FEMS Yeast Research, 21, 2021, foab059

https://doi.org/10.1093/femsyr/foab059
Advance Access Publication Date: 25 November 2021
Minireview

MINIREVIEW

What makes *Komagataella phaffii* non-conventional?

Özge Ata[1,2], Burcu Gündüz Ergün[3,4,†], Patrick Fickers[5], Lina Heistinger[1,2,6,‡], Diethard Mattanovich[1,2,§], Corinna Rebnegger[1,2,7] and Brigitte Gasser[1,2,4,*,¶]

[1]Department of Biotechnology, Institute of Microbiology and Microbial Biotechnology, University of Natural Resources and Life Sciences Vienna (BOKU), Muthgasse 18, 1190 Vienna, Austria, [2]Austrian Centre of Industrial Biotechnology (ACIB), Muthgasse 11, 1190 Vienna, Austria, [3]UNAM-National Nanotechnology Research Center, Bilkent University, Ankara, Turkey, [4]Biotechnology Research Center, Ministry of Agriculture and Forestry, Ankara, Turkey, [5]Microbial Processes and Interactions, TERRA Teaching and Research Centre, Gembloux Agro-Bio Tech, University of Liège, Av. de la Faculté 2B, 5030 Gembloux, Belgium, [6]Christian Doppler Laboratory for Innovative Immunotherapeutics, University of Natural Resources and Life Sciences (BOKU), Muthgasse 18, 1190 Vienna, Austria and [7]Christian Doppler Laboratory for Growth-Decoupled Protein Production in Yeast, University of Natural Resources and Life Sciences Vienna (BOKU), Muthgasse 18, 1190 Vienna, Austria

*Corresponding author: Department of Biotechnology, Institute of Microbiology and Microbial Biotechnology, University of Natural Resources and Life Sciences Vienna (BOKU); Muthgasse 18, 1190 Vienna, Austria. E-mail: brigitte.gasser@boku.ac.at

One sentence summary: Non-conventional features render the yeast *K. phaffii* an attractive model organism and an efficient host for biotechnology applications.

Editor: R Ledesma-Amaro

†Burcu Gündüz Ergün, https://orcid.org/0000-0002-9344-1336
‡Lina Heistinger, https://orcid.org/0000-0002-0935-4652
§Diethard Mattanovich, https://orcid.org/0000-0002-0907-4167
¶Brigitte Gasser, https://orcid.org/0000-0003-2881-6370

ABSTRACT

The important industrial protein production host *Komagataella phaffii* (syn *Pichia pastoris*) is classified as a non-conventional yeast. But what exactly makes *K. phaffii* non-conventional? In this review, we set out to address the main differences to the 'conventional' yeast *Saccharomyces cerevisiae*, but also pinpoint differences to other non-conventional yeasts used in biotechnology. Apart from its methylotrophic lifestyle, *K. phaffii* is a Crabtree-negative yeast species. But even within the methylotrophs, *K. phaffii* possesses distinct regulatory features such as glycerol-repression of the methanol-utilization pathway or the lack of nitrate assimilation. Rewiring of the transcriptional networks regulating carbon (and nitrogen) source utilization clearly contributes to our understanding of genetic events occurring during evolution of yeast species. The mechanisms of mating-type switching and the triggers of morphogenic phenotypes represent further examples for how *K. phaffii* is distinguished from the model yeast *S. cerevisiae*. With respect to heterologous protein production, *K. phaffii* features high secretory capacity but secretes only low amounts of endogenous proteins. Different to *S. cerevisiae*, the Golgi apparatus of *K. phaffii* is stacked like in mammals. While it is tempting to speculate that Golgi architecture is correlated to the high secretion levels or the different N-glycan structures observed in *K. phaffii*, there is recent evidence against this. We

Received: 18 August 2021; **Accepted:** 23 November 2021

conclude that *K. phaffii* is a yeast with unique features that has a lot of potential to explore both fundamental research questions and industrial applications.

Keywords: *Komagataella phaffii*; non-conventional yeast; biotechnology; protein production; carbon metabolism; methylotrophy

INTRODUCTION

During the course of evolution, many different lifestyles emerged among yeasts. Thanks to their diversity, the world is surrounded by distinct yeast species inhabiting many different environments, metabolizing several carbon sources and producing a variety of metabolites (Kurtzman, Fell and Boekhout 2011). *Komagataella phaffii* (formerly known as *Pichia pastoris*) is among the Ascomycota yeasts from the Saccharomycetes class (Heistinger, Gasser and Mattanovich 2020). Approximately, 5200 genes are encoded on four rather large chromosomes, with a total genome size of 9.4 Mb. *Komagataella phaffii* is a methylotrophic yeast that can utilize methanol as the sole carbon and energy source. Besides methanol, it can grow on a number of carbon sources including glucose, glycerol, ethanol, trehalose, L-rhamnose, mannitol, sorbitol, D-glucitol, lactic acid, succinic acid, acetic acid and citric acid (varies among different strains; Sreekrishna *et al.* 1997; Kurtzman, Fell and Boekhout 2011; Sahu and Rangarajan 2016).

Since the 1990s, *K. phaffii* is among the preferred hosts for recombinant protein production, and more recently, *K. phaffii* has also been employed for non-protein products (Peña *et al.* 2018; Werten *et al.* 2019; Zhu *et al.* 2019; Karbalaei, Rezaee and Farsiani 2020; Duman-Özdamar and Binay 2021; Gao, Jiang and Lian 2021). Furthermore, *K. phaffii* serves as a model organism in biomedical research and basic cell biology (Bernauer *et al.* 2020). Its attractivity as a biotechnological host and model organism is strongly connected to features that distinct *K. phaffii* from other yeasts and classify it as a non-conventional yeast.

METHYLOTROPHY

The most obvious non-conventional feature of *K. phaffii* is its ability to metabolize methanol. Recent phylogenetic analyses of the budding yeasts show that all methylotrophic yeasts cluster in one clade (Shen *et al.* 2018), indicating that methylotrophy has probably evolved only once. The first enzymatic step of methanol utilization (MUT), the oxidation of methanol to formaldehyde, is catalyzed by an alcohol oxidase (Aox1/2). Sequence similarity proposes that this enzyme belongs to the group of glucose–methanol–choline (GMC) oxidoreductases (Ozimek, Veenhuis and van der Klei 2005). Different to other methylotrophic yeasts, *K. phaffii* possesses two genes encoding AOX, which are 97% identical on the amino acid level and form a 600-kDa homo-octamer containing eight FAD cofactors (Vonck, Parcej and Mills 2016; Ito *et al.* 2007). Formaldehyde is then assimilated by a transketolase, forming dihydroxyacetone and glyceraldehyde-3-phosphate from xylulose-5-phosphate and formaldehyde. This enzyme called dihydroxyacetone synthase is believed to have evolved from transketolase, however, having gained the ability to accept formaldehyde as an acceptor substrate (Kato *et al.* 1982). Xylulose-5-phosphate is recycled by sugar phosphate interconversions. In principle, all enzymes necessary to perform this recycling are present in the pentose phosphate pathway (PPP) of yeasts, as well as the reactions for the dissimilation of formaldehyde to CO_2, by which *K. phaffii* generates energy. Evolution of methylotrophy, however, modified also the spatial distribution of the pathway enzymes to peroxisomes, creating methanol assimilation organelles. Rußmayer *et al.* (2015) showed that the entire xylulose monophosphate (XuMP) cycle for methanol assimilation is localized in the peroxisomes. Instead of using the canonical cytosolic PPP enzymes, the XuMP cycle is catalyzed by a set of isoenzymes (Das1/2, Fba1-2, Rki1-2, Rpe1-2 and Tal1-2), encoded by duplicated gene copies with peroxisomal targeting signals. This leads to the speculation whether the peroxisomal localization is an essential feature to enable methylotrophy in yeasts. This idea is supported by the fact that, despite substantial effort it has not been successful so far to create methylotrophy in bakers yeast *Saccharomyces cerevisiae* by simply transferring alcohol oxidase and dihydroxyacetone synthase to it. It may well be that the other peroxisomal pathway enzymes were missing, or specific features of peroxisomes that methylotrophic yeasts have likely evolved. With a transposon tagged knockout strategy, Zhu *et al.* (2018) have searched for the essential gene set that is specific for growth on methanol. The annotation of this gene set is poor, but could be improved by a meta-analysis with the functional annotation by Valli *et al.* (2016). Methanol-only essential genes included the expected pathway enzyme genes, a set of genes responsible for peroxisomal protein targeting, as well as genes encoding peroxisomal structural proteins. As all these genes were not essential for growth on glucose it can be speculated that they have a different function, are less abundantly expressed or are absent in non-methylotrophic yeasts. Espinosa *et al.* (2020) have recently reported that *S. cerevisiae* could be evolved to assimilate methanol in addition to glucose. While the exact pathway remains unresolved, these data indicate (1) that a basic enzymatic pattern for MUT is encoded in the budding yeast genome, and (2) that the native enzymes are not sufficient (at least in their localization) to support growth, so that the uniqueness of the methanol pathway in methylotrophic yeasts is underlined.

CRABTREE PHENOTYPE AND DIFFERENT PATHWAYS FOR CARBON UTILIZATION

Glucose is one of the carbon sources on which *K. phaffii* can readily grow and form biomass (Kurtzman, Fell and Boekhout 2011). However, glucose uptake is limited compared to Crabtree-positive yeasts such as *Saccharomyces cerevisiae* ($q_{Smax} \approx 0.35$–0.60 $g_S/g_X/h$ for *K. phaffii* and 2.88–2.16 $g_S/g_X/h$ for *S. cerevisiae* in aerobic cultures (Diderich *et al.* 1999; Otterstedt *et al.* 2004; Maurer *et al.* 2006; Ata *et al.* 2018). It is argued that the high number of hexose transporters might be one of the reasons for the superior glucose uptake metabolism of Crabtree-positive (respiro-fermenting) yeasts. Accordingly, *S. cerevisiae* possesses more than 15 hexose transporters which allows it to transport glucose at a high rate (Boles and Hollenberg 1997; Elbing *et al.* 2004). Eventually, the high glucose flux exceeds the respiration capacity and leads to overflow metabolism of glycolysis which results in reduced biomass yields and ethanol production even at aerobic conditions (Fig. 1). In contrast, *K. phaffii* has a reduced number of hexose transporters similar to other respiratory yeasts such as *Kluyveromyces lactis*, *Hansenula polymorpha*

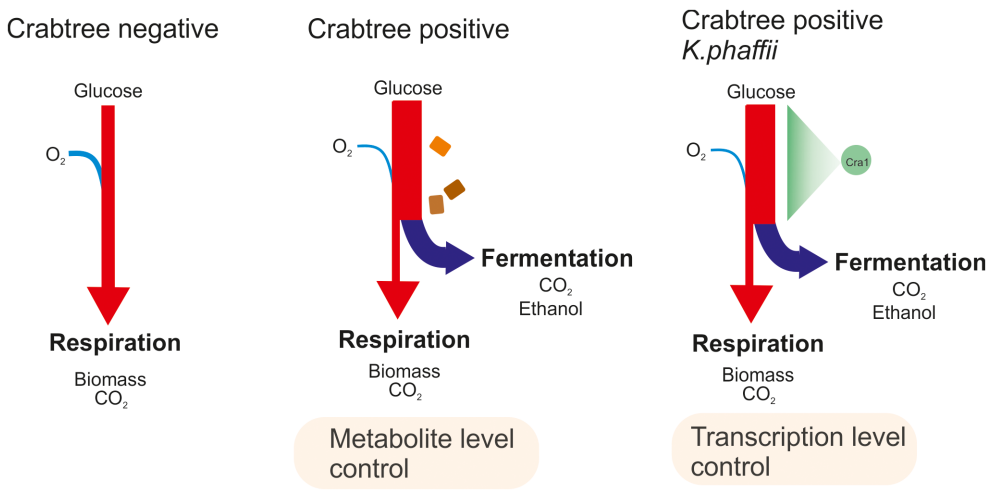

Figure 1. Regulation of the Crabtree phenotype in yeasts. (Left) Crabtree negative yeasts have a limited glucose uptake capacity and exclusively ferment under aerobic conditions. (Middle) Due to the overflow metabolism, Crabtree positive yeasts have a respiro-fermentative metabolism which is controlled at the metabolite level. (Right) Overexpression of *CRA1* upregulates the glycolytic genes leading to overflow metabolism, and demonstrating that glycolysis in *K. phaffii* is controlled at the transcriptional level.

(syn *Ogataea polymorpha and Ogataea parapolymorpha*) and *Schefferomyces* (*Pichia*) *stipitis* (Mattanovich *et al.* 2009). It possesses one high affinity transporter (*GTH1*) and three *HXT* isozymes (*HXT1*, *HXT2* and *HXT400*; Valli *et al.* 2016). This prevents the overflow metabolism of glycolysis so that *K. phaffii* exclusively performs respiration under fully aerobic conditions (Ata *et al.* 2018). Therefore, it is classified as canonical Crabtree-negative yeast. Additionally, the fraction of carbon entering the PPP is higher in *K. phaffii*, in contrast to *S. cerevisiae* where the main flux is through glycolysis (see below). Recently, a transcription factor (TF), *CRA1* (a Gal4-like TF, homolog of Sc*GAL4*) was identified to be controlling glycolysis and fermentation metabolism of *K. phaffii* (Ata *et al.* 2018). Constitutive overexpression of this TF resulted in increased glucose uptake and ethanol production rates, as well as upregulation of glycolytic genes, which switched the Crabtree-negative phenotype of *K. phaffii* to Crabtree-positive (Fig. 1). Interestingly, the hexose transporters were not regulated by *CRA1* overexpression suggesting that hexose transporters are not responsible for the increased glycolytic flux in *K. phaffii*.

Komagataella phaffii cannot metabolize galactose but according to sequence homology it contains a *GAL10* gene in addition to *CRA1* (homolog of Sc*GAL4*; Valli *et al.* 2016). In a comparative genomics analysis, it was shown that the ability of galactose utilization has been lost at least seven times during yeast evolution (Riley *et al.* 2016). It might be expected that a few *GAL* homologs in *K. phaffii* might have been inherited from an ancestral yeast that could consume galactose and eventually were lost during its evolution. The function of *GAL10* in *K. phaffii* is not known, but the Sc*GAL4* homolog *CRA1* is apparently controlling glycolysis in *K. phaffii*, in contrast to *S. cerevisiae* where Gal4 regulates the *GAL* genes (Ata *et al.* 2018). The whole genome duplication (WGD) marks a transition in the role of Gal4 where its function switched from a generalist TF controlling glycolysis to a specialist TF regulating galactose metabolism (Choudhury and Whiteway 2018). In yeasts which originate prior to this transition like *Candida albicans*, *Debaryomyces hansenii*, or *Schizosaccharomyces pombe*, the Leloir pathway for galactose utilization is controlled by Cph1. In post-WGD yeasts, Cph1 was eventually lost, Mig1 and Gal4 (along with other regulatory proteins) have been recruited to co-regulate the Leloir pathway, while in pre-WGD yeasts like *K. phaffii* and *C. albicans*, Gal4 has a regulatory function in the central carbon metabolism (Martchenko, Levitin and Whiteway 2007; Askew *et al.* 2009; Ata *et al.* 2018).

There is a high diversity among different yeast species on the ability of growing on glycerol as the sole carbon source (Kurtzman, Fell and Boekhout 2011). Unlike *S. cerevisiae*, *K. phaffii* can grow on glycerol in minimal media without requiring any additional supplements (μ_{max} of 0.26/h, $q_{Glycerol\_max} = 0.37$ $g_S/g_X/h$; Jahic *et al.* 2002). Possession of four H^+/glycerol symporters in addition to a Fps1-type glycerol facilitator demonstrates the superiority of the glycerol uptake metabolism compared to many other yeasts (Lages, Silva-Graça and Lucas 1999; Mattanovich *et al.* 2009). Of these four H^+/glycerol transporters, GT1 (encoded by *STL1-1*, *PP7435_Chr1-0321*) was found to be one of the factors affecting the crosstalk between the glycerol and methanol metabolism in *K. phaffii* (Zhan *et al.* 2016). While the deletion of GT1 (*STL1-1*) did not cause a significant growth impairment in glycerol-based medium, it relieved glycerol repression on P_{AOX1}. Additionally, overexpression of *GT1* repressed *MXR1* and *AOX1* expression, whereas *MXR1* overexpression repressed *GT1* (Zhan *et al.* 2016, 2017; Li *et al.* 2018).

In addition to regulating the MUT pathway, glycerol metabolism in *K. phaffii* seems to have a global effect on NADPH balance. Metabolic flux analysis with ^{13}C-labeled glycerol showed that the major cytosolic NADPH source might be the glycerol catabolic pathways (Tomàs-Gamisans *et al.* 2019). In contrast to the previous assumption that the oxidative branch of the PPP is the main source for cytosolic NADPH, it was demonstrated that the flux through the PPP is almost negligible. Furthermore, it was hypothesized that, among the different glycerol catabolic pathways (Klein *et al.* 2017), the NADP-dependent glycerol oxidation pathway is the major cytosolic NADPH source in glycerol grown cultures. However, PPP seems to have a more profound effect on NADPH generation when the cells are grown on glucose: The split ratio of PPP flux is around 40–55% in *K. phaffii* (Baumann *et al.* 2010; Nocon *et al.* 2016; Ata *et al.* 2018) as opposed to *S. cerevisiae* which has a split ratio of 4–17% (Gombert *et al.* 2001; Maaheimo *et al.* 2001; Velagapudi *et al.* 2007).

L-rhamnose metabolism occurs in fungi *via* an oxidative (non-phosphorylated) pathway (Koivistoinen *et al.* 2012). A comparative genomics study showed that the genes of L-rhamnose

metabolism are clustered among the yeasts analysed in the study (Riley *et al.* 2016). It appears that *K. phaffii* has five genes (*LRA1*, *LRA2*, *LRA3*, *LRA4* and *TRC1*) associated with L-rhamnose metabolism (Liu, Styles and Fink 2016; Valli *et al.* 2016). However, *PP7435_Chr1-0845 (LRA3)* does not cluster with the other genes, contrary to the gene arrangement in *S. stipitis*. *Komagataella phaffii* can grow on L-rhamnose as the sole carbon source with a slightly lower growth rate than on glucose and the expression of the related genes are induced by rhamnose and repressed in the presence of glucose (Liu, Styles and Fink 2016).

GENE REGULATION AND TFs

Methylotrophic yeasts offer a repertoire of regulated, strong promoters that are naturally regulating MUT pathway genes (Ergün *et al.* 2019b). Typically, the promoters of MUT pathway genes are tightly repressed on repressing carbon sources such as glucose and strongly induced when shifted to methanol. However, they demonstrate different modes of derepression among methylotrophic yeasts. The promoter of the *K. phaffii AOX1* gene P_{AOX1} is strongly induced by methanol while repressed by ethanol, glucose and glycerol (Ergün *et al.* 2020). Derepressed *K. phaffii* cells (repressing carbon source is depleted or non-repressing carbon source is present) display approximately 2% *AOX1* transcriptional activity of the methanol-induced level, whereas methanol-induced cells display more than 1000-fold higher activity than fully repressed (glucose-grown) cells (Lin-Cereghino *et al.* 2006). In contrast, promoters of orthologous genes *AOD* (alcohol oxidase) in *Candida boidinii*, *MOX* (methanol oxidase) in *H. polymorpha* and *MOD1* (methanol oxidase 1) in *Pichia methanolica* show respectively 3–30%, 60–70% and 60–70% derepression in glycerol compared to their methanol-induced levels (Hartner and Glieder 2006). Substantial amounts of heterologous protein were expressed by P_{AOX1} in *H. polymorpha* on glycerol, and P_{AOX1} was regulated in the same manner as P_{MOX} (Raschke *et al.* 1996). These findings suggest that rather than promoter *cis*-acting DNA elements, components of the cellular transcriptional machinery determine the expression mode of P_{AOX1} on glycerol. Understanding the transcriptional regulation of P_{AOX1} and other MUT pathway genes has paramount importance to enhance control over natural promoters and design new expression mechanisms in *K. phaffii*.

An overview of the characterized transcription factors (TFs) involved in regulation of carbon source utilization and their respective function in *K. phaffii* is given in Table 1 and Fig. 2.

S. cerevisiae Adr1 (alcohol dehydrogenase synthesis regulator) has a pivotal role in the activation of glucose repressible genes, peroxisomal protein genes and ethanol, glycerol and fatty acid utilization pathway genes (Young *et al.* 2003). Alcohol dehydrogenase Adh2 is the first enzyme of the ethanol utilization pathway that catalyzes oxidation of ethanol to acetaldehyde. Adr1 and Cat8 synergistically activate *S. cerevisiae ADH2* transcription and many other ethanol utilization pathway genes (Young *et al.* 2003). The *K. phaffii* Adr1-homologue Mxr1 (methanol expression regulator 1) has gained new functions and lost others through evolution as a result of changes in the environmental conditions, cell physiology and spectrum of genes that it controls. Deletion of *MXR1* caused total transcriptional shut down of *K. phaffii* P_{AOX1} and cells cannot grow on methanol, while it had less detrimental effect on *K. phaffii* P_{ADH2} activation and growth on ethanol compared to *S. cerevisiae* transcriptional regulation (Ergün 2018). *MXR1* is constitutively expressed at low levels and activates MUT pathway and peroxisome biogenesis (PEX) genes (Lin-Cereghino *et al.* 2006). Mxr1 is cytoplasmic in glucose-grown cells but localized to the nucleus in cells cultured on gluconeogenic substances (Lin-Cereghino *et al.* 2006). Similar to *S. cerevisiae* Adr1, the 14–3-3 protein directly interacts with Mxr1 by phosphorylation and inhibits its activity (Parua *et al.* 2012). The ethanol-repressible nature of P_{AOX1} has been investigated through Mxr1, however, the answer came from a different side. Promoter engineering of P_{AOX1} by introducing a Cat8 *cis*-acting DNA motif converted ethanol-repressible P_{AOX1} to the ethanol inducible $P_{AOX1/Cat8\text{-}L3}$ variant (Ergün *et al.* 2020). This demonstrates that the ethanol repressible nature of P_{AOX1} is due to the absence of an ethanol responsive *cis*-acting element. Addition of further Cat8 *cis*-acting motifs enhanced *K. phaffii* P_{ADH2} expression 4.8-fold on ethanol (Ergün *et al.* 2019a), while the *K. phaffii* Δ*cat8-1*Δ*cat8-2* mutant lost its ability to grow on ethanol (Ergün 2018; Barbay *et al.* 2021), emphasizing the importance of the *K. phaffii* Cat8-1 and Cat8-2 TFs for ethanol regulation.

In the methylotrophic yeast *C. boidinii* Trm1 and Trm2 (transcriptional regulation of methanol induction) are essential TFs for the expression of MUT pathway and PEX genes (Sasano *et al.* 2008, 2010). Trm2 is the homologue of *K. phaffii* Mxr1 and responsible for the activation of methanol-inducible genes by relieving glucose repression, and also essential for Trm1-dependent gene activation (Sasano *et al.* 2010). *Komagataella phaffii* Trm1, the homologue of *C. boidinii* Trm1, is another positively acting TF of MUT pathway and PEX genes. The respective Δ*trm1* mutant showed impaired P_{AOX1} activity and growth on methanol (Sahu, Krishna Rao and Rangarajan 2014), which could be rescued by *MIT1* overexpression (Wang *et al.* 2016b).

Mit1 (methanol-induced TF1) positively regulates MUT pathway genes, but not peroxisomal genes and activates P_{AOX1} in response to methanol while it represses P_{AOX1} transcription on glycerol (Wang *et al.* 2016b). *Komagataella phaffii* Δ*mit1* cannot grow on methanol and P_{AOX1} is not active. Complementation of the Δ*mit1* mutant with *H. polymorpha* Mpp1 (methylotrophic peroxisomal protein) restored growth and P_{AOX1} activity on methanol, and furthermore, lead to remarkable *AOX1* expression levels on glycerol (Wang *et al.* 2016b). Structural differences between Mit1 and Mpp1 likely contribute to the differential expression mode of P_{AOX1} and P_{MOX}.

Mxr1, Mit1 and Trm1 are binding to P_{AOX1} at different sites and cooperatively activate P_{AOX1}, but overexpression of Mit1 or Trm1 does not restore P_{AOX1} activity in the Δ*mxr1* mutant. Firstly, derepression of P_{AOX1} is mediated by the master TF Mxr1, then Trm1 and Mit1 contribute to promoter activation (Wang *et al.* 2016b). *TRM1* is expressed constitutively in glucose, glycerol and methanol, while *MIT1* expression is strongly induced on methanol. Both Mit1 and Trm1 are localized to the nucleus in glucose, glycerol and methanol conditions. Mit1 binds to P_{AOX1} in the presence of all three carbon sources, however, Trm1 only binds when cells are grown on methanol or glycerol but not on glucose (Wang *et al.* 2016b).

Rop1 (repressor of phosphoenolpyruvate carboxykinase; Kumar and Rangarajan 2012) and Nrg1 (Wang *et al.* 2016a) are negative regulators of *K. phaffii* MUT pathway and PEX genes. Rop1 and Mxr1 function antagonistically but exhibit the same DNA binding specificity whereby Rop1 binds to DNA with higher affinity than Mxr1 (Kumar and Rangarajan 2012). *S. cerevisiae* Nrg1 and Nrg2 (negative regulator of glucose-repressed genes) mediate glucose repression (Zhou and Winston 2001). *K. phaffii* Nrg1 represses MUT pathway and PEX genes in glucose and glycerol conditions. It directly binds to five positions on P_{AOX1}, two of which overlap with MXREs (Wang *et al.* 2016a). The Δ*nrg1* mutant showed a growth defect when cultivated on glucose,

Table 1. List of TFs that are experimentally confirmed to regulate carbon source utilization in *K. phaffii*.

TF	Uniprot ID*	Homologs in other yeasts	Function	References
Aft1	F2QPE8_KOMPC		Regulates genes of carbohydrate metabolism and recombinant protein secretion	Ruth *et al.* (2014)
Cat8-1	F2QS26_KOMPC	*S. cerevisiae* Cat8, *K. lactis* Cat8	Activates glyoxylate cycle and EUT pathway in ethanol grown *K. phaffii*, required for growth on acetate	Barbay *et al.* (2021)
Cat8-2	F2QYX3_KOMPC	*S. cerevisiae* Cat8/Sip4, *K. lactis* Cat8/Sip4	Activates carnitine shuttle and EUT pathway in ethanol grown *K. phaffii*	Barbay *et al.* (2021)
Cra1	F2QQF5_KOMPC	*S. cerevisiae* Gal4	Controls glycolysis and fermentation metabolism	Ata *et al.* (2018)
Flo8	F2QYE9_KOMPC	*S. cerevisiae* Flo8	Master regulator of filamentous growth and surface adherence, also involved in glucose repression	Rebnegger *et al.* (2016)
Mig1-1	F2QZJ1_KOMPC	*S. cerevisiae* Mig1, *H. polymorpha* Mig1	Repressor of MUT pathway and PEX genes	Wang *et al.* (2017); Shi *et al.* (2018)
Mig1-2	F2QPW6_KOMPC	*S. cerevisiae* Mig1, *H. polymorpha* Mig2	Repressor of MUT pathway and PEX genes	Wang *et al.* (2017); Shi *et al.* (2018)
Mit1	F2QV89_KOMPC	*H. polymorpha* Mpp1	Activator of MUT pathway but not PEX genes on methanol, represses P_{AOX1} in response to glycerol	(Wang *et al.* (2016b)
Mxr1	F2QZ27_KOMPC	*S. cerevisiae* Adr1, *C. boidinii* Trm2	Activator of MUT pathway and PEX genes	Lin-Cereghino *et al.* (2006)
Trm1	F2QZY1_KOMPC	*C. boidinii* Trm1	Activator of MUT pathway and PEX genes	Sahu *et al.* (2014)
Trc1	F2QZI4_PICP7	*S. stipitis* Trc1	TF suggested to be involved in the regulation of *LRA* genes of the rhamnose metabolism	Liu, Styles and Fink (2016)
Nrg1	F2QUX2_KOMPC	*S. cerevisiae* Nrg1/2	Repressor of MUT pathway and PEX genes	Wang *et al.* (2016a)
Rop1	F2QW29_KOMPC		Repressor of MUT pathway, PEX genes and phosphoenolpyruvate carboxykinase	Kumar and Rangarajan (2012)

*For some TF genes different annotations are used in literature. In order to avoid ambiguity, their UniProt IDs are provided.

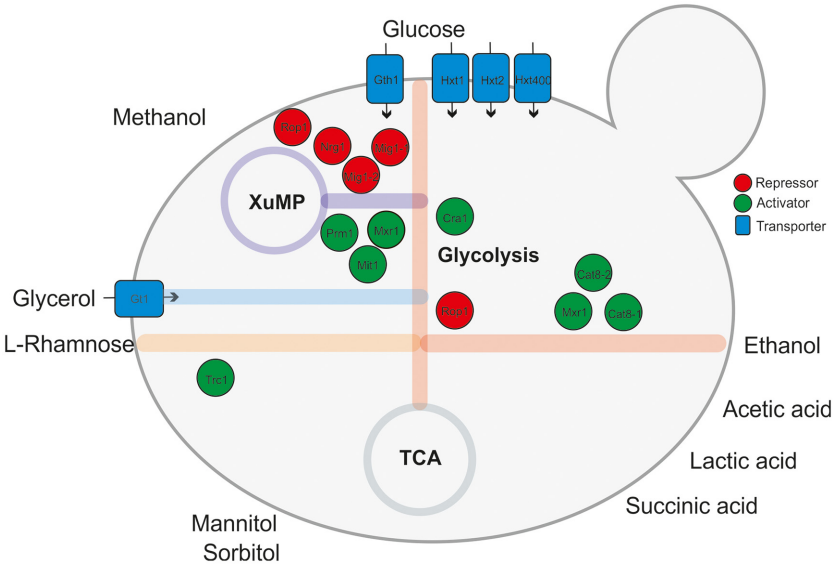

Figure 2. TFs involved in the regulation of *K. phaffii*'s central carbon metabolism. Blue rectangles represent the transporters. Activators and repressors are shown in green and red, respectively. See Table 1 for detailed information on the TFs. XuMP: Xylulose monophosphate pathway. TCA: Tricarboxylic acid cycle.

glycerol or methanol. Unexpectedly, overexpression of *NRG1* was found to enhance secretion of recombinant proteins through a yet unidentified mechanism (Stadlmayr *et al.* 2010).

S. cerevisiae Mig1 and Mig2 are involved in glucose repression of different carbon metabolism genes (Schüller 2003). *Komagataella phaffii* Mig1-1 and Mig1-2 are mainly localized in the nucleus in glucose or glycerol conditions, while they translocate to the cytosol when cells are grown on methanol (Wang *et al.* 2017). Glycerol-induced suppression of P_{AOX1} is partially removed in the Δ*mig1-1* mutant, while no effect is observed in the Δ*mig1-2* mutant. The double knock-out Δ*mig1-1*Δ*mig1-2* mutant showed increased transcriptional activation (Wang *et al.* 2017), which seems to be mediated through activation of Mit1 (Shi *et al.* 2018). On the other hand, neither deletion of *MIG1-1* nor *MIG1-2* deregulated P_{AOX1} on glucose. The double deletion of *MIG1-1* and *MIG1-2* led to a growth defect on glycerol and glucose (Shi *et al.* 2018).

MATING AND MATING-TYPE SWITCHING

K. phaffii is a preferentially haploid yeast usually propagating by mitotic cell division. Sexual reproduction (mating and spore formation) is possible but can only be observed under nitrogen starvation conditions (Feng *et al.* 2020). In contrast to *S. cerevisiae*, where haploid cells mate spontaneously to form stable diploids, the mating-type (*MAT*) genes and most other mating-relevant genes of *K. phaffii* are not expressed in rich medium (Heistinger, Gasser and Mattanovich 2018). Once mating has occurred, diploid *K. phaffii* cells rapidly undergo meiosis and sporulation if no selective pressure is applied. This coupling of mating and sporulation has also been observed in other yeasts like *Candida lusitaniae* or *K. lactis* but is absent in *Saccharomyces* species, where the lack of nitrogen and presence of a non-fermentable carbon source act as a trigger for sporulation of diploid cells (Booth, Tuch and Johnson 2010; Merlini, Dudin and Martin 2013; Sherwood *et al.* 2014; Hanson and Wolfe 2017).

Although the mating-type and mating behavior of yeasts is generally regulated by the *MAT* genes, there are significant differences in the mechanism of mating-type switching and the control of cell type regulations between different species. *K. phaffii* has a two-locus mating-type system, where both loci, containing either *MAT***a**1 and *MAT***a**2 or *MATα*1 and *MATα*2, are located at the beginning of chromosome 4 (Fig. 3). The two loci are flanked by inverted repeat (IR) sequences containing one (*DIC1*) and three (*SLA2*, *SUI1* and *CWC25*) genes and are separated by around 135 kb of DNA sequence also containing the centromere (Hanson, Byrne and Wolfe 2014). Under mating conditions, the *MAT* locus next to the telomeric region remains silenced, while the genes in the second *MAT* locus are transcribed and thereby determine the mating-type of the cell (Heistinger, Gasser and Mattanovich 2018).

K. phaffii is a secondary homothallic yeast and mating-type switching takes place by homologous recombination at the 'outer' inverted repeat region containing the *DIC1* genes. This leads to an inversion of the whole chromosomal region between the two *MAT* loci and a swap of the *MAT* allele in the active mating-type locus (Hanson, Byrne and Wolfe 2014). A similar, well studied mating-type system is found in the methylotrophic yeast *O. polymorpha*, where the mating loci are flanked by one inverted repeat region and silencing of the silent *MAT* locus is mediated by its proximity to centromeric heterochromatin (Hanson, Byrne and Wolfe 2014; Maekawa and Kaneko 2014). A recent study analysing the *MAT* loci of more than 300 budding yeast species found that such two-locus flip-flop switching mechanisms have evolved independently at least ten times, while the three-locus mating type system as it is found in *S. cerevisiae* and closely related species seems to have evolved only once within the budding yeasts (Krassowski *et al.* 2019). In *S. cerevisiae*, two silent *MAT* loci (*HMLα* and *HMR***a**) serve as template for mating-type switching via a synthesis dependent strand annealing mechanism. Switching is initiated by the HO endonuclease, which introduces a double strand break at the active *MAT* locus (Strathern *et al.* 1982; Ira, Satory and Haber 2006; Haber 2012; Hanson and Wolfe 2017). The genomes of methylotrophic yeasts do not contain orthologs of HO endonuclease and although it has been shown that nitrogen starvation induces switching via an Rme1 and Ste12 dependent pathway in *O. polymorpha*, it remains unclear by which molecular mechanism mating-type switching is initiated (Hanson, Byrne and Wolfe 2017; Yamamoto *et al.* 2017).

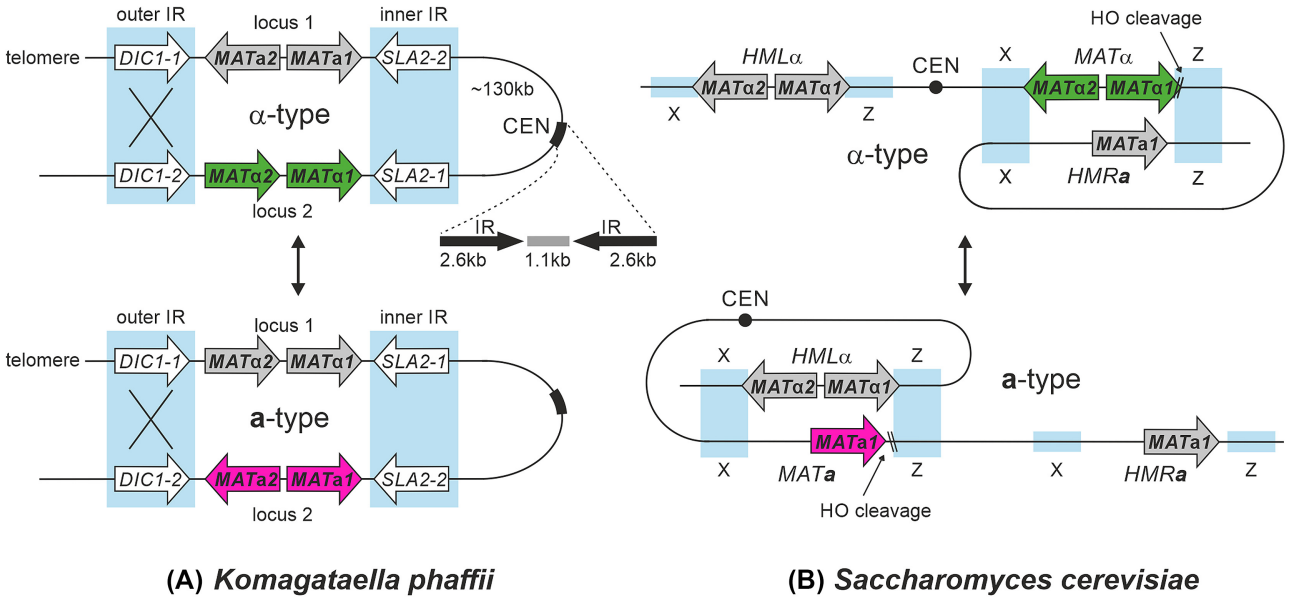

Figure 3. Mating-type systems of *K. phaffii* and *S. cerevisiae*. **(A)** Homologous recombination at the outer inverted repeat (IR) region results in mating-type switching by inversion of the genomic region between the two *MAT* loci, including the centromere of chromosome 4 in *K. phaffii*. **(B)** In *S. cerevisiae*, HO endonuclease initiates mating-type switching. The two silent *MAT* loci (*HMLα* and *HMRa*) serve as template for mating-typ switching *via* a synthesis dependent strand annealing mechanism. The concept to display the mating type switches is based on Hanson and Wolfe (2017).

As described above, the Mat TFs encoded in the active *MAT* locus determine the mating-type of a cell. In *K. phaffii*, Mat**a**2 and Matα1 in **a**- and α-type cells, respectively, activate the expression of mating-type specific genes like the pheromone and pheromone surface receptor genes and are thereby essential for mating. In diploid cells, *MAT***a**1 and *MAT*α2 are required for sporulation and the repression of mating (Heistinger, Gasser and Mattanovich 2018). In *S. cerevisiae* and other post-WGD yeasts, a rewiring of this regulatory network resulted in the loss of **a**-specific gene activation by Mat**a**2 and the requirement for **a**-specific gene repression by Matα2 in α-type cells (Tsong *et al.* 2006; Sorrells *et al.* 2015).

Due to the requirements of prolonged nitrogen starvation and solid media, mating of *K. phaffii* is unlikely to occur under industrial production conditions, even though there are always cells of both mating-types present in the culture when working with homothallic strains. However, mating represents a useful tool for strain development as it can be applied to combine interesting traits, investigate genetic determinants of relevant traits by methods like quantitative trait loci (QTL) mapping or the generation of combinatorial libraries (Chen *et al.* 2012; Liti and Louis 2012; Swinnen, Thevelein and Nevoigt 2012). The use of heterothallic strains, such as the *K. phaffii* CBS2612 Δ*dic1-2* strains generated by the deletion of the 'outer' homologous region required for mating-type switching, further enables the use of *K. phaffii* for classical genetic studies (Heistinger, Gasser and Mattanovich 2018).

CENTROMERES

The centromeres of *K. phaffii* consist of a 2–2.7 kb inverted repeat region with a central core sequence of around 1 kb. Each of the four chromosomes has one centromere which is unique in sequence. The centromere of chromosomes 3 and 4 are found close to one chromosome end, with the centromere on chromosome 4 being located within the invertible region flanked by the *MAT* loci (Figure 3). The centromeres were identified as large non-transcribed regions and replicate early during cell division (Coughlan *et al.* 2016; Sturmberger *et al.* 2016). A ChIP-seq experiment showed that the centromere-specific histone variant Cse4 is most abundant in the core sequence but its signal can be detected all along the non-transcribed centromeric region. The orientation of the centromeres is variable in different *K. phaffii* isolates, indicating that recombination at the inverted repeats can occur (Coughlan *et al.* 2016). The inverted repeat centromeres of *K. phaffii* are highly different from the small point centromeres found in *S. cerevisiae* and closely related species, which are only around 125 bp long and defined by a clear consensus sequence (Hegemann and Fleig 1993). Although inverted repeat centromeres are also found in other yeasts like *Candida tropicalis*, the small genome and efficient tool for manipulation make *K. phaffii* an interesting system to study centromere function (Chatterjee *et al.* 2016). So far, two studies have reported the characterization of *K. phaffii* plasmid vectors carrying whole centromeric sequences. Those plasmids were found to increase mitotic stability while maintaining a low copy number when compared to classical ARS plasmids (Nakamura *et al.* 2018; Piva *et al.* 2020).

MORPHOLOGY SWITCHES

Upon experiencing adverse environmental conditions, budding yeasts can switch from unicellular to multicellular lifestyle, leading to flocculation, pseudohyphae formation or invasive growth (Brückner and Mösch 2012). These morphogenetic events give rise to subpopulations of cells exhibiting different phenotypes, thus providing advantage for adaptation to environmental changes and increasing chances of survival. In *S. cerevisiae* these morphology switches are associated with the flocculin (*FLO*) gene family, which has five dominant members encoding GPI-anchored cell-wall proteins (Verstrepen and Klis 2006; Willaert 2018). Out of these Flo1, Flo5, Flo9 and Flo10 are involved in flocculation, while Flo11 is responsible for filamentous growth (Guo *et al.* 2000; Van Mulders *et al.* 2009; Goossens and Willaert

2012). Most *S. cerevisiae* laboratory strains are devoid of flocculation and filamentous growth, as there is a defect in the *FLO8* gene encoding the master transcriptional activator (Liu, Styles and Fink 1996).

Morphological differentiations have also been observed in *Komagataella* species, however, until recently not much was known of the genetic and biochemical basis underlying these phenotypes. *K. phaffii* possesses an expanded *FLO* gene family consisting of 12 members containing different Flo-domains, and a Flo8-type TF (De *et al.* 2020). Pseudohyphae formation and surface adherence are absent when Flo8 is deleted, indicating that Flo8 is also the major regulator of filamentous growth in *K. phaffii* (Rebnegger *et al.* 2016; De *et al.* 2020). In both, *S. cerevisiae* and *C. albicans*, Flo8 was demonstrated to form a heterodimer with Mss11 *via* their N-terminal LisH domains that cooperatively regulates filamentous growth (Su *et al.* 2009; Kim *et al.* 2014). However, while also in *K. phaffii* Flo8 contains an N-terminal LisH domain no Mss11 homolog was identified (De *et al.* 2020), suggesting that Flo8 acts as a homodimer in this organism. While there is a clear ortholog of *FLO11*, for the other structural *FLO* genes no distinct homologs can be identified (Kock *et al.* 2018; Brückner *et al.* 2020; De *et al.* 2020). Kock *et al.* (2018) studied the nine predicted adhesins with a lectin-like PA14 domain present in *Komagataella* spp, and identified that the *Komagataella* adhesins form a unique clade in the fungal kingdom, thus stressing that it is hard to infer *FLO* gene function from one yeast species to another. Among them KpFlo1 (also named Cea1 or Flo5-2) was discovered as the first fungal adhesin showing high specificity for terminal β-GlcNAc capped glycans including chitinous polymers. The authors speculate that the *Komagataella* PA14 domain containing proteins evolved to adapt the cells to their specific habitat to govern cell-substrate interactions different from flocculation.

Only very little is known about the mechanisms leading to floc formation in *K. phaffii*. Flocculation is caused by homotypic cell-cell adhesion, whereby yeast cells aggregate into multicellular masses (flocs) that sediment out of the medium (Soares 2011). In *S. cerevisiae*, flocculation is triggered by carbon source limitation, pH variations, external stressors or the presence of ethanol or ions e.g. Ca^{2+} (Soares 2011). Cell–cell adhesion is depending on the N-terminal PA14 lectin domains present in *S. cerevisiae* Flo1, Flo5, Flo9 and Flo10 and their reversible binding of cell wall mannans (Goossens *et al.* 2015). The respective genes are located adjacent to telomeres, and are silenced through their subtelomeric localization during normal growth conditions (Soares 2011; Cullen and Sprague 2012). It is not known so far which members of the *FLO* gene family are responsible for flocculation in *K. phaffii*. Unlike in *S. cerevisiae*, no silencing of the subtelomeric *FLO* genes was observed in exponential growth conditions at pH 5.0 (De *et al.* 2020).

Mbawala *et al.* (1990) showed that cells exhibiting higher flocculation show elongated mannose chains containing alpha-1,2 and beta-1,2 linkages, indicating that also in *K. phaffii* lectin-like mechanisms are involved in this cell–cell adhesion process. Addition of 2 mM EDTA which captures Ca^{2+} involved in glycan cross-linking was reported to reduce floc formation observed during growth of *K. phaffii* in unbuffered YPD media (Tanneberger *et al.* 2007). In our experience, floc formation can be observed macroscopically at a pH around 4.0, and is reversible upon shifting the pH (De 2019).

Importantly, flocculation and sedimentation can be advantageous in industrial bioprocesses, as a rapid and efficient means of separation of the biomass from the product containing supernatant (e.g. in industrial ethanol fermentation processes of *S. cerevisiae*, (Soares 2011)). In this respect, *K. phaffiis* cells engineered for increased rhamnose metabolic flux were shown to exhibit strong flocculation and sedimentation in rhamnose-containing media (Yan *et al.* 2018).

During pseudohyphal growth, filament-like structures are formed as cells divide but remain attached to each other (Cullen and Sprague 2012). If the filaments extend into a solid substrate, the phenomenon is termed invasive growth. In *S. cerevisiae* , pseudohyphal growth is dependent on Flo11 and more prevalent in diploid cells while invasive growth is more prevalent in haploid cells (Soares 2011; Cullen and Sprague 2012; Song and Kumar 2012). Despite being haploid, invasive growth has so far not been observed in *K. phaffii* . Also the environmental triggers leading to pseudohyphal growth are different between the two species: In contrast to *S. cerevisiae* , *K. phaffii* morphology is not affected by fusel alcohols or nitrogen starvation (De *et al.* 2020). So far, an elongated phenotype representing pseudohyphae was only observed when *K. phaffii* was cultivated at slow growth rates below $\mu = 0.075$/h in glucose-limited chemostats (Rebnegger *et al.* 2014; De *et al.* 2020). The transition of *K. phaffii* from yeast to pseudohyphal form is associated with transcriptional changes of at least three *FLO* genes (*FLO11* , *FLO400* and *FLO5-1* , all under control of Flo8) as well as chromatin remodeling. In contrast to *S. cerevisiae* , deletion of *FLO11* reduced but did not abolish pseudohyphae formation in *K. phaffii* . On the other hand, deletion of either *FLO400* or *FLO5-1* prevented the morphological changes. This was associated with a lack of *FLO11* induction upon switching to slow growth rates in glucose-limited chemostats, suggesting that *K. phaffii* Flo400 and/or Flo5-1 act as upstream signals for the activation of *FLO11* . However, it is not known which signaling cascades are responsible (De *et al.* 2020). Surprisingly, some strategies preventing morphological differentiations also resulted in higher productivity of secreted recombinant proteins (Gasser, Mattanovich and Buchetics 2014). Representative microscopic images of different morphological states of *K. phaffii* as described in this and preceding chapters are shown in Fig. 4.

NITROGEN METABOLISM

Although nitrogen is an essential macronutrient for yeasts, its metabolism has been poorly investigated. Moreover, as most of available data concern *S. cerevisiae*, highlighting the non-conventional traits of *K. phaffii* is less obvious as compared to carbon metabolism for instance. In both *K. phaffii* and *S. cerevisiae*, ammonia and organic nitrogen are the main N-sources, and N-metabolism is mainly based on deamination and transamination reactions with ammonia, glutamate and glutamine as key compounds. From BLASTp searches (E value cut-off 10^{-6}), all transporters for ammonia (three genes; Marini *et al.* 1997) and amino acids (21 genes; Bianchi *et al.* 2019) characterized so far in *S. cerevisiae* seem to have a counterpart in *K. phaffii* (data not shown). However, despite this similarity, it turned out that one needs to be careful by drawing conclusions from *S. cerevisiae*. For example, branched-chain amino acid (BCAA) biosynthesis is differentially compartmentalized in *K. phaffii*, with solely cytosolic α-isopropylmalate synthase Leu4 and BCAA aminotransferase Bat1 (Förster *et al.* 2014). Furthermore, enzymes involved in alanine (Alt1), aspartate (Aat2), glutamate (Gdh2, Gdh3 and Glt1) and lysine (Lys20 and Lys21) synthesis are found primarily in the cytosol of *K. phaffii*, whereas they localize also to the mitochondria in *S. cerevisiae* (Valli *et al.* 2020). In *S. cerevisiae,* two NADPH-glutamate dehydrogenases encoded by *GDH1* and *GDH3* catalyze the transamination of α-ketoglutarate with the formation of glutamate. *K. phaffii* possesses only one Gdh enzyme,

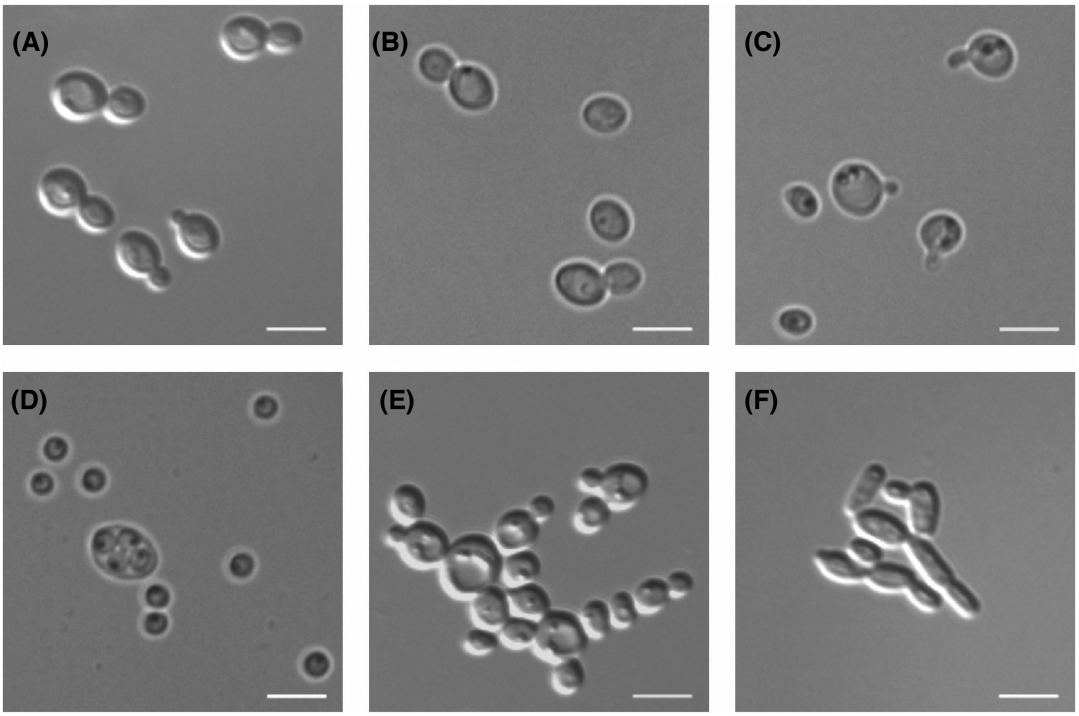

Figure 4. Diverse morphology of *K. phaffii*. **(A)** and **(B)** budding cells in rich medium, **(C)** shmoo formation, **(D)** sporulation, **(E)** flocculation and **(F)** pseudohyphal growth. Images were taken in brightfield or DIC mode. Scale bar 5 μm.

termed Gdh3 (CP014584, with about 70% of identity in amino acid sequence to *S. cerevisiae* Gdh1 and Gdh3), similarly to *Yarrowia lipolytica* (Trotter *et al.* 2020) and *S. stipitis* (Freese *et al.* 2011). Glutamate can be further transaminated by glutamine synthase (gene *GLN1*) to yield glutamine. Inversely, ammonia can be released by deamination of glutamine or glutamate by glutamate synthase (gene *GLT1*) and NAD-glutamate dehydrogenase (gene *GDH2*), respectively (Magasanik 2003). These three key metabolites are the major precursors of amino acid biosynthesis while a-ketoglutarate is the hinge between C- and N-metabolisms.

In some methylotrophic yeasts such as *H. polymorpha,* nitrate assimilation occurs by its reduction into ammonium through the consecutive action of nitrate and nitrite reductases encoded by the genes *YNR1* and *YNL1*, respectively (Siverio 2002). This is not the case for *K. phaffii* that does not assimilate nitrate or possess nitrate reductase activity (Unkles *et al.* 2004). Even though the majority of media contain ammonium as nitrogen source, *K. phaffii* can also efficiently grow on urea at least in complex medium (Guo *et al.* 2012), and the addition of urea to the fermentation medium was discussed to be beneficial for protein production purposes (Adivitiya, Mohanty and Khasa 2021). As most of the other hemiascomycetes, *K. phaffii* contains a clear homolog of the bifunctional biotin-dependent urea amidolyase Dur1,2, but not nickel-dependent urease Ure1 or the corresponding Ni^{2+}/Co^{2+} transporter Nic1 (Navarathna *et al.* 2010). Correspondingly, depletion of Co^{2+} from the medium had no effect on growth or productivity of *K. phaffii* (Pekarsky *et al.* 2020). Furthermore, there are up to five putative urea transporters found in the *K. phaffii* genome (four homologs of the *S. cerevisiae* Dur3 and one of *S. stipitis* Dur8), but nothing is known on their functionality or regulation so far (Valli *et al.* 2016, 2020).

Some specificities of *K. phaffii* could be also highlighted from amino acid catabolic pathways. For instance, the asparagine degradation in *S. cerevisiae* is composed by a single copy *ASP1* gene and four repeated copies of *ASP2* encoding cytosolic and cell-wall asparaginases (League, Slot and Rokas 2012). These enzymes catalyze the deamination of asparagine into aspartate with the release of ammonium. From BLASTp analysis, no clear Asp2 counterpart could be identified in the *K. phaffii* genome, while sequence NC_012965.1 from strain GS115 shows 57% identity with Asp1 (data not shown).

By contrast to *S. cerevisiae, K. phaffii* and other non-conventional yeasts are able to grow in minimal medium containing aspartate or glutamate as sole N- and C-sources (Sahu and Rangarajan 2016). This ability has been found to be correlated to the activity of Gdh2. A Δ*gdh2* mutant of *S. stipitis* cannot utilize glutamate as C-source while in *Y. lipolytica* the activity of Gdh2 increases up to 18-fold as compared to Gdh1 when glutamate is both, C- and N-source (Trotter *et al.* 2020). In *K. phaffii*, TFs that regulate methanol metabolism have also been found involved in this process. For instance, Mrx1 but not Trm1 or Rop1, regulates the activity of *GDH2* expression at post-transcriptional level. Mxr1 also regulates at the transcriptional level the genes *AAT1* and *AAT2* encoding mitochondrial and cytosolic aspartate aminotransferase, respectively, and the gene *GLN1*. Mxr1 Response Elements (MXREs) have been found in the promoter sequence of *AAT2* and *GLN1*. Therefore, methanol metabolism, a peculiar feature of *K. phaffii*, also controls N-metabolism. Beside this, nitrogen sources such as casamino acid have also been reported to regulate methanol metabolism (Velastegui *et al.* 2019). Indeed, expression of the genes *AOX1*, *DAS1* and *FLD1* is reduced in the presence of 0.1% of casamino acid.

Recently, a genomic survey of nitrogen assimilation pathways in budding yeast has been published (Linder 2019). By contrast to *S. cerevisiae*, the genome of *K. phaffii* contains the gene *AMO1* encoding amine oxidase that catalyzes the deamination

of aliphatic primary amine (R-NH_2) with the release of ammonia. It also contains the gene *AOC1* encoding lysyl oxidase. The corresponding enzyme has been biochemically characterized in detail (Kucha and Dooley 2001). Regarding uracil catabolism, *K. phaffii* genome contains genes *URC1* and *URC4* encoding putative cyclohydrolase and ribosyl-urea degrading enzymes which are missing in *S. cerevisiae*. In budding yeast, purines, uric acid and allantoin are all catabolized in a common pathway with ammonia as the final product. A total of 10 genes are involved in this pathway namely, *XAN1*, *XAN2*, *URO1*, *URO2*, *URO3*, *DAL1*, *DAL2*, *DAL3* and *DUR1,2*, although some of them are not present in the *S. cerevisiae* genome. These are the two *XAN* genes and the three *URO* genes that are putatively involved in the conversion of xanthine into uric acid and of uric acid into allantoin. Xanthine oxidoreductase (*XAN* genes) are known to require a molybdenum cofactor (MoCo) to be active (Mendel 2013). At least six genes, *MOC1–MOC6* are involved in molybdenum cofactor biosynthesis in eukaryotic cells. The homologs of these six genes were identified in *K. phaffii* but not in *S. cerevisiae* (Linder 2019).

PROTEIN SECRETION

Based on the status of *K. phaffii* as a popular recombinant protein production platform, a large body of dedicated research aimed at elucidating (or manipulating) the molecular mechanics governing protein synthesis, secretion and post-translational modifications (PTMs) such as disulfide bond formation, proteolytic processing as well as N- and O-glycosylation. According to published literature, secretion yields of recombinant products produced in *K. phaffii* often exceed those of *S. cerevisiae* (e.g. Morton and Potter 2000; Dragosits *et al.* 2011; Tran *et al.* 2017). Similar to *S. cerevisiae*, about 10% of the total genes in the *P. pastoris* genome are predicted to have a function in the secretory pathway, including genes annotated to ER, protein folding, glycosylation, proteolytic processing, ERAD, Golgi, the vacuole, SNAREs and other genes involved in vesicle-mediated transport (Delic *et al.* 2013). An extensive review comparing the canonical protein secretion pathway of seven different yeast species with that of *S. cerevisiae* on a genomic level was published by Delic *et al.* (2013), highlighting some noteworthy differences between the different species. One aspect that is specifically relevant to the production of biopharmaceuticals are differences in N- and O-glycosylation, as both types of PTM have been demonstrated to affect pharmacokinetics and pharmacodynamics of recombinantly produced proteins (De Wachter, Van Landuyt and Callewaert 2018; Zhou and Qiu 2019). N-glycosylation plays a very important role in the folding and quality control (QC) process of glycosylated proteins. The calnexin cycle, a crucial component of this QC process, is characterized in mammals and *S. pombe*. Homologs of the key enzyme UGGT (UDP-glucose:glycoprotein glucosyltransferase) also exist in several *Saccharomycetales* including *K. phaffii* and *Y. lipolytica*, but not in *S. cerevisiae* (Caramelo and Parodi 2008). Recently, it was demonstrated for *S. cerevisiae* that also O-glycosylation plays a decisive role in ER-quality control via a process termed unfolded protein O-mannosylation (UPOM), where proteins are only subjected to O-glycosylation if their correct confirmation is not attained for prolonged time periods. This increases their solubility and eventually leads to their degradation by the proteasome-dependent ERAD-pathway or, after exit from the ER, target them for post-ER degradation (Neubert *et al.* 2016).

Protein N-glycosylation in eukaryotes is initiated in the ER by the linking of the precursor $Glc_3Man_9GlcNAc_2$ to an asparagine residue in the consensus sequence asparagine-X-serine/threonine (Asn-X-Ser/Thr, where X is any amino acid except for proline). Subsequently, the terminal α-1,2 and α-1,3 glucose residues are removed by respective glucosidases and one α-1,2-mannose is removed by an ER-residing α-1,2-mannosidase, resulting in an $Man_8GlcNAc_2$ glycan. Further N-glycan modifications of properly folded proteins take place in the Golgi apparatus, where yeasts add mannose and mannosylphosphate sugars to the $Man_8GlcNAc_2$ glycan core, generating N-glycans of the high-mannose type (Hamilton and Gerngross 2007; De Wachter, Van Landuyt and Callewaert 2018). N-glycan structure and side chain composition can differ substantially between yeast species (Thak *et al.* 2018) but can also be very heterogeneous in regard to a specific N-glycosylation site. *S. cerevisiae* N-glycans typically carry longer mannose outer chains, containing up to 150–200 mannose residues in total, while in *K. phaffii* outer mannose chain length is shorter with a total of 8–18 mannose residues (Herscovics and Orlean 1993; Kang *et al.* 1998; Krainer *et al.* 2013; Thak *et al.* 2018). Different to *S. cerevisiae*, *K. phaffii* glycans do not have immunogenic terminal α-1,3-linked mannose residues, due to the lack of the corresponding Mnn1 enzyme family (Delic *et al.* 2013; Thak *et al.* 2018). Instead, *K. phaffii* possesses Bmt enzymes which catalyse the addition of beta-1,2 mannoses.

Mannose outer chain elongation is initiated by the introduction of an α-1,6-mannose residue by the mannosyltransferase Och1. Disruption of this gene in *K. phaffii* leads to a reduction from 10 to 8 mannose residues in the dominant glycan (Krainer *et al.* 2013) and co-overexpression of a recombinant α-1,2-mannosidase in the ER mainly yields $Man_5GlcNAc_2$ structures (De Wachter, Van Landuyt and Callewaert 2018). In contrast, in *S. cerevisiae* two more enzymes acting on $Man_8GlcNAc_2$ were required to be deleted, which are not present in the *K. phaffii* genome. While it was initially believed that *K. phaffii* lacking Och1 shows only a minor phenotype in contrast to *S. cerevisiae* $\Delta och1$, it was later elucidated that the true *K. phaffii* $\Delta och1$ knockout has a wrinkly morphology and a growth deficit (Krainer *et al.* 2013; De Wachter, Van Landuyt and Callewaert 2018), which was not observed in the initial insertional mutant (Choi *et al.* 2003; Vervecken *et al.* 2004). Nevertheless, in combination with the fact that less knockouts are required, this is probably one explanation why glycoengineering proceeded mainly with *K. phaffii* rather than *S. cerevisiae*. De Wachter, Van Landuyt and Callewaert (2018) comprehensively reviewed recent advances in N- and O-glycoengineering of yeasts. So far, most progress regarding the humanization of N-glycans has been made in *K. phaffii*, allowing even for the production of proteins containing complex-type sialylated N-glycans or mucin-type O-glycans. Nevertheless, to date no biopharmaceuticals produced in glycoengineered yeasts have reached the market.

In yeasts, O-glycosylation is also initiated in the ER by a family of protein O-mannosyltransferases (PMTs) that transfer mannose from dolichol phosphate β-D-mannose (Dol-P-Man) to Ser/Thr residues in the nascent proteins. Linear elongation of the O-glycans takes place in the Golgi, where mannosyltransferases catalyze the transfer of mannose from GDP mannose (Neubert *et al.* 2016). As true for many other genes, the number of PMT genes is reduced to five in *K. phaffii* compared to seven in *S. cerevisiae*. Similar to other yeasts, *K. phaffii PMT1* and *PMT2* appear both to have a predominant role in protein O-glycosylation, with *PMT2* showing the highest expression levels and the $\Delta pmt2$ mutant strain the most severe growth defect (Govindappa *et al.* 2013; Nett *et al.* 2013). Radoman *et al.* (2021) comprehensively analysed the occurrence and composition of

O-glycans of secreted proteins that were produced in *K. phaffii*, and found that the degree of O-mannosylation of a recombinant protein proved to be higher when methanol was used as a carbon source. The majority of O-glycans was composed of one mannose residue, while a maximum of five mannose residues was observed. It is not known, however, which enzymes are responsible for chain elongation, as *K. phaffii* possesses six homologs of the Golgi-resident Ktr/Kre-family mannosyltransferases compared to nine family members in *S. cerevisiae*, making their correct functional assignment as well as their disruption more difficult in *K. phaffii*.

Major differences between *S. cerevisiae* and non-conventional yeasts are seen in the structural organization of the Golgi apparatus, which consists of disk-shaped membranes called cisterna. Golgi resident proteins, mainly involved in N- and O-glycan as well as lipid processing, are spatially organized, depending on their respective function. Freshly synthesized proteins that pass the ER-quality control, exit the ER *via* COPII-coated transport vesicles at so called transitional ER (tER) sites and enter the Golgi at the *cis* site, move through medial cisternae and eventually arrive at the *trans* compartment (Papanikou and Glick 2009; Suda and Nakano 2011). In most eukaryotes, Golgi cisternae form stacks. In vertebrates, the structural organization of the Golgi is even more complex, appearing as a twisted, ribbon-like network (Wei and Seemann 2010). *S. cerevisiae* is one of only a few known eukaryotic organisms where the respective Golgi compartments are not organized in stacks but scattered across the cytoplasm (Mowbrey and Dacks 2009). *K. phaffii* on the other hand, contains 2–5 stacks of ca. 4 cisternae each per cell and shares several other Golgi characteristics with mammalian cells such as the presence of a cisternae-surrounding matrix as well as fenestration and tubular extension of the cisternae (Papanikou and Glick 2009). It, therefore, serves as a model organism for Golgi-related research. Disruption of Golgi stacking in mammalian cells has been demonstrated to reduce total N-linked protein glycosylation and decrease N-glycan complexity. Reversible unstacking of the Golgi in *K. phaffii* was observed in a *SEC16* temperature sensitive mutant (Connerly *et al.* 2005) as well as the KO mutants of *PpLMH1*, encoding a GRIP domain Golgin (Jain, Dahara and Bhattacharyya 2019). Permanent unstacking under physiological conditions was achieved by the disruption of the genes *RSN1*, *CSC1-2* or *TVP18*, which are thought to be involved in calcium transport or signaling. However, unstacking of the Golgi by disruption of these genes did not lead to significant changes in the cellular N-glycome nor in N-glycan abundance or composition of glycoGFP (Aw *et al.* 2021).

One of the major advantages attributed to *K. phaffii* is that it secretes comparatively few host cell proteins, facilitating down-stream processing, and thereby reducing production costs. According to Lum and Min (Lum and Min 2011) the predicted secretome size of *S. cerevisiae* is 50%, and that of *Y. lipolytica*, another popular host for recombinant protein production, is even 300% larger than that of *K. phaffii*. However, these numbers do not account for growth conditions or for proteins regularly observed extracellularly such as metabolic enzymes, chaperones or proteins involved in translation that are not actively secreted but reach the extracellular space either by passive transport, unconventional secretion or cell lysis (Nombela, Gil and Chaffin 2006; Miura and Ueda 2018). Several studies have been conducted in order to characterize the secretome of *K. phaffii* under industrially relevant conditions (Dragosits *et al.* 2009; Huang *et al.* 2011; Burgard *et al.* 2020). The most recent study carried out by Burgard *et al.* (2020) investigated how the secretome of *K. phaffii* changes throughout a typical recombinant protein production process and how the choice of carbon source (glucose or glycerol/methanol) affects native protein secretion. In total 51 proteins were identified, concordant with previous observations. A similar study conducted in *K. lactis*, whose predicted secretome size varies between 113 and 178 proteins (Swaim *et al.* 2008; Lum and Min 2011), found up to 120 proteins when cells were grown on galactose and a total of 151 proteins across all growth conditions. In both, *K. lactis* and *K. phaffii*, a majority of identified proteins possessed a predicted signal peptide. While the core secretome (proteins identified in every condition) of *K. phaffii* mainly consisted of cell wall proteins, in *K. lactis* also many proteins with a function in glycosylation, carbohydrate metabolism and proteolysis were identified. To the authors knowledge no studies investigating the full secretome of other important yeast recombinant protein production hosts like *S. cerevisiae*, *H. polymorpha* or *Y. lipolytica* under industrially relevant conditions have been published so far, hindering relevant comparisons to these organisms.

CONCLUSIONS

Genetic diversity is large among budding yeasts. Their morphological similarity should not make us believe that they function very similarly. Partly, differences between *S. cerevisiae* and non-conventional yeasts can be explained by the WGD event and subsequent functional specialization that *S. cerevisiae* and its close relatives went through. Non-conventional species such as *K. phaffii* have rather adapted their proteome to the ecological niches they habitate. But even where gene sets are similar, different transcriptional control creates multitute. Such regulatory differentiations have recently been identified in carbon metabolism and contribute to the different physiology of *K. phaffii* in comparison to *S. cerevisiae*. Also, differential localization of proteins was observed. It remains to be identified in future if such regulatory differences also contribute to the higher secretion efficiency and different glycan pattern that make *K. phaffii* an excellent protein production host. Overall, we strongly recommend being careful when drawing analogies solely based on sequence analysis.

FUNDING INFORMATION

BG, OA and DM acknowledge funding by the Austrian BMK, BMDW, SFG, Standortagentur Tirol, Government of Lower Austria, Vienna Business Agency and BOKU through the FFG—COMET Funding Program. OA acknowledges funding by the Austrian Science Fund FWF (grant number M2891). BG, LH, DM and CR acknowledge funding by the Austrian BMDW and Nationalstiftung FTE through the Christian Doppler Research Association. This article is based upon work from the Yeast4Bio COST Action 18229, supported by COST (European Cooperation in Science and Technology).

AUTHOR CONTRIBUTIONS

All authors contributed equally to the manuscript. OA and LH prepared the final figures. All authors have read and agreed to the final version of the manuscript.

Conflicts of Interest. The authors declare no conflicts of interest.

REFERENCES

Adivitiya B, Mohanty S, Khasa Y. Nitrogen supplementation ameliorates product quality and quantity during high cell density bioreactor studies of *Pichia pastoris*: a case study with proteolysis prone streptokinase. *Int J Biol Macromol* 2021;**180**:760–70.

Askew C, Sellam A, Epp E *et al*. Transcriptional regulation of carbohydrate metabolism in the human pathogen *Candida albicans*. *PLoS Pathog* 2009;**5**:e1000612.

Ata Ö, Rebnegger C, Tatto NE *et al*. A single Gal4-like transcription factor activates the Crabtree effect in *Komagataella phaffii*. *Nat Commun* 2018;**9**:4911.

Aw R, De Wachter C, Laukens B *et al*. Knockout of *RSN1*, *TVP18* or *CSC1-2* causes perturbation of Golgi cisternae in *Pichia pastoris*. *Traffic* 2021;**22**:48–63.

Barbay D, Mačáková M, Sützl L *et al*. Two homologs of the Cat8 transcription factor are involved in the regulation of ethanol utilization in *Komagataella phaffii*. *Curr Genet* 2021;**67**:641–61.

Baumann K, Carnicer M, Dragosits M *et al*. A multi-level study of recombinant *Pichia pastoris* in different oxygen conditions. *BMC Syst Biol* 2010;**4**:141.

Bernauer L, Radkohl A, Lehmayer LGK *et al*. *Komagataella phaffii* as emerging model organism in fundamental research. *Front Microbiol* 2020;**11**:607028.

Bianchi F, Van't Klooster J, Ruiz S *et al*. Regulation of amino acid transport in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* 2019;**83**:e00024–00019.

Boles E, Hollenberg CP. The molecular genetics of hexose transport in yeasts. *FEMS Microbiol Rev* 1997;**21**:85–111.

Booth LN, Tuch BB, Johnson AD. Intercalation of a new tier of transcription regulation into an ancient circuit. *Nature* 2010;**468**:959–63.

Brückner S, Mösch HU. Choosing the right lifestyle: adhesion and development in *Saccharomyces cerevisiae*. *FEMS Microbiol Rev* 2012;**36**:25–58.

Brückner S, Schubert R, Kraushaar T *et al*. Kin discrimination in social yeast is mediated by cell surface receptors of the Flo11 adhesin family. *eLife* 2020;**9**:e55587.

Burgard J, Grünwald-Gruber C, Altmann F *et al*. The secretome of *Pichia pastoris* in fed-batch cultivations is largely independent of the carbon source but changes quantitatively over cultivation time. *Microb Biotechnol* 2020;**13**:479–94.

Caramelo J, Parodi A. Getting in and out from calnexin/calreticulin cycles. *J Biol Chem* 2008;**283**:10221–5.

Chatterjee G, Sankaranarayanan S, Guin K *et al*. Repeat-associated fission yeast-like regional centromeres in the ascomycetous budding yeast *Candida tropicalis*. *PLos Genet* 2016;**12**:e1005839.

Chen MT, Lin S, Shandil I *et al*. Generation of diploid *Pichia pastoris* strains by mating and their application for recombinant protein production. *Microb Cell Fact* 2012;**11**:91.

Choi BK, Bobrowicz P, Davidson RC *et al*. Use of combinatorial genetic libraries to humanize N-linked glycosylation in the yeast *Pichia pastoris*. *Proc Natl Acad Sci* 2003;**100**:5022–7.

Choudhury BI, Whiteway M. Evolutionary transition of GAL regulatory circuit from generalist to specialist function in ascomycetes. *Trends Microbiol* 2018;**26**:692–702.

Connerly P, Esaki M, Montegna E *et al*. Sec16 is a determinant of transitional ER organization. *Curr Biol* 2005;**15**:1439–47.

Coughlan AY, Hanson SJ, Byrne KP *et al*. Centromeres of the yeast *Komagataella phaffii* (*Pichia pastoris*) have a simple inverted-repeat structure. *Genome Biol Evolut* 2016;**8**:2482–92.

Cullen PJ, Sprague GF, Jr. The regulation of filamentous growth in yeast. *Genetics* 2012;**190**:23–49.

De S, Rebnegger C, Moser J *et al*. Pseudohyphal differentiation in *Komagataella phaffii*: investigating the FLO gene family. *FEMS Yeast Res* 2020;**20**:foaa044.

De S. Epigenetic Regulationin *Komagataella phaffii: Investigating Pseudohyphal Growth and Flocculation*., BOKU-Universität für Bodenkultur, Vienna, 2019.

De Wachter C, Van Landuyt L, Callewaert N. Engineering of yeast glycoprotein expression. *Adv Biochem Eng Biotechnol* 2018;**175**:93–135.

Delic M, Valli M, Graf AB *et al*. The secretory pathway: exploring yeast diversity. *FEMS Microbiol Rev* 2013;**37**:872–914.

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

Dragosits M, Frascotti G, Bernard-Granger L *et al*. Influence of growth temperature on the production of antibody Fab fragments in different microbes: a host comparative analysis. *Biotechnol Prog* 2011;**27**:38–46.

Dragosits M, Stadlmann J, Albiol J *et al*. The effect of temperature on the proteome of recombinant *Pichia pastoris*. *J Proteome Res* 2009;**8**:1380–92.

Duman-Özdamar ZE, Binay B. Production of industrial enzymes via *Pichia pastoris* as a cell factory in bioreactor: current status and future aspects. *Protein J* 2021;**40**:367–76.

Elbing K, Larsson C, Bill R *et al*. Role of hexose transport in control of glycolytic flux in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2004;**70**:5323–30.

Ergün BG, Demir İ, Özdamar T *et al*. Engineered deregulation of expression in yeast with designed hybrid-promoter architectures in coordination with discovered master regulator transcription factor. *Adv Biosyst* 2020;**4**:e1900172.

Ergün BG, Gasser B, Mattanovich D *et al*. Engineering of alcohol dehydrogenase 2 hybrid-promoter architectures in *Pichia pastoris* to enhance recombinant protein expression on ethanol. *Biotechnol Bioeng* 2019a;**116**:2674–86.

Ergün BG, Hüccetoğulları D, Öztürk S *et al*. Established and upcoming yeast expression systems. *Methods Mol Biol* 2019b;**1923**:1–74.

Ergün BG. *Transcriptional engineering of Pichia pastoris alcohol dehydrogenase 2 and alcohol oxidase 1 promoters for recombinant protein production*, PhD thesis, Middle East Technical University, Ankara, 2018.

Espinosa M, Gonzalez-Garcia R, Valgepea K *et al*. Adaptive laboratory evolution of native methanol assimilation in *Saccharomyces cerevisiae*. *Nat Commun* 2020;**11**:5564.

Feng D, Stoyanov A, Olliff J *et al*. Carbon source requirements for mating and mating-type switching in the methylotrophic yeasts *Ogataea* (*Hansenula*) *polymorpha* and *Komagataella phaffii* (*Pichia pastoris*). *Yeast* 2020;**37**:237–45.

Förster J, Halbfeld C, Zimmermann M *et al*. A blueprint of the amino acid biosynthesis network of hemiascomycetes. *FEMS Yeast Res* 2014;**14**:1090–100.

Freese S, Vogts T, Speer F *et al*. C- and N-catabolic utilization of tricarboxylic acid cycle-related amino acids by *Scheffersomyces stipitis* and other yeasts. *Yeast* 2011;**28**:375–90.

Gao J, Jiang L, Lian J. Development of synthetic biology tools to engineer *Pichia pastoris* as a chassis for the production of natural products. *Synth Syst Biotechnol* 2021;**6**:110–9.

Gasser B, Mattanovich D, Buchetics M. *Recombinant Host Cell Engineered to Overexpress Helper Proteins*: Boehringer Ingelheim RCV GmbH & Co KG; Sandoz AG ; VTU Technology GmbH;

Lonza Ltd; Universitat fur Bodenkultur Wien (BOKU); Biomin Holding GmbH, 2014.
Gombert A, Moreira dos Santos M, Christensen B *et al.* Network identification and flux quantification in the central metabolism of *Saccharomyces cerevisiae* under different conditions of glucose repression. *J Bacteriol* 2001;**183**:1441–51.
Goossens K, Ielasi F, Nookaew I *et al.* Molecular mechanism of flocculation self-recognition in yeast and its role in mating and survival. *mBio* 2015;**6**:e00427–00415.
Goossens KV, Willaert RG. The N-terminal domain of the Flo11 protein from *Saccharomyces cerevisiae* is an adhesin without mannose-binding activity. *FEMS Yeast Res* 2012;**12**:78–87.
Govindappa N, Hanumanthappa M, Venkatarangaiah K *et al.* PMT1 gene plays a major role in O-mannosylation of insulin precursor in *Pichia pastoris*. *Protein Expr Purif* 2013;**88**:164–71.
Guo B, Styles C, Feng Q *et al.* A *Saccharomyces* gene family involved in invasive growth, cell-cell adhesion, and mating. *Proc Natl Acad Sci* 2000;**97**:12158–63.
Guo C, Huang Y, Zheng H *et al.* Secretion and activity of antimicrobial peptide cecropin D expressed in *Pichia pastoris*. *Exp Therap Med* 2012;**4**:1063–8.
Haber JE. Mating-type genes and MAT switching in *Saccharomyces cerevisiae*. *Genetics* 2012;**191**:33–64.
Hamilton S, Gerngross T. Glycosylation engineering in yeast: the advent of fully humanized yeast. *Curr Opin Biotechnol* 2007;**18**:387–92.
Hanson SJ, Byrne KP, Wolfe KH. Flip/flop mating-type switching in the methylotrophic yeast *Ogataea polymorpha* is regulated by an Efg1-Rme1-Ste12 pathway. *PLos Genet* 2017;**13**:e1007092.
Hanson SJ, Byrne KP, Wolfe KH. Mating-type switching by chromosomal inversion in methylotrophic yeasts suggests an origin for the three-locus *Saccharomyces cerevisiae* system. *Proc Natl Acad Sci* 2014;**111**:E4851–4858.
Hanson SJ, Wolfe KH. An evolutionary perspective on yeast mating-type switching. *Genetics* 2017;**206**:9–32.
Hartner FS, Glieder A. Regulation of methanol utilisation pathway genes in yeasts. *Microb Cell Fact* 2006;**5**:39.
Hegemann J, Fleig U. The centromere of budding yeast. *Bioessays* 1993;**15**:451–60.
Heistinger L, Gasser B, Mattanovich D. Creation of stable heterothallic strains of *Komagataella phaffii* enables dissection of mating gene regulation. *Mol Cell Biol* 2018;**38**:e00398–17.
Heistinger L, Gasser B, Mattanovich D. Microbe profile: *K omagataella phaffii*: a methanol devouring biotech yeast formerly known as *Pichia pastoris*. *Microbiology* 2020;**166**:614–6.
Herscovics A, Orlean P. Glycoprotein biosynthesis in yeast. *FASEB J* 1993;**7**:540–50.
Huang CJ, Damasceno LM, Anderson KA *et al.* A proteomic analysis of the *Pichia pastoris* secretome in methanol-induced cultures. *Appl Microbiol Biotechnol* 2011;**90**:235–47.
Ira G, Satory D, Haber JE. Conservative inheritance of newly synthesized DNA in double-strand break-induced gene conversion. *Mol Cell Biol* 2006;**26**:9424–9.
Ito T, Fujimura S, Uchino M *et al.* Distribution, diversity and regulation of alcohol oxidase isozymes, and phylogenetic relationships of methylotrophic yeasts. *Yeast* 2007;**24**:523–32.
Jahic M, Rotticci-Mulder J, Martinelle M *et al.* Modeling of growth and energy metabolism of *Pichia pastoris* producing a fusion protein. *Bioprocess Biosys Eng* 2002;**24**:385–93.
Jain B, Dahara R, Bhattacharyya D. The golgin *Pp*Imh1 mediates reversible cisternal stacking in the Golgi of the budding yeast *Pichia pastoris*. *J Cell Sci* 2019;**132**:jcs230672.
Kang H, Sohn J, Choi E *et al.* Glycosylation of human alpha 1-antitrypsin in *Saccharomyces cerevisiae* and methylotrophic yeasts. *Yeast* 1998;**14**:371–81.
Karbalaei M, Rezaee SA, Farsiani H. *Pichia pastoris*: a highly successful expression system for optimal synthesis of heterologous proteins. *J Cell Physiol* 2020;**235**:5867–81.
Kato N, Higuchi T, Sakazawa C *et al.* Purification and properties of a transketolase responsible for formaldehyde fixation in a methanol-utilizing yeast, *Candida boidinii* (*Kloeckera* sp.) No. 2201. *Biochim Biophys Acta* 1982;**715**:143–50.
Kim HY, Lee SB, Kang HS *et al.* Two distinct domains of Flo8 activator mediates its role in transcriptional activation and the physical interaction with Mss11. *Biochem Biophys Res Commun* 2014;**449**:202–7.
Klein M, Swinnen S, Thevelein JM *et al.* Glycerol metabolism and transport in yeast and fungi: established knowledge and ambiguities. *Environ Microbiol* 2017;**19**:878–93.
Kock M, Brückner S, Wozniak N *et al.* Structural and functional characterization of PA14/Flo5-Like adhesins from *Komagataella pastoris*. *Front Microbiol* 2018;**9**:2581.
Koivistoinen O, Arvas M, Headman J *et al.* Characterisation of the gene cluster for l-rhamnose catabolism in the yeast *Scheffersomyces* (*Pichia*) *stipitis*. *Gene* 2012;**492**:177–85.
Krainer FW, Gmeiner C, Neutsch L *et al.* Knockout of an endogenous mannosyltransferase increases the homogeneity of glycoproteins produced in *Pichia pastoris*. *Sci Rep* 2013;**3**:3279.
Krassowski T, Kominek J, Shen X *et al.* Multiple reinventions of mating-type switching during budding yeast evolution. *Curr Biol* 2019;**29**:2555–62. e2558.
Kucha JA, Dooley DM. Cloning, sequence analysis, and characterization of the 'lysyl oxidase' from *Pichia pastoris*. *J Inorg Biochem* 2001;**83**:193–204.
Kumar NV, Rangarajan PN. The zinc finger proteins Mxr1p and repressor of phosphoenolpyruvate carboxykinase (ROP) have the same DNA binding specificity but regulate methanol metabolism antagonistically in *Pichia pastoris*. *J Biol Chem* 2012;**287**:34465–73.
Kurtzman CP, Fell JW, Boekhout T. *The Yeasts: A Taxonomic Study*. Elsevier, 2011.
Lages F, Silva-Graça M, Lucas C. Active glycerol uptake is a mechanism underlying halotolerance in yeasts: a study of 42 species. *Microbiology* 1999;**145**:2577–85.
League G, Slot J, Rokas A. The *ASP3* locus in *Saccharomyces cerevisiae* originated by horizontal gene transfer from *Wickerhamomyces*. *FEMS Yeast Res* 2012;**12**:859–63.
Li X, Yang Y, Zhan C *et al.* Transcriptional analysis of impacts of glycerol transporter 1 on methanol and glycerol metabolism in *Pichia pastoris*. *FEMS Yeast Res* 2018;**18**:fox081.
Lin-Cereghino GP, Godfrey L, de la Cruz BJ *et al.* Mxr1p, a key regulator of the methanol utilization pathway and peroxisomal genes in *Pichia pastoris*. *Mol Cell Biol* 2006;**26**:883–97.
Linder T. A genomic survey of nitrogen assimilation pathways in budding yeasts (sub-phylum Saccharomycotina). *Yeast* 2019;**36**:259–73.
Liti G, Louis EJ. Advances in quantitative trait analysis in yeast. *PLos Genet* 2012;**8**:e1002912.
Liu B, Zhang Y, Zhang X *et al.* Discovery of a rhamnose utilization pathway and rhamnose-inducible promoters in *Pichia pastoris*. *Sci Rep* 2016;**6**:27352.
Liu H, Styles CA, Fink GR. *Saccharomyces cerevisiae* S288C has a mutation in *FLO8*, a gene required for filamentous growth. *Genetics* 1996;**144**:967–78.
Lum G, Min XJ. FunSecKB: the Fungal Secretome KnowledgeBase. *Database* 2011;**2011**:bar001.

Maaheimo H, Fiaux J, Cakar ZP *et al.* Central carbon metabolism of *Saccharomyces cerevisiae* explored by biosynthetic fractional (13)C labeling of common amino acids. *Eur J Biochem* 2001;**268**:2464–79.

Maekawa H, Kaneko Y. Inversion of the chromosomal region between two mating type loci switches the mating type in *Hansenula polymorpha*. *PLos Genet* 2014;**10**:e1004796.

Magasanik B. Ammonia assimilation by *Saccharomyces cerevisiae*. *Eukar Cell* 2003;**2**:827–9.

Marini A, Soussi-Boudekou S, Vissers S *et al.* A family of ammonium transporters in *Saccharomyces cerevisiae*. *Mol Cell Biol* 1997;**17**:4282–93.

Martchenko M, Levitin A, Whiteway M. Transcriptional activation domains of the *Candida albicans* Gcn4p and Gal4p homologs. *Eukar Cell* 2007;**6**:291–301.

Mattanovich D, Graf A, Stadlmann J *et al.* Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. *Microb Cell Fact* 2009;**8**:29.

Maurer M, Kühleitner M, Gasser B *et al.* Versatile modeling and optimization of fed batch processes for the production of secreted heterologous proteins with *Pichia pastoris*. *Microb Cell Fact* 2006;**5**:37.

Mbawala A, al Mahmood S, Loppinet V *et al.* Acetolysis and 1H NMR studies on mannans isolated from very flocculent and weakly flocculent cells of *Pichia pastoris* IFP 206. *J Gen Microbiol* 1990;**136**:1279–84.

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

Merlini L, Dudin O, Martin SG. Mate and fuse: how yeast cells do it. *Open Biol* 2013;**3**:130008.

Miura N, Ueda M. Evaluation of unconventional protein secretion by *Saccharomyces cerevisiae* and other fungi. *Cells* 2018;**7**:128.

Morton CL, Potter PM. Comparison of *Escherichia coli*, *Saccharomyces cerevisiae*, *Pichia pastoris*, *Spodoptera frugiperda*, and COS7 cells for recombinant gene expression. Application to a rabbit liver carboxylesterase. *Mol Biotechnol* 2000;**16**:193–202.

Mowbrey K, Dacks JB. Evolution and diversity of the Golgi body. *FEBS Lett* 2009;**583**:3738–45.

Nakamura Y, Nishi T, Noguchi R *et al.* A stable, autonomously replicating plasmid vector containing *Pichia pastoris* centromeric DNA. *Appl Environ Microbiol* 2018;**84**:e02882–02817.

Navarathna D, Harris S, Roberts D *et al.* Evolutionary aspects of urea utilization by fungi. *FEMS Yeast Res* 2010;**10**:209–13.

Nett JH, Cook WJ, Chen MT *et al.* Characterization of the *Pichia pastoris* protein-O-mannosyltransferase gene family. *PLoS ONE* 2013;**8**:e68325.

Neubert P, Halim A, Zauser M *et al.* Mapping the O-mannose glycoproteome in *Saccharomyces cerevisiae*. *Mol Cell Proteomics* 2016;**15**:1323–37.

Nocon J, Steiger M, Mairinger T *et al.* Increasing pentose phosphate pathway flux enhances recombinant protein production in *Pichia pastoris*. *Appl Microbiol Biotechnol* 2016;**100**:5955–63.

Nombela C, Gil C, Chaffin WL. Non-conventional protein secretion in yeast. *Trends Microbiol* 2006;**14**:15–21.

Otterstedt K, Larsson C, Bill RM *et al.* Switching the mode of metabolism in the yeast *Saccharomyces cerevisiae*. *EMBO Rep* 2004;**5**:532–7.

Ozimek P, Veenhuis M, van der Klei IJ. Alcohol oxidase: a complex peroxisomal, oligomeric flavoprotein. *FEMS Yeast Res* 2005;**5**:975–83.

Papanikou E, Glick BS. The yeast golgi apparatus: insights and mysteries. *FEBS Lett* 2009;**583**:3746–51.

Parua PK, Ryan PM, Trang K *et al.* *Pichia pastoris* 14-3-3 regulates transcriptional activity of the methanol inducible transcription factor Mxr1 by direct interaction. *Mol Microbiol* 2012;**85**:282–98.

Pekarsky A, Mihalyi S, Weiss M *et al.* Depletion of boric acid and cobalt from cultivation media: impact on recombinant protein production with *Komagataella phaffii*. *Bioeng* 2020;**7**:161.

Peña DA, Gasser B, Zanghellini J *et al.* Metabolic engineering of *Pichia pastoris*. *Metab Eng* 2018;**50**:2–15.

Piva L, De Marco J, Moraes L *et al.* Construction and characterization of centromeric plasmids for *Komagataella phaffii* using a color-based plasmid stability assay. *PLoS ONE* 2020;**15**:e0235532.

Radoman B, Grünwald-Gruber C, Schmelzer B *et al.* The degree and length of O-glycosylation of recombinant proteins produced in *Pichia pastoris* depends on the nature of the protein and the process type. *Biotechnol J* 2021;**16**:e2000266.

Raschke WC, Neiditch BR, Hendricks M *et al.* Inducible expression of a heterologous protein in *Hansenula polymorpha* using the alcohol oxidase 1 promoter of *Pichia pastoris*. *Gene* 1996;**177**:163–7.

Rebnegger C, Graf AB, Valli M *et al.* In *Pichia pastoris*, growth rate regulates protein synthesis and secretion, mating and stress response. *Biotechnol J* 2014;**9**:511–25.

Rebnegger C, Vos T, Graf AB *et al.* *Pichia pastoris* exhibits high viability and a low maintenance energy requirement at near-zero specific growth rates. *Appl Environ Microbiol* 2016;**82**:4570–83.

Riley R, Haridas S, Wolfe KH *et al.* Comparative genomics of biotechnologically important yeasts. *Proc Natl Acad Sci* 2016;**113**:9882–7.

Rußmayer H, Buchetics M, Gruber C *et al.* Systems-level organization of yeast methylotrophic lifestyle. *BMC Biol* 2015;**13**:80.

Ruth C, Buchetics M, Vidimce V *et al.* *Pichia pastoris* Aft1 - a novel transcription factor, enhancing recombinant protein secretion. *Microb Cell Fact* 2014;**13**:120.

Sahu U, Krishna Rao K, Rangarajan PN. Trm1p, a Zn(II)(2)Cys(6)-type transcription factor, is essential for the transcriptional activation of genes of methanol utilization pathway, in *Pichia pastoris*. *Biochem Biophys Res Commun* 2014;**451**: 158–64.

Sahu U, Rangarajan P. Methanol expression regulator 1 (Mxr1p) is essential for the utilization of amino acids as the sole source of carbon by the methylotrophic yeast, *Pichia pastoris*. *J Biol Chem* 2016;**291**:20588–601.

Sasano Y, Yurimoto H, Kuriyama M *et al.* Trm2p-dependent derepression is essential for methanol-specific gene activation in the methylotrophic yeast *Candida boidinii*. *FEMS Yeast Res* 2010;**10**:535–44.

Sasano Y, Yurimoto H, Yanaka M *et al.* Trm1p, a Zn(II)2Cys6-type transcription factor, is a master regulator of methanol-specific gene activation in the methylotrophic yeast *Candida boidinii*. *Eukaryotic Cell* 2008;**7**:527–36.

Schüller HJ. Transcriptional control of nonfermentative metabolism in the yeast *Saccharomyces cerevisiae*. *Curr Genet* 2003;**43**:139–60.

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

Sherwood RI, Hashimoto T, O'Donnell CW *et al.* Discovery of directional and nondirectional pioneer transcription factors by modeling DNase profile magnitude and shape. *Nat Biotechnol* 2014;**32**:171–8.

Shi L, Wang X, Wang J *et al.* Transcriptome analysis of Δ*mig1*Δ*mig2* mutant reveals their roles in methanol catabolism, peroxisome biogenesis and autophagy in methylotrophic yeast *Pichia pastoris*. *Genes Genomics* 2018;**40**: 399–412.

Siverio J. Assimilation of nitrate by yeasts. *FEMS Microbiol Rev* 2002;**26**:277–84.

Soares EV. Flocculation in *Saccharomyces cerevisiae*: a review. *J Appl Microbiol* 2011;**110**:1–18.

Song Q, Kumar A. An overview of autophagy and yeast pseudohyphal growth: integration of signaling pathways during nitrogen stress. *Cells* 2012;**1**:263–83.

Sorrells TR, Booth LN, Tuch BB *et al.* Intersecting transcription networks constrain gene regulatory evolution. *Nature* 2015;**523**:361–5.

Sreekrishna K, Brankamp RG, Kropp KE *et al.* Strategies for optimal synthesis and secretion of heterologous proteins in the methylotrophic yeast *Pichia pastoris*. *Gene* 1997;**190**:55–62.

Stadlmayr G, Benakovitsch K, Gasser B *et al.* Genome-scale analysis of library sorting (GALibSo): isolation of secretion enhancing factors for recombinant protein production in *Pichia pastoris*. *Biotechnol Bioeng* 2010;**105**:543–55.

Strathern JN, Klar AJ, Hicks JB *et al.* Homothallic switching of yeast mating type cassettes is initiated by a double-stranded cut in the MAT locus. *Cell* 1982;**31**:183–92.

Sturmberger L, Chappell T, Geier M *et al.* Refined *Pichia pastoris* reference genome sequence. *J Biotechnol*. 2016;**235**:121–31.

Su C, Li Y, Lu Y *et al.* Mss11, a transcriptional activator, is required for hyphal development in *Candida albicans*. *Eukar Cell* 2009;**8**:1780–91.

Suda Y, Nakano A. The yeast golgi apparatus. *Traffic* 2011;**13**: 505–10.

Swaim CL, Anton BP, Sharma SS *et al.* Physical and computational analysis of the yeast *Kluyveromyces lactis* secreted proteome. *Proteomics* 2008;**8**:2714–23.

Swinnen S, Thevelein JM, Nevoigt E. Genetic mapping of quantitative phenotypic traits in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2012;**12**:215–27.

Tanneberger K, Kirchberger J, Bar J *et al.* A novel form of 6-phosphofructokinase. Identification and functional relevance of a third type of subunit in *Pichia pastoris*. *J Biol Chem* 2007;**282**:23687–97.

Thak E, Kim J, Lee D *et al.* Structural analysis of N-/O-glycans assembled on proteins in yeasts. *J Microbiol* 2018;**56**:11–23.

Tomàs-Gamisans M, Ødum A, Workman M *et al.* Glycerol metabolism of *Pichia pastoris* (*Komagataella* spp.) characterised by 13 C-based metabolic flux analysis. *New Biotechnol* 2019;**50**:52–59.

Tran AM, Nguyen TT, Nguyen CT *et al.* *Pichia pastoris* versus *Saccharomyces cerevisiae*: a case study on the recombinant production of human granulocyte-macrophage colony-stimulating factor. *BMC Res Notes* 2017;**10**:148.

Trotter P, Juco K, Le H *et al.* Glutamate dehydrogenases in the oleaginous yeast *Yarrowia lipolytica*. *Yeast* 2020;**37**:103–15.

Tsong AE, Tuch BB, Li H *et al.* Evolution of alternative transcriptional circuits with identical logic. *Nature* 2006;**443**:415–20.

Unkles S, Wang R, Wang Y *et al.* Nitrate reductase activity is required for nitrate uptake into fungal but not plant cells. *J Biol Chem* 2004;**279**:28182–6.

Valli M, Grillitsch K, Grünwald-Gruber C *et al.* A subcellular proteome atlas of the yeast *Komagataella phaffii*. *FEMS Yeast Res* 2020;**20**:foaa001.

Valli M, Tatto NE, Peymann A *et al.* Curation of the genome annotation of *Pichia pastoris* (*Komagataella phaffii*) CBS7435 from gene level to protein function. *FEMS Yeast Res* 2016;**16**:fow051.

Van Mulders S, Christianen E, Saerens S *et al.* Phenotypic diversity of Flo protein family-mediated adhesion in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2009;**9**:178–90.

Velagapudi VR, Wittmann C, Schneider K *et al.* Metabolic flux screening of *Saccharomyces cerevisiae* single knockout strains on glucose and galactose supports elucidation of gene function. *J Biotechnol* 2007;**132**:395–404.

Velastegui E, Theron C, Berrios J *et al.* Downregulation by organic nitrogen of *AOX1* promoter used for controlled expression of foreign genes in the yeast *Pichia pastoris*. *Yeast* 2019;**36**:297–304.

Verstrepen KJ, Klis FM. Flocculation, adhesion and biofilm formation in yeasts. *Mol Microbiol* 2006;**60**:5–15.

Vervecken W, Kaigorodov V, Callewaert N *et al.* In vivo synthesis of mammalian-like, hybrid-type N-glycans in *Pichia pastoris*. *Appl Environ Microbiol* 2004;**70**:2639–46.

Vonck J, Parcej DN, Mills DJ. Structure of alcohol oxidase from *Pichia pastoris* by cryo-electron microscopy. *PLoS ONE* 2016;**11**:e0159476.

Wang J, Wang X, Shi L *et al.* Methanol-independent protein expression by *AOX1* promoter with trans-acting elements engineering and glucose-glycerol-shift induction in *Pichia pastoris*. *Sci Rep* 2017;**7**:41850.

Wang X, Cai M, Shi L *et al.* PpNrg1 is a transcriptional repressor for glucose and glycerol repression of *AOX1* promoter in methylotrophic yeast *Pichia pastoris*. *Biotechnol Lett* 2016a;**38**:291–8.

Wang X, Wang Q, Wang J *et al.* Mit1 transcription factor mediates methanol signaling and regulates the alcohol oxidase 1 (*AOX1*) promoter in *Pichia pastoris*. *J Biol Chem* 2016b;**291**:6245–61.

Wei JH, Seemann J. Unraveling the Golgi ribbon. *Traffic* 2010;**11**:1391–400.

Werten MW, Eggink G, Cohen Stuart MA *et al.* Production of protein-based polymers in *Pichia pastoris*. *Biotechnol Adv* 2019;**37**:642–66.

Willaert RG. Adhesins of yeasts: protein structure and interactions. *J Fungi* 2018;**4**:119.

Yamamoto K, Tran TNM, Takegawa K *et al.* Regulation of mating type switching by the mating type genes and *RME1* in *Ogataea polymorpha*. *Sci Rep* 2017;**7**:16318.

Yan C, Xu X, Zhang X *et al.* Decreased rhamnose metabolic flux improved production of target proteins and cell flocculation in *Pichia pastoris*. *Front Microbiol* 2018;**9**:1771.

Young ET, Dombek KM, Tachibana C *et al.* Multiple pathways are co-regulated by the protein kinase Snf1 and the transcription factors Adr1 and Cat8. *J Biol Chem* 2003;**278**: 26146–58.

Zhan C, Wang S, Sun Y *et al.* The *Pichia pastoris* transmembrane protein GT1 is a glycerol transporter and relieves the repression of glycerol on *AOX1* expression. *FEMS Yeast Res* 2016;**16**:fow033.

Zhan C, Yang Y, Zhang Z *et al.* Transcription factor Mxr1 promotes the expression of Aox1 by repressing glycerol transporter 1 in *Pichia pastoris*. *FEMS Yeast Res* 2017;**17**. DOI: 10.1093/femsyr/fox015.

Zhou H, Winston F. *NRG1* is required for glucose repression of the *SUC2* and *GAL* genes of *Saccharomyces cerevisiae*. *BMC Genet* 2001;**2**:5.

Zhou Q, Qiu H. The mechanistic impact of N-glycosylation on stability, pharmacokinetics, and immunogenicity of therapeutic proteins. *J Pharm Sci* 2019;**108**: 1366–77.

Zhu J, Gong R, Zhu Q *et al.* Genome-wide determination of gene essentiality by transposon insertion sequencing in yeast *Pichia pastoris*. *Sci Rep* 2018;**8**:10223.

Zhu T, Sun H, Wang M *et al.* *Pichia pastoris* as a versatile cell factory for the production of industrial enzymes and chemicals: current status and future perspectives. *Biotechnol J* 2019;**14**:e1800694.