiol. 2008;62:93-

111.
407. Schwarz G, Mendel RR, Ribbe MW. Molybdenum cofactors, enzymes and pathways. Nature. 2009;460:839-47.
408. Mendel RR. The molybdenum cofactor. J Biol Chem. 2013;288:13165-72.
409. Iobbi-Nivol C, Leimkuhler S. Molybdenum enzymes, their maturation and molybdenum cofactor biosynthesis in *Escherichia coli*. Biochim Biophys Acta. 2013;1827:1086-101.
410. Zhang Y, Rump S, Gladyshev VN. Comparative genomics and evolution of Molybdenum utilization. Coord Chem Rev. 2011;255:1206-17.
411. Peng T, Xu Y, Zhang Y. Comparative genomics of molybdenum utilization in prokaryotes and eukaryotes. BMC Genomics. 2018;19:691.
412. Leimkuhler S, Buhning M, Beilschmidt L. Shared sulfur mobilization routes for tRNA thiolation and molybdenum cofactor biosynthesis in prokaryotes and eukaryotes. Biomolecules. 2017;7.
413. Hille R. Molybdenum and tungsten in biology. Trends Biochem Sci. 2002;27:360-7.
414. Hille R, Hall J, Basu P. The mononuclear molybdenum enzymes. Chem Rev. 2014;114:3963-4038.
415. Palmer T, Santini CL, Iobbi-Nivol C, Eaves DJ, Boxer DH, Giordano G. Involvement of the narJ and mob gene products in distinct steps in the biosynthesis of the molybdoenzyme nitrate reductase in *Escherichia coli*. Mol Microbiol. 1996;20:875-84.
416. Lake MW, Temple CA, Rajagopalan KV, Schindelin H. The crystal structure of the *Escherichia coli* MobA protein provides insight into molybdopterin guanine dinucleotide biosynthesis. J Biol Chem. 2000;275:40211-7.
417. Reschke S, Duffus BR, Schrapers P, Mebs S, Teutloff C, Dau H, et al. Identification of YdhV as the first molybdoenzyme binding a Bis-Mo-MPT cofactor in *Escherichia coli*. Biochemistry. 2019;58:2228-42.
418. Cvetkovic A, Menon AL, Thorgersen MP, Scott JW, Poole FL, 2nd, Jenney FE, Jr., et al. Microbial metalloproteomes are largely uncharacterized. Nature. 2010;466:779-82.
419. Nielsen J. Yeast systems biology: model organism and cell factory. Biotechnol J. 2019;14:e1800421.
420. Ostergaard S, Olsson L, Nielsen J. Metabolic engineering of *Saccharomyces cerevisiae*. Microbiol Mol Biol Rev. 2000;64:34-50.
421. de Barros Pita W, Leite FC, de Souza Liberal AT, Simoes DA, de Morais MA, Jr. The ability to use nitrate confers advantage to *Dekkera bruxellensis* over *S. cerevisiae* and can explain its adaptation to industrial fermentation processes. Antonie Van Leeuwenhoek. 2011;100:99-107.
422. da Silva TC, Leite FC, De Morais MA, Jr. Distribution of *Dekkera bruxellensis* in a sugarcane-based fuel ethanol fermentation plant. Lett Appl Microbiol. 2016;62:354-8.
423. Juergens H, Varela JA, de Vries ARG, Perli T, Gast VJM, Gyurchev NY, et al. Genome editing in *Kluyveromyces* and *Ogataea* yeasts using a broad-host-range Cas9/gRNA co-expression plasmid. Fems Yeast Res. 2018;18.
424. Winzeler EA, Richards DR, Conway AR, Goldstein AL, Kalman S, McCullough MJ, et al. Direct allelic variation scanning of the yeast genome. Science. 1998;281:1194-7.
425. Tejada-Jimenez M, Llamas A, Sanz-Luque E, Galvan A, Fernandez E. A high-affinity molybdate transporter in eukaryotes. Proc Natl Acad Sci U S A. 2007;104:20126-30.
426. Ravin NV, Eldarov MA, Kadnikov VV, Beletsky AV, Schneider J, Mardanova ES, et al. Genome sequence and analysis of methylotrophic yeast *Hansenula polymorpha* DL1. BMC Genomics. 2013;14:837.
427. Smith MT, Yamazaki M, Poot GA. *Dekkera*, *Brettanomyces* and *Eeniella*: electrophoretic comparison of enzymes and DNA-DNA homology. Yeast. 1990;6:299-310.
428. Solis-Escalante D, Kuijpers NG, Bongaerts N, Bolat I, Bosman L, Pronk JT, et al. amdSYM, a new dominant recyclable marker cassette for *Saccharomyces cerevisiae*. Fems Yeast Res. 2013;13:126-39.
429. Jespersen L, Jakobsen M. Specific spoilage organisms in breweries and laboratory media for their detection. Int J Food Microbiol. 1996;33:139-55.
430. Suh SO, Zhou JLJ. Methylotrophic yeasts near *Ogataea* (*Hansenula*) *polymorpha*: a proposal of *Ogataea angusta* comb. nov and *Candida parapolymorpha* sp nov. Fems Yeast Res. 2010;10:631-8.
431. Saraya R, Gidijala L, Veenhuis M, van der Klei IJ. Tools for genetic engineering of the yeast *Hansenula polymorpha*. Methods Mol Biol. 2014;1152:43-62.
432. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403-10.

433. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Res. 2019;47:W636-W41.
434. Gorter de Vries AR, de Groot PA, van den Broek M, Daran JG. CRISPR-Cas9 mediated gene deletions in lager yeast *Saccharomyces pastorianus*. Microb Cell Fact. 2017;16:222.
435. Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. Plos One. 2008;3:e3647.
436. Raymond CK, Pownder TA, Sexson SL. General method for plasmid construction using homologous recombination. Biotechniques. 1999;26:134-8, 40-1.
437. van Rossum HM, Kozak BU, Niemeijer MS, Duine HJ, Luttik MAH, Boer VM, et al. Alternative reactions at the interface of glycolysis and citric acid cycle in *Saccharomyces cerevisiae*. Fems Yeast Res. 2016;16.
438. Boonekamp FJ, Dashko S, van den Broek M, Gehrmann T, Daran JM, Daran-Lapujade P. The genetic makeup and expression of the glycolytic and fermentative pathways are highly conserved within the *Saccharomyces* genus. Front Genet. 2018;9:504.
439. Perli T, Vos AM, Bouwknegt J, Dekker WJC, Wiersma SJ, Mooiman C, et al. Identification of oxygen-independent pathways for pyridine-nucleotide and Coenzyme-A synthesis in anaerobic fungi by expression of candidate genes in yeast. bioRxiv. 2020:2020.07.06.189415.
440. Boonekamp FJ, Dashko S, Duiker D, Gehrmann T, van den Broek M, den Ridder M, et al. Design and experimental evaluation of a minimal, innocuous watermarking strategy to distinguish near-Identical DNA and RNA sequences. ACS Synth Biol. 2020;9:1361-75.
441. Llamas A, Mendel RR, Schwarz G. Synthesis of adenylated molybdopterin: an essential step for molybdenum insertion. J Biol Chem. 2004;279:55241-6.
442. Llamas A, Otte T, Multhaup G, Mendel RR, Schwarz G. The Mechanism of nucleotide-assisted molybdenum insertion into molybdopterin. A novel route toward metal cofactor assembly. J Biol Chem. 2006;281:18343-50.
443. Marelja Z, Stocklein W, Nimtz M, Leimkuhler S. A novel role for human Nfs1 in the cytoplasm: Nfs1 acts as a sulfur donor for MOCS3, a protein involved in molybdenum cofactor biosynthesis. J Biol Chem. 2008;283:25178-85.
444. Teschner J, Lachmann N, Schulze J, Geisler M, Selbach K, Santamaria-Araujo J, et al. A novel role for *Arabidopsis* mitochondrial ABC transporter ATM3 in molybdenum cofactor biosynthesis. Plant Cell. 2010;22:468-80.
445. Kispal G, Csere P, Prohl C, Lill R. The mitochondrial proteins Atm1p and Nfs1p are essential for biogenesis of cytosolic Fe/S proteins. Embo J. 1999;18:3981-9.
446. Marelja Z, Mullick Chowdhury M, Dosche C, Hille C, Baumann O, Lohmannsroben HG, et al. The L-cysteine desulfurase NFS1 is localized in the cytosol where it provides the sulfur for molybdenum cofactor biosynthesis in humans. Plos One. 2013;8:e60869.
447. Campbell WH. Nitrate reductase structure, function and regulation: bridging the gap between biochemistry and physiology. Annu Rev Plant Physiol Plant Mol Biol. 1999;50:277-303.
448. Guerrero MG, Vega JM, Losada M. The assimilatory nitrate-reducing system and its regulation. Annu Rev Plant Phys. 1981;32:169-204.
449. Zagorec M, Buhler JM, Treich I, Keng T, Guarente L, Labbe-Bois R. Isolation, sequence, and regulation by oxygen of the yeast *HEM13* gene coding for coproporphyrinogen oxidase. J Biol Chem. 1988;263:9718-24.
450. Camadro JM, Thome F, Brouillet N, Labbe P. Purification and properties of protoporphyrinogen oxidase from the yeast *Saccharomyces cerevisiae*. Mitochondrial location and evidence for a precursor form of the protein. J Biol Chem. 1994;269:32085-91.
451. Pena-Moreno IC, Castro Parente D, da Silva JM, Andrade Mendonca A, Rojas LAV, de Morais Junior MA, et al. Nitrate boosts anaerobic ethanol production in an acetate-dependent manner in the yeast *Dekkera bruxellensis*. J Ind Microbiol Biotechnol. 2019;46:209-20.
452. Protchenko O, Shakoury-Elizeh M, Keane P, Storey J, Androphy R, Philpott CC. Role of *PUG1* in inducible porphyrin and heme transport in *Saccharomyces cerevisiae*. Eukaryot Cell. 2008;7:859-71.
453. de Souza Liberal AT, Basilio AC, do Monte Resende A, Brasileiro BT, da Silva-Filho EA, de Morais JO, et al. Identification of *Dekkera bruxellensis* as a major contaminant yeast in continuous fuel ethanol fermentation. J Appl Microbiol. 2007;102:538-47.

454. Abbott DA, Hynes SH, Ingledew WM. Growth rates of *Dekkera*/*Brettanomyces* yeasts hinder their ability to compete with *Saccharomyces cerevisiae* in batch corn mash fermentations. Appl Microbiol Biotechnol. 2005;66:641-7.
455. Crauwels S, Van Assche A, de Jonge R, Borneman AR, Verreth C, Troels P, et al. Comparative phenomics and targeted use of genomics reveals variation in carbon and nitrogen assimilation among different *Brettanomyces bruxellensis* strains. Appl Microbiol Biot. 2015;99:9123-34.
456. Parente DC, Cajueiro DBB, Moreno ICP, Leite FCB, Pita WD, De Morais MA. On the catabolism of amino acids in the yeast *Dekkera bruxellensis* and the implications for industrial fermentation processes. Yeast. 2018;35:299-309.
457. Maia LB, Moura JJ, Moura I. Molybdenum and tungsten-dependent formate dehydrogenases. J Biol Inorg Chem. 2015;20:287-309.
458. Wierckx N, Koopman F, Ruijssenaars HJ, de Winde JH. Microbial degradation of furanic compounds: biochemistry, genetics, and impact. Appl Microbiol Biotechnol. 2011;92:1095-105.
459. Schaedler TA, Thornton JD, Kruse I, Schwarzlander M, Meyer AJ, van Veen HW, et al. A conserved mitochondrial ATP-binding cassette transporter exports glutathione polysulfide for cytosolic metal cofactor assembly. J Biol Chem. 2014;289:23264-74.
460. Leighton J, Schatz G. An ABC transporter in the mitochondrial inner membrane is required for normal growth of yeast. EMBO J. 1995;14:188-95.
461. Kispal G, Csere P, Guiard B, Lill R. The ABC transporter Atm1p is required for mitochondrial iron homeostasis. FEBS Lett. 1997;418:346-50.
462. Deves R, Boyd CA. The determination of kinetic parameters for carrier-mediated transport of non-labelled substrate analogues: a general method applied to the study of divalent anion transport in placental membrane vesicles. Proc R Soc Lond B Biol Sci. 1989;237:85-97.
463. Fitzpatrick KL, Tyerman SD, Kaiser BN. Molybdate transport through the plant sulfate transporter SHST1. FEBS Lett. 2008;582:1508-13.
464. Truong HN, Meyer C, Daniel-Vedele F. Characteristics of *Nicotiana tabacum* nitrate reductase protein produced in *Saccharomyces cerevisiae*. Biochem J. 1991;278 (Pt 2):393-7.
465. Siverio JM. Assimilation of nitrate by yeasts. FEMS Microbiol Rev. 2002;26:277-84.
466. de Barros Pita W, Tiukova I, Leite FC, Passoth V, Simoes DA, de Morais MA, Jr. The influence of nitrate on the physiology of the yeast *Dekkera bruxellensis* grown under oxygen limitation. Yeast. 2013;30:111-7.
467. Facklam TJ, Marzluf GA. Nitrogen regulation of amino acid catabolism in *Neurospora crassa*. Biochem Genet. 1978;16:343-54.
468. Costello C, Griffin WM, Landis AE, Matthews HS. Impact of biofuel crop production on the formation of hypoxia in the Gulf of Mexico. Environ Sci Technol. 2009;43:7985-91.
469. Bakker BM, Overkamp KM, van Maris AJ, Kotter P, Luttik MA, van Dijken JP, et al. Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. FEMS Microbiol Rev. 2001;25:15-37.
470. Galafassi S, Capusoni C, Moktaduzzaman M, Compagno C. Utilization of nitrate abolishes the "Custers effect" in *Dekkera bruxellensis* and determines a different pattern of fermentation products. J Ind Microbiol Biotechnol. 2013;40:297-303.
471. Blomqvist J, Nogue VS, Gorwa-Grauslund M, Passoth V. Physiological requirements for growth and competitiveness of *Dekkera bruxellensis* under oxygen-limited or anaerobic conditions. Yeast. 2012;29:265-74.
472. Tiukova IA, Moller-Hansen I, Belew ZM, Darbani B, Boles E, Nour-Eldin HH, et al. Identification and characterisation of two high-affinity glucose transporters from the spoilage yeast *Brettanomyces bruxellensis*. FEMS Microbiol Lett. 2019;366.
473. Barth G, Gaillardin C. Physiology and genetics of the dimorphic fungus *Yarrowia lipolytica*. FEMS Microbiol Rev. 1997;19:219-37.
474. Tsugawa R, Nakase T, Kobayash.T, Yamashit.K, Okumura S. Fermentation of *n*-paraffins by yeast .3. alpha-Ketoglutarate productivity of various yeast. Agricultural and Biological Chemistry. 1969;33:929-&.
475. Marella ER, Dahlin J, Dam MI, ter Horst J, Christensen HB, Sudarsan S, et al. A single-host fermentation process for the production of flavor lactones from non-hydroxylated fatty acids. Metab Eng. 2020;61:427-36.
476. Holkenbrink C, Ding BJ, Wang HL, Dam MI, Petkevicius K, Kildegaard KR, et al. Production of moth sex pheromones for pest control by yeast fermentation. Metab Eng. 2020;62:312-21.

477. Rumelhard M, Hosako H, Eurlings IM, Westerink WM, Staska LM, van de Wiel JA, et al. Safety evaluation of rebaudioside A produced by fermentation. Food Chem Toxicol. 2016;89:73-84.
478. Perli T, van der Vorm DNA, Wassink M, van den Broek M, Pronk JT, Daran J-M. Engineering heterologous molybdenum-cofactor-biosynthesis and nitrate-assimilation pathways enables nitrate utilization by *Saccharomyces cerevisiae*. Metab Eng. 2021;65:11-29.
479. Fischer K, Llamas A, Tejada-Jimenez M, Schrader N, Kuper J, Ataya FS, et al. Function and structure of the molybdenum cofactor carrier protein from *Chlamydomonas reinhardtii*. J Biol Chem. 2006;281:30186-94.
480. Chen DC, Beckerich JM, Gaillardin C. One-step transformation of the dimorphic yeast *Yarrowia lipolytica*. Appl Microbiol Biot. 1997;48:232-5.
481. Bitinaite J, Rubino M, Varma KH, Schildkraut I, Vaisvila R, Vaiskunaite R. USER friendly DNA engineering and cloning method by uracil excision. Nucleic Acids Res. 2007;35:1992-2002.
482. Holkenbrink C, Dam MI, Kildegaard KR, Beder J, Dahlin J, Domenech Belda D, et al. EasyCloneYALI: CRISPR/Cas9-based synthetic toolbox for engineering of the yeast *Yarrowia lipolytica*. Biotechnol J. 2018;13:e1700543.
483. Andersen JB, Sternberg C, Poulsen LK, Bjorn SP, Givskov M, Molin S. New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Appl Environ Microbiol. 1998;64:2240-6.
484. Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, et al. Genome evolution in yeasts. Nature. 2004;430:35-44.
485. Kerscher S, Durstewitz G, Casaregola S, Gaillardin C, Brandt U. The complete mitochondrial genome of *Yarrowia lipolytica*. Comp Funct Genomics. 2001;2:80-90.
486. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics. 2011;27:2156-8.
487. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly. 2012;6:80-92.
488. Almagro Armenteros JJ, Salvatore M, Emanuelsson O, Winther O, von Heijne G, Elofsson A, et al. Detecting sequence signals in targeting peptides using deep learning. Life Sci Alliance. 2019;2.
489. Almagro Armenteros JJ, Sonderby CK, Sonderby SK, Nielsen H, Winther O. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics. 2017;33:3387-95.
490. Fukasawa Y, Tsuji J, Fu SC, Tomii K, Horton P, Imai K. MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites. Mol Cell Proteomics. 2015;14:1113-26.
491. Small I, Peeters N, Legeai F, Lurin C. Predotar: A tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics. 2004;4:1581-90.
492. Petsalaki EI, Bagos PG, Litou ZI, Hamodrakas SJ. PredSL: a tool for the N-terminal sequence-based prediction of protein subcellular localization. Genomics Proteomics Bioinformatics. 2006;4:48-55.
493. Avram D, Bakalinsky AT. *SSU1* encodes a plasma membrane protein with a central role in a network of proteins conferring sulfite tolerance in *Saccharomyces cerevisiae*. J Bacteriol. 1997;179:5971-4.
494. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, et al. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science. 2014;343:1247997.
495. Vahisalu T, Kollist H, Wang YF, Nishimura N, Chan WY, Valerio G, et al. SLAC1 is required for plant guard cell S-type anion channel function in stomatal signalling. Nature. 2008;452:487-91.
496. Cabrera E, Gonzalez-Montelongo R, Giraldez T, Alvarez de la Rosa D, Siverio JM. Molecular components of nitrate and nitrite efflux in yeast. Eukaryot Cell. 2014;13:267-78.
497. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell. 2014;156:935-49.
498. Fonfara I, Le Rhun A, Chylinski K, Makarova KS, Lecrivain AL, Bzdrenga J, et al. Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9 among orthologous type II CRISPR-Cas systems. Nucleic Acids Res. 2014;42:2577-90.
499. Lobs AK, Schwartz C, Wheeldon I. Genome and metabolic engineering in non-conventional yeasts: Current advances and applications. Synth Syst Biotechnol. 2017;2:198-207.

500. Cai P, Gao J, Zhou Y. CRISPR-mediated genome editing in non-conventional yeasts for biotechnological applications. Microb Cell Fact. 2019;18:63.
501. Jensen ED, Ferreira R, Jakociunas T, Arsovska D, Zhang J, Ding L, et al. Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. Microb Cell Fact. 2017;16:46.
502. Deaner M, Holzman A, Alper HS. Modular Ligation Extension of Guide RNA Operons (LEGO) for Multiplexed dCas9 Regulation of Metabolic Pathways in Saccharomyces cerevisiae. Biotechnol J. 2018;13:e1700582.
503. Dominguez AA, Lim WA, Qi LS. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. Nat Rev Mol Cell Biol. 2016;17:5-15.
504. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. Nat Rev Genet. 2016;17:333-51.
505. Jain M, Olsen HE, Paten B, Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biol. 2016;17:239.
506. Jain M, Tyson JR, Loose M, Ip CLC, Eccles DA, O'Grady J, et al. MinION Analysis and Reference Consortium: Phase 2 data release and analysis of R9.0 chemistry. F1000Res. 2017;6:760.
507. Rang FJ, Kloosterman WP, de Ridder J. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. Genome Biol. 2018;19:90.
508. Bainomugisa A, Duarte T, Lavu E, Pandey S, Coulter C, Marais BJ, et al. A complete high-quality MinION nanopore assembly of an extensively drug-resistant *Mycobacterium tuberculosis* Beijing lineage strain identifies novel variation in repetitive PE/PPE gene regions. Microb Genom. 2018;4.
509. Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, et al. Highly parallel direct RNA sequencing on an array of nanopores. Nat Methods. 2018;15:201-6.
510. Howorka S, Siwy ZS. Reading amino acids in a nanopore. Nat Biotechnol. 2020;38:159-60.
511. Restrepo-Perez L, Joo C, Dekker C. Paving the way to single-molecule protein sequencing. Nat Nanotechnol. 2018;13:786-96.
512. Palluk S, Arlow DH, de Rond T, Barthel S, Kang JS, Bector R, et al. *De novo* DNA synthesis using polymerase-nucleotide conjugates. Nat Biotechnol. 2018;36:645-50.
513. Jensen MA, Davis RW. Template-independent enzymatic oligonucleotide synthesis (TiEOS): its history, prospects, and challenges. Biochemistry. 2018;57:1821-32.
514. Hughes RA, Ellington AD. Synthetic DNA synthesis and assembly: putting the synthetic in synthetic biology. Cold Spring Harb Perspect Biol. 2017;9.
515. Ranaghan MJ, Li JJ, Laprise DM, Garvie CW. Assessing optimal: inequalities in codon optimization algorithms. BMC Biol. 2021;19:36.
516. Martens JH, Barg H, Warren MJ, Jahn D. Microbial production of vitamin B12. Appl Microbiol Biotechnol. 2002;58:275-85.
517. Eisenstein M. Enzymatic DNA synthesis enters new phase. Nat Biotechnol. 2020;38:1113-5.
518. Hillson N, Caddick M, Cai Y, Carrasco JA, Chang MW, Curach NC, et al. Building a global alliance of biofoundries. Nat Commun. 2019;10:2040.
519. May M. A DIY approach to automating your lab. Nature. 2019;569:587-8.
520. Barthels F, Barthels U, Schwickert M, Schirmeister T. FINDUS: An Open-Source 3D Printable Liquid-Handling Workstation for Laboratory Automation in Life Sciences. SLAS Technol. 2020;25:190-9.
521. Lesnick J, Lejeune-Dodge A, Ruppert N, Jarman C. High-precision cell dispensing with the Labcyte Echo liquid handler. Sunnyvale, CA: Labcyte, Aug. 2016.
522. He R, Ding R, Heyman JA, Zhang D, Tu R. Ultra-high-throughput picoliter-droplet microfluidics screening of the industrial cellulase-producing filamentous fungus *Trichoderma reesei*. J Ind Microbiol Biotechnol. 2019;46:1603-10.
523. Huang M, Bai Y, Sjostrom SL, Hallstrom BM, Liu Z, Petranovic D, et al. Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. Proc Natl Acad Sci U S A. 2015;112:E4689-96.
524. Wehrs M, Tanjore D, Eng T, Lievense J, Pray TR, Mukhopadhyay A. Engineering robust production microbes for large-scale cultivation. Trends Microbiol. 2019;27:524-37.

525. Rugbjerg P, Myling-Petersen N, Porse A, Sarup-Lytzen K, Sommer MOA. Diverse genetic error modes constrain large-scale bio-based production. Nat Commun. 2018;9:787.
526. Junker BH. Scale-up methodologies for *Escherichia coli* and yeast fermentation processes. J Biosci Bioeng. 2004;97:347-64.
527. Kensy F, Zang E, Faulhammer C, Tan RK, Buchs J. Validation of a high-throughput fermentation system based on online monitoring of biomass and fluorescence in continuously shaken microtiter plates. Microb Cell Fact. 2009;8:31.
528. Wong BG, Mancuso CP, Kiriakov S, Bashor CJ, Khalil AS. Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER. Nat Biotechnol. 2018;36:614-23.
529. D'Ambrosio V, Jensen MK. Lighting up yeast cell factories by transcription factor-based biosensors. FEMS Yeast Res. 2017;17.
530. Ozbal CC, LaMarr WA, Linton JR, Green DF, Katz A, Morrison TB, et al. High throughput screening via mass spectrometry: a case study using acetylcholinesterase. Assay Drug Dev Technol. 2004;2:373-81.
531. Sinclair I, Stearns R, Pringle S, Wingfield J, Datwani S, Hall E, et al. Novel acoustic loading of a mass spectrometer: toward next-generation high-throughput MS screening. J Lab Autom. 2016;21:19-26.
532. St John PC, Bomble YJ. Approaches to computational strain design in the multiomics era. Front Microbiol. 2019;10:597.
533. Noe F, De Fabritiis G, Clementi C. Machine learning for protein folding and dynamics. Curr Opin Struct Biol. 2020;60:77-84.
534. Lai HY, Zhang ZY, Su ZD, Su W, Ding H, Chen W, et al. iProEP: a computational predictor for predicting promoter. Mol Ther-Nucl Acids. 2019;17:337-46.
535. Tang H, Wu Y, Deng J, Chen N, Zheng Z, Wei Y, et al. Promoter architecture and promoter engineering in *Saccharomyces cerevisiae*. Metabolites. 2020;10.
536. Lynd LR. Overview and evaluation of fuel ethanol from cellulosic biomass: technology, economics, the environment, and policy. Annu Rev Energ Env. 1996;21:403-65.
537. Chen GQ. New challenges and opportunities for industrial biotechnology. Microb Cell Fact. 2012;11:111.

Acknowledgements

During the ups and lows of my 4-years long PhD journey, and more in general, during my scientific career, I had the pleasure to meet a bunch of smart and pleasant folks who helped me shape into who I am today. Although it is impossible for me to herewith address each one of them, I would like to thank a few in particular.

First and foremost, I want to express my sincere gratitude to **Jean-Marc** for his daily supervision and guidance. I really appreciate the freedom and independence you have provided me over the years, while also never failing to toss out fresh ideas and offer possible solutions to experimental challenges. Your practical and laid-back approach made it very easy for me to get through any situation, even the most difficult ones. I feel very proud that you considered me for the position within the HTS team. That means a lot to me and I really look forward to continuing working with you. **Jack**, thanks for being such an inspiring co-promotor. The positive and energetic way you engage in every scientific discussion really amazes me and the fact that you genuinely care about the personal development of each one of us in the group is something that I really value and respect.

During my Ph.D., I was able to build on the knowledge and expertise that I had obtained from previous supervisors. Therefore, I would like to thank you all for helping me grow my interests and develop as a scientist. **Cristina and Paola**, you provided me a first look at the world of synthetic biology. I had a lot of fun working with you and the team during that wild summer packed with experiments and take-away pizzas. The joyous moment during the iGEM award ceremony was the icing on the cake, and it amply repaid all the efforts; I will never forget that. **Sheref**, working in your lab for my master's thesis was a great pleasure. I appreciate all of the flexibility you offered me during the research; it truly helped me grow as an independent scientist and not be afraid to step outside of my comfort zone. Also, thank you for supporting all of my subsequent applications; I am certain that your letters were extremely helpful. **Michael**, thanks for teaching me the dark art of microfluidics back at CFEL. It was really fun to learn all those hetching and moulding techniques that fall more in the material engineering rather than biology. I would have not been able to accomplish much during in my project without your help. Thank you, **Pamela**, **Steph**, and **Isaac**, for having me at such a terrific lab at HMS. Despite the fact that it was one of my first trips overseas, you assisted me in settling in and provided me with the opportunity to learn about a variety of cutting-edge synthetic

biology projects, which boosted my creativity. **Massimo**, I remember the first day we met in Boston as if it was yesterday. You really opened my mind and introduced me to the world of industrial research. Thank you for believing in me and for giving me the chance to join your team at Ginkgo Bioworks. Under your guidance, I really got passionate about microbial biotechology and I am really glad that you advised me to start a PhD. If you are reading this book is also because of you! The year I spent at Ginkgo was one of the most formative times in my life. Working at such an innovative company was mind-blowing and I would like to thank all the extremely talented Ginks I had the chance to work or simply share a drink with.

Throughout my PhD I had the great pleasure to supervise six students and, although I was initially skeptical, this was one of the activities I enjoyed the most. **Mats**, **Dewi**, **Mees**, **Ellen**, **Daan** and **Lilian**: I want to thank you for choosing to work with me and for all the energy and dedication with which you worked during your projects. Being the first time for me in such role, I have to say that I really learned a lot from all of you and you made me realize that this is something I would like to do more in the future. I hope you enjoyed your time at IMB as much I enjoyed my time with you! I wish you all the best for your future career steps.

Completing a PhD is never a one-person task and during my time at IMB I had the pleasure to collaborate with many other people in the group and I would like to thank them here. **Anna**, my PAcMEN fellow, colleague and friend, I am really glad I got the chance to share all our travel adventures and also collaborate on different projects with you. Your positive energy and determination will be missed. Thank you for all the fun moments and fruitful discussions, work and non-work related. **Wijb**, it took some time but we eventually got to know each other and I am super glad we did! I have immense respect for you both as a person and as a scientist. I really enjoyed working with you in the lab and although we could not take the project through the finish line, I had a great time and learned a lot in the process. Also, **Anna** and **Wijb** thank you for accepting to be my paranymphs. You are among the people who know me best and I cannot think of anyone else who would be better suited for this role. **Aurin**, as you joined the group it was like a breath of fresh air. You were always able to cheer everyone up with some jokes or absurd fun discussions. Collaborating on the PaNic project with you and the ElOxy team was lots of fun. **Raul**, I am really glad that someone with your skills joined the group. Without you, it would have been impossible for me to make sense of all those protein sequences! You are a super chill guy and I wish you all the best. **Jasmijn**, although we did not share any scientific project, my PhD would have not been the same without you around. All the coffee breaks, discussions, shared successes and failures made me feel like I was not alone

in all this. Thank you.

Something that can be easily overlooked when talking about an academic groups but of crucial importance in keeping things running smoothly and in providing long-term knowledge transfer, is the presence of a great staff technicians team. **Pilar**, **Marijke**, and **Erik**, and **Marcel**, thank you for keeping the lab afloat! **Pilar**, thank you for regularly reminding me to register my strains' information in the database and for never getting mad at me for forgetting to do so! You are super friendly and easy to approach, I was really sad when you decided to leave the group but, I am sure you made the right choice and I wish you all the best back in Spain! **Marijke**, thank you for all your help during my enzyme assay experiments and thank you for genuinely care about the well being of all of us. I think many of us would agree if I say that you are the mom of the group! **Erik**, thank you for helping me setting up fermentation experiments and for always trying your best to keep the fermentation lab tidy and organized. Without you it would just be a mass of tangled tubings! **Christiaan** and **Flip**, thank you for taking the time to answer every stupid question I had about fermentation processes. Before starting at IMB, I set a goal for myself to learn more about fermentation and microbial physiology during my PhD. Well, you all played a huge role in helping me achieving that goal, thank you! **Marcel**, thank you for patiently supporting me and explaining me every step in the genomics and transcriptomics data analysis. I learned a lot from you and I really think you are an essential asset for the group.

Performing reproducible microbiology experiments in sterile conditions requires a lot of time investment to prepare media, sterile glassware and reactors. I felt really lucky when I heard I could just ask **Apilena**, **Astrid** and **Jannie** to help me with this, your work at the MSD kitchen is invaluable to IMB. Thank you for always kindly accepting my last-minute requests and for always being so flexible and available.

Since it would be impossible to thank each one of you individually here, I want you to know that I am extremely thankfull to everyone who contributed in making **IMB** such a pleasant work environment. I always felt like I was part of a big family and I am sincerely going to miss many of you.

Cari **mamma** e **papà** vorrei ringraziarvi per avermi incoraggiato a studiare e per il vostro grande supporto in tutte le mie scelte. Non dev'essere stato facile sopportarmi durante la mia adolescenza e poi osservarmi andar via all'estero, prima negli Stati Uniti, poi nei Paesi Bassi. Spero che leggere questa dedica e vedere fin dove sono arrivato vi renda un po' orgogliosi e ripaghi in parte lo sforzo. Non c'è giorno che passi che non vi penso.

Grazie **Matteo** e **Michela** per mettermi sempre un sorriso sulla bocca ogni volta che ci chiamiamo e cerchiamo di comunicare con quelle due bestioline. Grazie per prendervi cura di mamma e papà e delle nonne **Luigia** e **Rosina**. Non vedo l'ora di potervi ospitare tutti quanti qui a Delft per farvi vivere un po' della mia quotidianità. Un grazie speciale va anche alla mia famiglia acquisita **Antonio**, **Marisa** e **Simone**. Grazie per farmi sempre sentire come uno dei vostri e per preoccuparvi sempre per noi.

Finally, I'd like to express my gratitude to **Elisa**, my wife. Thank you so much for always encouraging and motivating me. I am extremely proud of who you have become. A brilliant scientist as well as a warm and compassionate person. I'm excited to embark on many new adventures with you.

Curriculum vitae

Thomas Perli was born on the 12th of March 1990 in Trento, Italy. Thomas grew up in, Zambana, a small town near Trento and completed his high-school diploma at Liceo Scientifico Leonardo da Vinci, Trento in 2010. Soon after, Thomas enrolled in the Scienze e Technologie Biomolecolari bachelor programme at the University of Trento and performed his bachelor thesis as part of the UNITN-2013 iGEM team, engineering a bacteria for the production of ethylene, that was used to boost fruit ripening. In 2013, Thomas started a master in Cellular and Molecular Biotechnology at the University of Trento and performed a 9-months research project in the Reconstructive and Synthetic Biology group headed by Prof. Sheref Mansy, studying *in vitro* protein expression systems and cellular communication using microfluidic devices. During that time, Thomas was awarded with an Erasmus+ scholarship that supported a three-months traineeship at the Coherent Imaging Division, CFEL (Hamburg, Germany), under the supervision of Dr. Michael Heymann, where he learned different microfabrication and lithographich techniques. During his master, Thomas was selected for the national Giovanni Armenise-Harward Foundation summer fellowship programme and spent two months in Prof. Pamela Silver laboratory at Harvard Medical School (Boston, USA) where he designed and tested synthetic inducible promoters for cyanobacteria. After completing his studies, Thomas performed a 1-year internship at Ginkgo Bioworks (Boston, USA) where he worked on improving performances of a microbial production strain though metabolic engineering under the supervision of Dr. Massimo Merighi. In 2017, Thomas moved to the Netherlands and started a PhD, as part of the MSCA PAcMEN ITN programme, at the Industrial Microbiology group, Delft University of Technology, under the co-supervision of Dr. Jean-Marc Daran and Prof. Jack Pronk. During his PhD, Thomas supervised six student research projects, co-authored various scientific publications, and was inventor on a patent application. The results generated during Thomas's PhD project form the basis of this thesis. In March 2021, Thomas accepted a postdoc position at the Industrial Microbiology Group.

List of publications

Digiacomo, F., Girelli, G., Aor, B., Marchioretti, C., Pedrotti, M., **Perli, T.**, Tonon, E., Valentini, V., Avi, D., Ferrentino, G., Dorigato, A., Torre, P., Jousson, O., Mansy, S. S., Del Bianco, C., 2014. **Ethylene-producing bacteria that ripen fruit.** *ACS Synth Biol*. 3.

Dassi, E., Ferretti, P., Covello, G., Specchер, A., Migazzi, A., Bosco, B., Rajashekar, B., Zarbo, C., Ballabio, C., Rossetto, D., Maino, E., Corradi, E., Costa, F., Precazzini, F., Rashid, H. O., Nicolussi, M., Bolzan, M., Demozzi, M., Olivieri, M., Zordan, N., Pedron, R., Manara, S., Rezvan, S., Naik, S. N., Khaghani, S., Masella, S., **Perli, T.**, Pierini, V., Bertorelli, R., Denti, M. A., De Sanctis, V., Tett, A., Segata, N., Htm, C. M. B., 2018. **The short-term impact of probiotic consumption on the oral cavity microbiome.** *Sci Rep*. 8.

Juergens, H., Varela, J. A., Gorter de Vries, A. R., **Perli, T.**, Gast, V. J. M., Gyurchev, N. Y., Rajkumar, A. S., Mans, R., Pronk, J. T., Morrissey, J. P., Daran, J. G., 2018. **Genome editing in *Kluyveromyces* and *Ogataea* yeasts using a broad-host-range Cas9/gRNA co-expression plasmid.** *FEMS Yeast Res.* 18.

Perli, T.*, Wronska, A. K.*, Ortiz-Merino, R. A., Pronk, J. T., Daran, J. M., 2020. **Vitamin requirements and biosynthesis in *Saccharomyces cerevisiae*.** *Yeast*. 37.

Perli, T., Moonen, D. P. I., van den Broek, M., Pronk, J. T., Daran, J. M., 2020. **Adaptive laboratory evolution and reverse engineering of single-vitamin prototrophies in *Saccharomyces cerevisiae*.** *Appl Environ Microbiol.* 86.

Perli, T., van der Vorm, D. N. A., Wassink, M., van den Broek, M., Pronk, J. T., Daran, J.-M., 2021. **Engineering heterologous molybdenum-cofactor-biosynthesis and nitrate-assimilation pathways enables nitrate utilization by *Saccharomyces cerevisiae*.** *Metab Eng*. 65.

Wronska, A. K., van den Broek, M., **Perli, T.**, de Hulster, E. A. F., Pronk, J. T., Daran, J.-M., 2021. **Engineering oxygen-independent biotin biosynthesis in *Saccharomyces cerevisiae*.** *Metab Eng.* 67.

Perli, T.*, Vos, A. M.*, Bouwknegt, J., Dekker, W. J. C., Wiersma, S. J., Mooiman, C., Ortiz-Merino, R. A., Daran, J.-M., Pronk, J. T., 2021. **Identification of oxygen-independent pathways for pyridine-nucleotide and Coenzyme-A synthesis in anaerobic fungi by expression of candidate genes in yeast.** *mBio*. 12.

Perli, T., Borodina I., Daran, J.-M., 2021 **Engineering of molybdenum-cofactor-dependent nitrate assimilation in *Yarrowia lipolytica*.** *Manuscript submitted*

* These authors contributed equally.

List of patents

Pronk, J. T., **Perli, T.**, & Daran, J-M. G., 2020. **Yeast with engineered molybdenum co-factor biosynthesis.** WO2020209718, IPC No. C07K, C12N.