RESEARCH ARTICLE

Engineering of molybdenum-cofactor-dependent nitrate assimilation in Yarrowia lipolytica

Thomas Perli¹,†, Irina Borodina²,‡ and Jean-Marc Daran¹,*,§

¹Department of Biotechnology, Delft University of Technology, Van der Maasweg 9, 2629 HZ, Delft, The Netherlands and ²The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kemitorvet, Building 220, 2800 Kgs. Lyngby, Denmark

*Corresponding author: Department of Biotechnology, Delft University of Technology, Van der Maasweg 9, 2629 HZ, Delft, The Netherlands. Tel: +31-15-27-82412; E-mail: J.G.Daran@TUDelft.nl

One sentence summary: This study establishes Molybdenum cofactor biosynthesis in Yarrowia lipolytica to expand the nutrient range thereof.

Editor: Zongbao Zhao

†Thomas Perli, http://orcid.org/0000-0002-4910-3292
‡Irina Borodina, http://orcid.org/0000-0002-8452-1393
§Jean-Marc Daran, http://orcid.org/0000-0003-3136-8193

ABSTRACT

Engineering a new metabolic function in a microbial host can be limited by the availability of the relevant cofactor. 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 Ogataea parapolymorpha Moco biosynthesis pathway combined with the expression of a high affinity molybdate transporter could lead to the synthesis of Moco in Y. lipolytica. The functionality of Moco was demonstrated by expression of an active Moco-dependent nitrate assimilation pathway from the same yeast 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, reaching 0.13 h⁻¹. 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 nutrient range used by industrial yeasts.

Keywords: molybdenum cofactor; nitrate reductase; nitrate assimilation; Yarrowia lipolytica; metabolic engineering

INTRODUCTION

As Saccharomyces cerevisiae, the Dipodascaceae yeast Yarrowia lipolytica has gained interest as a model yeast for dimorphism studies and as an industrial work horse (Barth and Gaillardin 1997). 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 and α-ketoglutaric acids (Tsugawa et al. 1969).

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 specialty chemicals. Y. lipolytica has become
a reliable platform for metabolic engineering as illustrated by the synthesis of flavor compounds (Marella et al. 2020), insect sex pheromones (Holkenbrink et al. 2020), non-caloric sweeteners (Rumelhard et al. 2016), itaconic acid (Blazeck et al. 2015) and terpenoids (Arnesen et al. 2020) as a few examples.

Engineering of Y. lipolytica with increasingly more complex pathways 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 (Broderick 2001; Champe, Harvey and Ferrier 2005). 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 or too expensive, (2) is 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 Ni²⁺-transport, an inorganic cofactor of Ni-dependent urease (Mölne et al. 2015), or with the engineering of tetrahydrobipterin pathway, that was instrumental in the implementation of de novo biosynthesis of opioids (Galanie et al. 2015; Li and Smolke 2016) and melatonin (Germann et al. 2016). More recently, the molybdopterin cofactor biosynthesis pathway from the methylo-trophic 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 (Perli et al. 2021). 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. Enabling Moco biosynthesis and nitrate assimilation in a evolutionarily distant yeast such as Y. lipolytica would confirm that the selected gene-set may be used in wide range of yeast species but would also contribute to further expand the metabolic capabilities of this microbial cell factory. Previous research has shown how Y. lipolytica accumulates lipids in nitrogen starving conditions (Ratledge 2002). Introduction of a heterologous nitrogen assimilation pathway such as the nitrate assimilation pathways could be applied to mimic a nitrogen-starving condition without actually having to limit the amount of nitrogen in the media.

As S. cerevisiae and more than 75% of the Saccharomycotina yeast species, Y. lipolytica is not able to synthesize molybdopterin cofactor (Shen et al. 2018) which precludes the possibility to harness enzyme families of biotechnological relevance. Moco-dependent enzymes encompass more than 30 different catalytic activities divided in three main families, based on the Moco variant they require (Cvetkovic et al. 2010; Hille, Hall and Basu 2014). Moco can be directly used as cofactor or modified by either sulfuration of the molybdate ion to form sulfated Moco or, in prokaryotes, by the covalent attachment of either GDP or CDP to molybdopterin (MTP) to form Bis(Molybdopterin Guanine Dinucleotide) Molybdenum Cofactor or MTP-cytosine dinucleotide cofactor, respectively. In a large majority, Moco-dependent enzymes catalyse oxidoreduction 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 (Hille 2002).

Moco is composed of a molybdate \(\text{MoO}_4^{2-}\) oxianion coordinated by two sulfur atoms on MTP, a tricyclic pterin scaffold. Moco cannot be supplemented in the media since the molecule is too unstable due to its oxygen sensitivity (Fischer et al. 2006). 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 (Rajagopalan and Johnson 1992; Iobbi-Nivol and Leimkuhler 2013; Mendel 2013). The first step, which takes place in the mitochondria of eukaryotic cells, is catalysed by the heterodimer Cnx1/Cnx2, that cyclizes GTP onto 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/Cnx6) 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 adenytransferase Cnx4 and a cysteine desulfurase that is typically involved in iron–sulfur cluster biosynthesis and tRNA thiolation, Nfs1 (Leimkuhler, Buhning and Beilschmidt 2017). In the final step, molybdate is inserted in a two-steps reaction catalyzed by the multi-domain protein Cnx3 (Iobbi-Nivol and Leimkuhler 2013). Previous studies showed that the yeast S. cerevisiae lacks a high-affinity molybdate transport system and that it is able to import the oxianion only with low affinity, unless a high affinity transporter such as CrMot1 from Chlamydomonas reinhardtii is expressed (Tejada-Jimenez et al. 2007; Perli et al. 2021).

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, a total of 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 (ALE) in synthetic media with nitrate as the sole nitrogen source. 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 ku70A::Spycas9-EcDsdAMX4; Marella et al. 2020). Yeast strains were grown on either YP (10 g/L Bacto yeast extract and 20 g/L Bacto peptone) or SM medium (Verduyn et al. 1992) with either 5 g/L KNO₃, or 2.3 g/L urea (SMNO₃ and SMurea, respectively) as sole nitrogen source. In all SM media variants, 6.6 g/L K₂SO₄ was added as a source of sulfate (Solis-Escalante et al. 2013). 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 et al. 1992). 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 SMYD 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...
**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 and        | Marella et al. (2020) |
|            | pGPD-OpCNX2::LIP2                                                                | ATTCC-20460             |                    |
| IMX2264    | MATa ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2 E4Δ::pTEFin-OpCNX1-tPEX10  | IMX2264                 | This study         |
|            | pGPD-OpCNX2::LIP2                                                                |                        |                    |
| IMX2265    | MATa ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2 E4Δ::pTEFin-OpCNX1-tPEX10  | IMX2265                 | This study         |
|            | pGPD-OpCNX2::LIP2                                                                |                        |                    |
| IMX2266    | MATa ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2 E4Δ::pTEFin-OpCNX1-tPEX10  | IMX2265                 | This study         |
|            | pGDP-OpCNX3::LIP2                                                                |                        |                    |
| IMX2267    | MATa ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2 E4Δ::pTEFin-OpCNX1-tPEX10  | IMX2266                 | This study         |
|            | pGDP-OpCNX2::LIP2                                                                |                        |                    |
| IMX2565    | MATa ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2 E4Δ::pTEFin-OpCNX1-tPEX10  | IMX2267                 | This study         |
|            | pGPD-OpCNX2::LIP2                                                                |                        |                    |
| IMX1174    | MATa ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2 E4Δ::pTEFin-OpCNX1-tPEX10  | IMX2565                 | This study         |
|            | pGPD-OpCNX2::LIP2                                                                |                        |                    |
| IMX1175    | MATa ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2 E4Δ::pTEFin-OpCNX1-tPEX10  | IMX2565                 | This study         |
|            | pGPD-OpCNX2::LIP2                                                                |                        |                    |
| IMX1176    | MATa ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2 E4Δ::pTEFin-OpCNX1-tPEX10  | IMX2565                 | This study         |
|            | pGPD-OpCNX2::LIP2                                                                |                        |                    |
| IMX1177    | MATa ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2 E4Δ::pTEFin-OpCNX1-tPEX10  | IMX2565                 | This study         |
|            | pGPD-OpCNX2::LIP2                                                                |                        |                    |
| IMX1178    | MATa ku70Δ::pTEF1-Spcas9-tTEF12::pGPD-EcdsdAMX4-tLIP2 E4Δ::pTEFin-OpCNX1-tPEX10  | IMX2565                 | This study         |
|            | pGPD-OpCNX2::LIP2                                                                |                        |                    |

(Evolved on SMDNO3 for 21 transfers. Line 2–Colony 1)

(Evolved on SMDNO3 for 21 transfers. Line 2–Colony 2)

(Evolved on SMDNO3 for 21 transfers. Line 2–Colony 3)

(Evolved on SMDNO3 for 21 transfers. Line 3–Colony 1)

(Evolved on SMDNO3 for 21 transfers. Line 3–Colony 2)
| Name       | Relevant genotype                                                                 | Parental strain | Reference |
|------------|-----------------------------------------------------------------------------------|-----------------|-----------|
| IMS1179    | MATa::pTEF1::Spca9-tTEF1::pGPD-EcdsdAMX4-tLIP2 E4A::pTEF1::OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2 C2A::pTEF1::OpNFS1-tPEX10 pGPD-OpCNX3-tLIP2 E1A::pTEF1::OpCNX3-tPEX10 pGPD-CrMoT1-tLIP2 C3A::pTEF1::OpYNR1-tPEX10 pGPD-OpYN1-tLIP2 pTEF1::OpCNX6-tPEX10 | IMX2565         | This study |
| IMS1180    | MATa::pTEF1::Spca9-tTEF1::pGPD-EcdsdAMX4-tLIP2 E4A::pTEF1::OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2 C2A::pTEF1::OpNFS1-tPEX10 pGPD-OpCNX3-tLIP2 E1A::pTEF1::OpCNX3-tPEX10 pGPD-CrMoT1-tLIP2 C3A::pTEF1::OpYNR1-tPEX10 pGPD-OpYN1-tLIP2 pTEF1::OpCNX6-tPEX10 | IMX2565         | This study |
| IMS1181    | MATa::pTEF1::Spca9-tTEF1::pGPD-EcdsdAMX4-tLIP2 E4A::pTEF1::OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2 C2A::pTEF1::OpNFS1-tPEX10 pGPD-OpCNX3-tLIP2 E1A::pTEF1::OpCNX3-tPEX10 pGPD-CrMoT1-tLIP2 C3A::pTEF1::OpYNR1-tPEX10 pGPD-OpYN1-tLIP2 pTEF1::OpCNX6-tPEX10 | IMX2565         | This study |
| IMS1182    | MATa::pTEF1::Spca9-tTEF1::pGPD-EcdsdAMX4-tLIP2 E4A::pTEF1::OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2 C2A::pTEF1::OpNFS1-tPEX10 pGPD-OpCNX3-tLIP2 E1A::pTEF1::OpCNX3-tPEX10 pGPD-CrMoT1-tLIP2 C3A::pTEF1::OpYNR1-tPEX10 pGPD-OpYN1-tLIP2 pTEF1::OpCNX6-tPEX10 | IMX2565         | This study |
| IMS1183    | MATa::pTEF1::Spca9-tTEF1::pGPD-EcdsdAMX4-tLIP2 E4A::pTEF1::OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2 C2A::pTEF1::OpNFS1-tPEX10 pGPD-OpCNX3-tLIP2 E1A::pTEF1::OpCNX3-tPEX10 pGPD-CrMoT1-tLIP2 C3A::pTEF1::OpYNR1-tPEX10 pGPD-OpYN1-tLIP2 pTEF1::OpCNX6-tPEX10 | IMX2565         | This study |
| IMS1184    | MATa::pTEF1::Spca9-tTEF1::pGPD-EcdsdAMX4-tLIP2 E4A::pTEF1::OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2 C2A::pTEF1::OpNFS1-tPEX10 pGPD-OpCNX3-tLIP2 E1A::pTEF1::OpCNX3-tPEX10 pGPD-CrMoT1-tLIP2 C3A::pTEF1::OpYNR1-tPEX10 pGPD-OpYN1-tLIP2 pTEF1::OpCNX6-tPEX10 | IMX2565         | This study |
| IMS1185    | MATa::pTEF1::Spca9-tTEF1::pGPD-EcdsdAMX4-tLIP2 E4A::pTEF1::OpCNX1-tPEX10 pGPD-OpCNX2-tLIP2 C2A::pTEF1::OpNFS1-tPEX10 pGPD-OpCNX3-tLIP2 E1A::pTEF1::OpCNX3-tPEX10 pGPD-CrMoT1-tLIP2 C3A::pTEF1::OpYNR1-tPEX10 pGPD-OpYN1-tLIP2 pTEF1::OpCNX6-tPEX10 | IMX2565         | This study |
Bacto tryptone, 5 g/L Bacto yeast extract and 5 g/L NaCl supplemented with 100 mg/L ampicillin. Yarrowia 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, St-Louis, MO) 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 and 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. Escherichia coli DHS-a (New England BioLabs, Ipswich, MA) was used for cloning procedures (Inoue, Nojima and Okayama 1990). Yeast genomic DNA used for diagnostic PCR reactions was isolated by using the SDS/LiAc protocol (Looke, Kristjuhan and Kristjuhan 2011). Yarrowia lipolytica transformation was performed with the LiAc method as previously described (Chen, Beckerich and Gaillardin 1997). A total of 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.

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 (OpYN1, OpYN1R and OpYN1I) were retrieved from O. parapolymorpha DL-1 genome sequence (Ravin et al. 2013; Perli et al. 2021) BioProject PRJNA60503). A gene coding for a previously characterized high-affinity molybdenum transporter, CrMoT1, from C. reinhardtii was also included in the gene-set (Tejada-Jimenez et al. 2007). 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 plasmids pUD1057 (OpCNX6), pUD1058 (OpCNX4), pUD1059 (OpCNX5), pUD1060 (OpNFS1), pUD1061 (OpCNX1), pUD1062 (OpCNX2), pUD1063 (CrMoT1) pUD1064 (OpCNX3), pUD1065 (OpYN1R), pUD1066 (OpYN1I) and pUD1067 (OpYN1I). Single-gene Biobricks compatible with USER cloning (New England BioLabs) were amplified from pUD1057, pUD1058, pUD1059, pUD1060, pUD1061, pUD1062, pUD1063, pUD1064, pUD1065, pUD1066 and pUD1067 using primer pairs 24487/24488, 24488/24486, 24489/24490, 24493/24448, 24479/24480, 24481/24482, 24491/24492, 24493/24494, 24495/24496, 24499/24500 and 24497/24498, respectively. Single promoter Biobricks pTEFin and pGPD were amplified from ST6512 genomic DNA using primer pairs 22956/24013 and 15529/15528, respectively. Then, a Biobrick carrying the back-to-back promoter pair pTEFin-pGPD was cloned by USER cloning (Bitinaite et al. 2007) fusion of the two single-promoter Biobricks and primer pair 22956/15528. A single promoter pTEFin 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 (Holkenbrink et al. 2018). A backbone E_4 Biobrick was combined with Biobricks pTEFin-pGPD, OpCNX1 and OpCNX2 in a USER cloning reactions as previously described (Holkenbrink et al. 2018) to yield plasmid CFB8966. Similarly, plasmids pCfB8896, pCfB8896, pCfB8896, and pCfB8897 were cloned by combining a pTEFin-pGPD promoter Biobrick with a C_2, E_1, E_3 or C_3 backbone, respectively, and gene Biobricks OpNFS1/OPCNX4, OpCNX6/OpCNX5, CrMoT1/OpCNX2 and OpYN1R/OpYN1I, respectively. Plasmid pCfB8897 was cloned by combining a D_1 backbone Biobrick with a pTEFin promoter and OpYN1I gene Biobricks. Gibson assembly (Gibson et al. 2009) was used to construct the integration plasmid pCfB8906 that carries expression 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 pCfB8896 as template. Secondly, the GFPmut3 spacer cassette (Andersen et al. 1998) was amplified from plasmid pY774 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 pCfB8907, carrying nitrate assimilation pathway was cloned by combining a OpYN1T1-OpYN1R1 cassette with an OpYN1I cassette and a GFPmut3 spacer. Plasmid pCfB8970 was linearized with primers 24524/24517 while the OpYN1I cassette was amplified with primers 24525/24518 and pCfB8897 as template. Plasmid pUD1264 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 pCfB8898 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 (Holkenbrink et al. 2018). 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 stock at −80°C. First, strain ST6512 was transformed with NotI-digested pCfB8896 and the E_4 targeting gRNA expression plasmid pCfB6638 to yield IMX2264. Then, IMX2264 was transformed with NotI-digested pCfB8896 and the C_2 targeting gRNA expression plasmid pCfB6682 to yield IMX2265. IMX2266 was obtained.
Table 2. Primers used in this study.

| Primer number | Primer sequence | Product(s)          |
|---------------|-----------------|---------------------|
| 22956         | AGTACTGCAAAGAUGCTTGGC   | PrTEFin-PrPD_USER_Biobrick |
| 24013         | ATCAAGACUGAGACCCGCTGGCG  | PrTEFin-PrPD_USER_Biobrick |
| 15529         | AGCTACTGAAAGCAGCTGACCTGCAG  | PrTEFin-PrPD_USER_Biobrick |
| 15528         | ATGACAGAUGTTTAGTGGTGGTGGTTATTCAAGAATG  | PrTEFin-PrPD_USER_Biobrick |
| 27208         | ACAGCAGCAGACCGAGGTTGGCGG  | PrTEFin_USER_Biobrick |
| 22956         | AGTACTGCAAAGAUGCTTGGC   | PrTEFin_USER_Biobrick |
| 24479         | ACTTTTTGCACTACAAACCAGAGCTACGGATCGAATTGGGAC  | OpCNX1_USER_Biobrick |
| 24480         | GTGCAATTAGTGGGAAATGAAGGTC  | OpCNX1_USER_Biobrick |
| 24481         | ATCGGCAAATTAGTGAGGCTATCCGAG  | OpCNX2_USER_Biobrick |
| 24482         | ACACGCAATTAGTGAGGCTATCCGAG  | OpCNX2_USER_Biobrick |
| 24483         | ACTTTTTGCACTACAAACCAGAGCTACGGATCGAATTGGGAC  | OpCNX3_USER_Biobrick |
| 24484         | CGTGGCAATTAGTGGGAAATGAAGGTC  | OpCNX3_USER_Biobrick |
| 24485         |ATCGGCAAATTAGTGAGGCTATCCGAG  | OpCNX4_USER_Biobrick |
| 24486         | CGTGGCAATTAGTGGGAAATGAAGGTC  | OpCNX4_USER_Biobrick |
| 24487         | ATCGGCAAATTAGTGAGGCTATCCGAG  | OpCNX5_USER_Biobrick |
| 24488         | CGTGGCAATTAGTGGGAAATGAAGGTC  | OpCNX5_USER_Biobrick |
| 24489         | ATCGGCAAATTAGTGAGGCTATCCGAG  | OpCNX5_USER_Biobrick |
| 24490         | CAGCCGAAATGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24491         | ACTTTTTGCACTACAAACCAGAGCTACGGATCGAATTGGGAC  | CrMoT1_USER_Biobrick |
| 24492         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24493         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24494         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24495         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24496         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24497         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24498         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24499         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24500         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24501         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24502         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24503         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24504         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24505         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24506         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24507         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24508         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24509         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24510         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24511         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24512         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24513         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24514         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24515         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24516         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24517         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24518         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24519         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24520         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24521         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24522         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24523         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24524         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 24525         | ATCGGCAAATTAGTGAGGCTATCCGAG  | CrMoT1_USER_Biobrick |
| 17887         | TCAGTCTGCGACCCAAGAATGTGGGAGTGAATCTACAGAAGGAT  | pUDI264_insert_Gibbon_fragment |
| 17888         | ATCGGCAAATTAGTGAGGCTATCCGAG  | pUDI264_insert_Gibbon_fragment |
| 17889         | ATCGGCAAATTAGTGAGGCTATCCGAG  | pUDI264_insert_Gibbon_fragment |
| 17890         | ATCGGCAAATTAGTGAGGCTATCCGAG  | pUDI264_insert_Gibbon_fragment |
Table 3. Plasmids used in this study.

| Name          | Characteristics                                      | Reference       |
|---------------|------------------------------------------------------|-----------------|
| pCB6371       | bla ColE1 NotIC, 3-3 homology tPEX20-tLIP2 C3-5 homology NotIC | Holkenbrink et al. (2018) |
| pCB6677       | bla ColE1 NotIE, 1-3 homology tPEX20-tLIP2 E1-5 homology NotIE | Holkenbrink et al. (2018) |
| pCB6679       | bla ColE1 NotED, 4-3 homology tPEX20-tLIP2 E4-5 homology NotED | Holkenbrink et al. (2018) |
| pCB6681       | bla ColE1 NotEB, 3-3 homology tPEX20-tLIP2 E3-5 homology NotEB | Holkenbrink et al. (2018) |
| pCB6682       | bla ColE1 NotEG, 2-3 homology tPEX20-tLIP2 E2-5 homology NotEG | Holkenbrink et al. (2018) |
| pCB6684       | bla ColE1 NotD, 3-3 homology tPEX20-tLIP2 D1-5 homology NotD | Holkenbrink et al. (2018) |
| pCB6627       | bla ColE1 NAT mRNA, C2                                    | Holkenbrink et al. (2018) |
| pCB6630       | bla ColE1 NAT mRNA, C3                                    | Holkenbrink et al. (2018) |
| pCB6631       | bla ColE1 NAT mRNA, D1                                    | Holkenbrink et al. (2018) |
| pCB6633       | bla ColE1 NAT mRNA, E1                                    | Holkenbrink et al. (2018) |
| pCB6637       | bla ColE1 NAT mRNA, E3                                    | Holkenbrink et al. (2018) |
| pCB6638       | bla ColE1 NAT mRNA, E4                                    | Holkenbrink et al. (2018) |
| pI774         | bla ColE1 Gfpmut3                                        | GeneArt         |
| 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 OpYNR11                                       | GeneArt         |
| pUD1067       | bla ColE1 OpYNR11                                       | GeneArt         |
| pCB8966       | bla ColE1 NotIE, 4-3 homology tPEX20-OpCNX1-TEF1in-pGPD-OpCNX2-tLIP2 | This study |
| pCB8967       | bla ColE1 NotIE, 2-3 homology tPEX20-OpNFS1-TEF1in-pGPD-OpCNX4-tLIP2 | This study |
| pCB8968       | bla ColE1 NotIE, 1-3 homology tPEX20-OpCNX6-TEF1in-pGPD-OpCNX5-tLIP2 | This study |
| pCB8969       | bla ColE1 NotIE, 3-3 homology tPEX20-CrMoT1-TEF1in-pGPD-OpCNX3-tLIP2 | This study |
| pCB8970       | bla ColE1 NotIE, 3-3 homology tPEX20-OpYNR1-TEF1in-pGPD-OpYNR11-tLIP2 | This study |
| pCB8971       | bla ColE1 NotIE, 2-3 homology tPEX20-OpYNR11-TEF1in-tLIP2 C3-5 homology NotIE | This study |
| pCB9006       | bla ColE1 NotIE, 3-3 homology tPEX20-OpCNX6-TEF1in-pGPD-OpCNX5-Gfpmut3-tLIP2 Gfpmut3 | This study |
| pCB9007       | bla ColE1 NotIE, 3-3 homology tPEX20-OpYNR1-TEF1in-pGPD-OpYNR11-tLIP2 Gfpmut3-tLIP2 Gfpmut3 | This study |
| pUD1264       | bla ColE1 NotIE, 3-3 homology tPEX20-OpCNX6-TEF1in-pGPD-OpCNX5-tLIP2 E3-5 homology NotIE | This study |

*Codon optimized for expression in Y. lipolytica.

by transforming IMX2265 with NotI-digested pCB9006 and the E1 targeting gRNA expression plasmid pCB6633. IMX2264 was transformed with NotI-digested pCB9007 and the C3 targeting gRNA expression plasmid pCB6630 to yield IMX2267. Whole genome re-sequencing of IMX2267 revealed that OpCNX5 and OpCNX5 were not integrated but instead CrMoT1 and OpCNX3 were present in two copies (PRJNA704845). To correct this absence, IMX2267 was subsequently transformed with NotI-digested pUD1264 together with the E3 targeting gRNA expression plasmid pCB6637, to yield the final strain IMX2565. The presence of all integrated genes in IMX2565 was confirmed by whole genome re-sequencing (PRJNA704845).

ALE

To evolve IMX2565 for fast growth in nitrate-containing medium, the strain was inoculated in triplicate in 100-mL flasks containing 20 mL SMDNO3. Flasks were incubated at 30 C, 200 rpm until OD600 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 was 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 SMDNO3. 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, colony 2), 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 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-μL shake flask containing 20 mL SMDNO3 and incubated overnight at 30°C, 200 rpm. The next day, each culture was centrifuged at 3000 g for 5 min, supernatant was discarded and cell pellets were resuspended in SMDNO3 to and OD660 of 5. Then, two 96 deep-well plates were filled with 250 μL of SMDNO3 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 OD660 equivalent, a calibration curve was prepared by correlating the ‘green’ value of a IMS2565 cultures in SMD100 at eight different OD660 values that were measured externally with a 7200 Jenway Spectrophotometer (Jenway, Stone, UK). 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 SMD100 of OD660 = 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, OD660 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 OD660 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, IMS1180, IMS1181, IMS1182, IMX1183, IMS1184 and IMS1185, frozen stock cultures were used to inoculate 20 mL SMDNO3 starter cultures. These were subsequently used to inoculate 100 mL SMDNO3 flask cultures to initial OD660 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 min at 14 000 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 and Cell Culture DNA Kit with 100G Genomics-tips (QIAGEN, Hilden, Germany) following manufacturer’s instructions. Illumina-based paired-end sequencing with 150-bp reads was performed on 550-bp TrueSeq 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 (Kerscher et al. 2001; Dujon et al. 2004) to which five extra contigs containing the relevant integration cassettes had been previously added.

Data processing and chromosome copy number variation determinations were done as previously described (Nijkamp et al. 2012; Perli et al. 2020) except for VCF file intersection and annotation that was performed with VCFtools 0.1.13 (vcf-iseq command) and Snpeff, respectively (Danecek et al. 2011; Cingolani et al. 2012).

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 (Perli et al. 2021). 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 flasks cultures on the same medium, to an initial OD660 of 0.2. Shake flasks were incubated until the OD660 exceeded 40. All strains were grown on SMDurea with the exception of the evolution lines 1, 2 and 3 that were grown on SMDNO3. 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 KPO4, 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 15 000 g on a tabletop centrifuge. Supernatant was recovered in a fresh tube and clarified by centrifuging for 1 h at 4°C, 20 000 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-dependant NADPH consumption at 340 nm using a spectrophotometer (Jasco, Easton, MA). Reactions were performed in 1 mL final volume of 100 mM KPO4 buffer pH 7 supplemented with 200 μM NADPH, 20 μM FAD, 1 mM KNO3 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 mgproteins−1.

In silico mitochondrial targeting prediction

The likelihood of mitochondrial targeting for the wild-type and mutated OpCnx1 and OpCnx2 sequences were calculated using five different web-based tools: TargetP 2.0 (Almagro Armenteros et al. 2019), DeepLoc 1.0 (Almagro Armenteros et al. 2017), MitoFates (Fukasawa et al. 2015), Predotar (Small et al. 2004) and PredSL (Petsalaki et al. 2006). 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, 0.5 g/L H2SO4 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 (Verhoeven et al. 2017). Nitrate, nitrite and ammonium concentrations in 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 (https://researchdata.4tu.nl/) under URL: https://doi.org/10.4121/14230238.v1. DNA sequencing data from Y. 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 Y. lipolytica

The absence of molybdenum-dependent enzymes in Yarrowia metabolism strongly suggested that the engineering of Moco biosynthesis genes may not only require functional expression of Moco biosynthesis genes, but also of a high-affinity molybdate transporter (Fig. 1). Based on a previous work in S. cerevisiae (Perli et al. 2021), it is not fewer than 11 genes from the yeast O. parapolymorpha and the algae C. 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), three genes coding for the nitrate assimilation pathway (OpYNT1, OpYNR1 and OpYNI1) from O. parapolymorpha, and one gene coding for the high-affinity molybdate transporter (CrMoT1) from C. reinhardtii (Fig. 1). The codon-optimized genes were integrated in the chromosome of the Yarrowia strain ST6512 (W29, ku70Δ::pTEF1-Cas9-tTEF12::pGPD-DsaA-tLIF2) by using CRISPR/Cas9 gene-editing and the EASY-Yali promoter parts and integrative plasmids (Holkenbrink et al. 2018). 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 (Holkenbrink et al. 2018). At each transformation, two genes were integrated, with the exception of the second last transformation in which all three genes encoding for nitrate assimilation pathway were 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, Fig. 2).

ALE 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 medium (Perli et al. 2021). We inoculated strain IMX2565 in SMDNO3 and incubated the flasks at 30°C in triplicate. Reproducibly, after 2 weeks the Yarrowia strain showed full growth (OD660 above 20) in all flasks. To improve the strain growth rate, IMX2565 was subjected to ALE by sequential transfers in flasks containing SMDNO3 for 335 generations (50 consecutive batches). Evolving cultures were periodically stocked and used to monitor the growth rate of the evolving populations (Fig. 3A). At the onset of the ALE experiment, the initial batches exhibited growth rates ranging 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 about 140 generations, three single colonies were isolated from each evolving culture and named IMS1180 (line 1, colony 1), IMS1181 (line 1, colony 2), IMS1182 (line 1, colony 3), IMS1174 (line 2, colony 1), IMS1175 (line 2, colony 2), IMS1176 (line 2, colony 3), IMS1177 (line 3, colony 1), IMS1178 (line 3, colony 2) and IMS1179 (line 3, colony 3). The specific growth rates of the single colony isolates on SMDNO3 ranged from 0.03 to 0.09 h⁻¹ (Fig. 3B). Clones IMS1180, IMS1175 and IMS1177 were empirically selected for further characterization based on the similar growth performance to the relative evolved populations. 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 (Fig. 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-read technology. After aligning the reads to the Y. lipolytica W29 reference genome (PRJNA604125; Kerscher et al. 2001; Dujon et al. 2004), 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 (Perli et al. 2021), 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 miscalcs 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 miscalcs, a total of five, two and four mutations were found in 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 miscalcs 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 miscalcs, a total of five, two and four mutations were found in evolved isolates IMS1175, IMS1177 and IMS1180, respectively (Fig. 3C and Table 4). While no genes were found mutated in all three independently evolved isolates, three genes (Spcas9, OpCNX1 and YALI0E24167g) 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 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 (YALI0E24167g) mutated in two different isolates IMS1180 and IMS1177 encoded a putative sulfite transporter that shared similarity to the S. cerevisiae sulfite efflux pump Ssu1 (Avram and Bakalinsky 1997). The nature of the mutations found in YALI0E24167g would suggest a loss of function. In IMS1177 the mutation G743A resulted in the introduction of a stop codon at position 248. Consequently, the mutated protein would be truncated to 45% of the original sequence and thus likely not functional.

The last gene found mutated in two different isolates was Spcas9. Although different, these two mutations occurred in a section of the gene encoding amino acids located in the same functional domain. The Spcas9 mutation in IMS1175 (G2959A) and in IMS1180 (C3020A) led to amino acids change in the RuvC-III domain of SpCas9 (Jinek et al. 2014).

Whereas IMS1175 and IMS1177 both harbored 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 YALI0D25784g encoding a hypothetical protein, YALI0A04697g a gene encoding a putative serine/threonine protein kinase, YALI0E00638g a gene encoding a putative methyl citrate synthase in IMS1175 and YALI0F05346g a gene encoding a protein exhibiting similarity with a cutinase from Fusarium solani cutinase.
ALE for fast growth on nitrate was associated with an increased Moco-dependent nitrate reductase activity

To investigate the effects of the ALE experiment, NADPH- and 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 SMDNO3, the parental un-evolved engineered strain IMX2565 and the Cas9-expressing parental strain ST6512 (Fig. 4). While as expected, cell extract from ST6512 which does not carry any of the heterologous genes for Moco biosynthesis and nitrate assimilation, showed no significant nitrate reductase activity, the engineered strain IMX2565 showed an activity of 0.004 ± 0.001 μM NADP⁺/min/mg total protein. 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/mg total protein. 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/mg total protein.

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 (SMDNO₃). The strains were grown in aerobic shake flasks and OD₆₆₀, glucose, nitrate, nitrite and ammonium levels were monitored over time (Fig. 5). As expected, the parental strain ST6512 did not show any growth and/or glucose/nitrate consumption (Fig. 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-well format (Fig. 3B), the ones measured in 100 mL shake flasks followed the same trend with IMS1180 being the slowest and IMS1175 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 (Fig. 5A). Notably, both evolved isolates and further evolved populations excreted moderate amounts of the intermediates nitrite and ammonia during the exponential phase of growth, suggesting that nitrate assimilation was not growth-limiting (Fig. SC–H). Similarly to what happened to engineered nitrate-assimilating S. cerevisiae strains (Perli et al. 2021), ammonia and nitrite were also excreted toward the end of the batch fermentation, resulting from continuous conversion or possible cell lysis.

**DISCUSSION**

Moco-dependent enzymes catalyse many redox reactions involved in the carbon, nitrogen and sulfur cycles (Leimkuhler and Iobbi-Nivol 2016) 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 host species would also suggest that this pathway engineering could be extrapolated to an even larger set of budding yeasts (Shen et al. 2018). 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 (Perli et al. 2021). 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 *Y. 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 OpCnx1 and OpCnx2 that catalyse the formation of cPMP in the very first step of Moco biosynthesis (Fig. 1). Evolved isolates IMS1175 and IMS1177 carried a mutation in the OpCnx1 whereas the third isolate IMS1180 carried a mutation in OpCnx2. 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 (Small et al. 2004; Petsalaki et al. 2006; Fukasawa et al. 2015; Almagro Armenteros et al. 2017, 2019; Table 5). These mutations could contribute to improve OpCnx1 and 2 translocations into the organelle and optimize 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 (Fig. 4). These results seem to indicate that although the *O. para-polyomorpha* 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
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 SMDNO3.

| 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) |
|              |               |             |                   | weakly similar to uniprot| O74782 Schizosaccharomyces pombe |
|              |               |             |                   | Hypothetical protein |
| IMS1177      | YALI0_D25784g | SNV         | T457C             | Ser151Pro |
|              | YALI0_E00638g | SNV         | G868T             | Gly290Cys |
|              |               |             |                   | |
|              | OpCNX1        | SNV         | G22A              | Glu8Lys |
|              | Spc9          | SNV         | G2959A            | Ala987Thr |
| IMS1180      | YALI0_E24167g | SNV         | G743A             | Trp248Stop |
|              |               |             |                   | Weakly similar to uniprot| Q9TEM3 Emericella nidulans MC3 |
|              |               |             |                   | Methylcitrate synthase precursor |
|              |               |             |                   | |
|              | OpCNX1        | SNV         | G22A              | Glu8Lys |
|              | IMS1180       | SNV         | A440G             | His147Arg |
|              | YALI0_E24167g | SNV         | G337A             | Ala113Thr |
|              |               |             |                   | |
|              | OpCNX2        | SNV         | G4A               | Val2Met |
|              | Spc9          | SNV         | C3020A            | Ala1007Asp |

Table 5. Likelihood of mitochondrial targeting in WT and mutated OpCnx1-OpCnx2 protein sequences across different prediction tools.

| Prediction tool | OpCnx1 WT | Mutated | OpCnx2 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  | 0.097147| 0.004088  | 0.004091|

characteristic of voltage-dependent anion transporters such as nitrate and sulfite exporters (Vahisalu et al. 2008). The protein is similar to the S. cerevisiae sulfite transporter Ssu1 which has been previously shown to be able to also export nitrate and nitrite (Cabrera et al. 2014). 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 by 180 amino acids, representing 45% of the protein length, that included not fewer than four transmembrane domains that undoubtedly caused
a loss of function. Assuming that YAL10_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 assimilation and growth rate after 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 Spcas9. 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 (Nishimasu et al. 2014) and all yielded either a decrease or an alteration of the endonuclease activity resulting in mutants able to only cleave one strand instead of two (Fonfara et al. 2014; Jinek et al. 2014). This was confirmed experimentally, as our attempts to further engineer 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 stepping stone 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.
Figure 5. Growth rates (A) and growth curves of ST6512 (B), IMS1180 (C), evolution line 1 T50 (D), IMS1175 (E), evolution line 2 T50 (F), IMS1177 (G) and evolution line 3 T50 (H) on SMDNO3. Symbols indicate biomass (•), glucose (♦), nitrate (◦), nitrite (□) and ammonium (△). 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).
SUPPLEMENTARY DATA

Supplementary data are available at FEMSYR online.

FUNDING

This work was supported by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska–Curie action PAcMEN (Grant agreement number 722287). IB acknowledges the financial support from the Novo Nordisk Foundation (Grant agreement numbers NN20OC0060809 and NN20OC0035580), and from the European Research Council under the European Union’s Horizon 2020 research and innovation program (YEAST-TRANS project, grant agreement number 757384).

AUTHOR CONTRIBUTIONS

TP and J-M.G.D. are inventors on a patent application related to this work (WO2020209718–Yeast with engineered molybdenum cofactor biosynthesis).

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 Professor Dr. Jack T. Prank for his support and contribution in the scientific discussion. The remaining authors have no competing interests to declare.

Conflicts of Interest. T.P. and J.-M.G.D. are inventors on a patent related to this work (WO2020209718–Yeast with engineered molybdenum cofactor biosynthesis).

REFERENCES

Almagro Armenteros JJ, Salvatore M, Emanuelsson O et al. Detecting sequence signals in targeting peptides using deep learning. Life Sci Allian 2019;2:e201900429.

Almagro Armenteros JJ, Sonderby CK, Sonderby SK et al. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics 2017;33:3387–95.

Andersen JB, Sternberg C, Poulsen LK et al. New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Appl Environ Microbiol 1998;64:2240–6.

Arnesen JA, Kildeggaard KR, Cernuda Pastor M et al. Yarrowia lipolytica strains engineered for the production of terpenoids. Front Bioeng Biotechnol 2020;8:945.

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.

Barth G, Gaillardin C. Physiology and genetics of the dimorphic fungus Yarrowia lipolytica. FEMS Microbiol Rev 1997;19:219–37.

Bitinaite J, Rubino M, Varma KH et al. USER friendly DNA engineering and cloning method by uracil excision. Nucleic Acids Res 2007;35:1992–2002.

Blazeck J, Hill A, Jamoussi M et al. Metabolic engineering of Yarrowia lipolytica for itaconic acid production. Metab Eng 2015;32:66–73.

Broderick JB. Coenzymes and Cofactors. Chichester: John Wiley & Sons, Ltd, 2001.

REFERENCES
Li Y, Smolke CD. Engineering biosynthesis of the anticancer alkaloid noscapine in yeast. Nat Commun 2016;7:12137.

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

Marella ER, Dahlin J, Dam MI et al. A single-host fermentation process for the production of flavor lactones from non-hydroxylated fatty acids. Metab Eng 2020;61:427–36.

Mendel RR. The molybdenum cofactor. J Biol Chem 2013;288:13165–72.

Milne N, Luttik MAH, Cueto Rojas HF et al. Functional expression of a heterologous nickel-dependent, ATP-independent urease in Saccharomyces cerevisiae. Metab Eng 2015;30:130–40.

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

Nishimasu H, Ran FA, Hsu PD et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell 2014;156:935–49.

Perli T, van der Vorm DA, Wassink M et al. Engineering heterologous molybdenum-cofactor-biosynthesis and nitrate-assimilation pathways enables nitrate utilization by Saccharomyces cerevisiae. Metab Eng 2021;65:11–29.

Perli T, Vos AM, Bouwknecht J 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. DOI: 10.1128/mBio.00967-21.

Petsalaki EI, Bagos PG, Litou ZI et al. PredSL: a tool for the N-terminal sequence-based prediction of protein subcellular localization. Genomics Proteomics Bioinformatics 2006;4:48–55.

Rajagopalan KV, Johnson JL. The pterin molybdenum cofactors. J Biol Chem 1992;267:10199–202.

Ratledge C. Regulation of lipid accumulation in oleaginous micro-organisms. Biochem Soc Trans 2002;30:1047–50.

Ravin NV, Eldarov MA, Kadnikov VV et al. Genome sequence and analysis of methylotrophic yeast Hansenula polymorpha DL1. BMC Genomics 2013;14:837.

Rumelhard M, Hosako H, Eurlings IM et al. Safety evaluation of rebaudioside A produced by fermentation. Food Chem Toxicol 2016;89:73–84.

Shen XX, Opulente DA, Kominek J et al. Tempo and mode of genome evolution in the budding yeast subphylum. Cell 2018;175:1533–45.

Small I, Peeters N, Legeai F et al. Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. Proteomics 2004;4:1581–90.

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

Tejada-Jimenez M, Llamas A, Sanz-Luque E et al. A high-affinity molybdate transporter in eukaryotes. Proc Natl Acad Sci 2007;104:20126–30.

Tsugawa R, Nakase T, Kobayash T et al. Fermentation of n-paraffins by yeast. 3. alpha-Ketoglutarate productivity of various yeast. Agr Biol Chem Tokyo 1969;33:929.

Vahisalu T, Kollist H, Wang YF et al. SLAC1 is required for plant guard cell S-type anion channel function in stomatal signalling. Nature 2008;452:487–91.

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

Verhoeven MD, Lee M, Kamoen L et al. Mutations in PMR1 stimulate xylose isomerase activity and anaerobic growth on xylose of engineered Saccharomyces cerevisiae by influencing manganese homeostasis. Sci Rep-Uk 2017;7:46155.