Carneiro *et al. AMB Express* (2024) 14:131
https://doi.org/10.1186/s13568-024-01795-0

AMB Express

ORIGINAL ARTICLE Open Access

Engineering *Komagataella phaffii* for ethylene glycol production from xylose

Clara Vida G. C. Carneiro[1,2], Débora Trichez[2], Jessica C. Bergmann[2], Viviane Castelo Branco Reis[2], Nils Wagner[3], Thomas Walther[3] and João Ricardo Moreira de Almeida[1,2*]

Abstract
Ethylene glycol (EG) is a versatile molecule produced in the petrochemical industry and is widely used to manufacture plastic polymers, anti-freeze, and automotive fluids. Biotechnological production of EG from xylose, a pentose present in lignocellulose biomass hydrolysates, has been achieved by the engineering of bacteria, such as *Escherichia coli* and *Enterobacter cloacae*, and the yeast *Saccharomyces cerevisiae* with synthetic pathways. In the present work, the Dahms pathway was employed to construct *Komagataella phaffii* strains capable of producing EG from xylose. Different combinations of the four enzymes that compose the synthetic pathway, namely, xylose dehydrogenase, xylonate dehydratase, dehydro-deoxy-xylonate aldolase, and glycolaldehyde reductase, were successfully expressed in *K. phaffii*. Increased production of EG (1.31 g/L) was achieved by employing a newly identified xylonate dehydratase (*xylD-HL*). This xylonate dehydratase allowed 30% higher EG production than a previously known xylonate dehydratase (*xylD-CC*). Further strain engineering demonstrated that *K. phaffii* possesses native glycolaldehyde reduction and oxidation activities, which lead to pathway deviation from EG to glycolic acid (GA) production. Finally, cultivation conditions that favor the production of EG over GA were determined.

Keywords Ethylene glycol, Xylose, Synthetic biology, Metabolic engineering, *K. Phaffii*

Introduction

Ethylene glycol (EG) is a 2-carbon alcohol primarily employed in the production of plastic polymers such as polyethylene terephthalate (PET), polyurethane (PU), and polyethylene succinate (PE) resins (Zhang et al. 2020). In addition, it has several industrial applications, such as producing antifreeze, hydraulic fluids, and surfactants. EG is one of the most promising building block compounds that can be produced from lignocellulosic biomass (de Jong et al. 2020). Lignocellulosic biomass use for bio-based products could reduce a significant amount of greenhouse gas emissions (15–66%) per year compared with the current value (de Jong et al. 2020). Thus, renewable production of EG would help to reduce greenhouse gas emissions. However, the current industrial production of EG occurs mainly by chemical synthesis through the hydration of ethylene oxide, a hazardous compound derived from the petrochemical industry (Cabulong et al. 2017a; Zhang et al. 2020). A minimal fraction of the total EG produced globally is produced from bioethanol conversion (de Jong et al. 2020).

Synthetic metabolic pathways to produce EG from biomass-derived xylose have been implemented in bacteria and yeast (Cam et al. 2016; Choi et al. 2017; Zhang et al. 2017; Valdehuesa et al. 2018). To date, there are three known synthetic metabolic pathways

*Correspondence:
João Ricardo Moreira de Almeida
joao.almeida@embrapa.br
[1]Graduate Program of Microbial Biology, Institute of Biology, University of Brasília, Brasília, Brazil
[2]Microbial Genetics and Biotechnology Laboratory, Embrapa Agroenergy, Brasília, Brazil
[3]Institute of Natural Materials Technology, TU Dresden, Dresden, Germany

to produce EG from xylose: (i) the Dahms pathway, (ii) xylulose-1 phosphate pathway, and (iii) ribulose-1 phosphate pathway (Stephen Dahms 1974; Cam et al. 2016; Salusjärvi et al. 2019). EG production from xylose was first achieved in *E. coli* using a modified Dahms pathway (Stephen Dahms 1974; Cam et al. 2016). In this pathway (Fig. 1), xylose is converted to xylonolactone (XLA) by a xylose dehydrogenase (XDH). Then, xylonolactone opens spontaneously or in a reaction catalyzed by lactonase into xylonate. The xylonate is dehydrated into 2-keto-3-deoxy-xylonate (KDX) with the action of a xylonate dehydratase (XD) enzyme. KDX is cleaved into glycolaldehyde and pyruvate by a dehydro-deoxy-xylonate aldolase (ALDO). The production of EG occurs through a reaction catalyzed by an aldehyde reductase (ALDR), which reduces glycolaldehyde to EG. Among the different metabolically engineered microorganisms, the highest EG production titers reached 108 g/L in *E. coli* (Chae et al. 2018) through the Dahms pathway. *Saccharomyces cerevisiae* was also engineered to produce EG using this pathway, although its production reached only 14 mg/L (Salusjärvi et al. 2017). The low efficiency of XD has been pointed out as the cause of low EG production titers in microorganisms converting the xylose via the Dahms pathway (Salusjärvi et al. 2017).

Komagataella phaffii, previously known as *Pichia pastoris*, is one of the yeasts with the most significant biotechnological potential in recombinant protein production (Yamada et al. 1995; Zahrl et al. 2017; Carneiro et al. 2022). More recently, *K. phaffii* has also gained attention as a microbial platform for the production of several chemical compounds of industrial interest (Carneiro et al. 2022; Gasser and Mattanovich 2018; Yang and Zhang 2018). This yeast can reach high cell biomass concentrations (40 g/L) with low nutrient requirements (Rebnegger et al. 2016), is acid-tolerant, has a Crabtree negative metabolism and genetic stability, tools for genetic manipulation and information about your physiology and metabolism (Carneiro et al. 2022). Different *Komagataella* species can naturally assimilate xylose but at different rates (Heistinger et al. 2022). Thus, considering the glucose and xylose contents in lignocellulosic hydrolysates, engineered *K. phaffii* strains should metabolize glucose for growth and maintenance while converting xylose to the desired metabolite. By implementing a newly identified bacterial xylose dehydrogenase, Ramos and colleagues recently constructed a *K. phaffii* strain, which produced up to 37 g/L xylonic acid with yields above 0.9 g/g (Ramos et al. 2021). This strain could also produce xylonic acid in sugarcane hydrolysate, demonstrating its relative tolerance to the hydrolysate. Previously, *K. phaffii* was engineered to assimilate xylose using a reductive pathway composed of xylose isomerase (Li et al. 2015). Another study also showed tolerance of *K. phaffii* to lignocellulose-derived inhibitors (Paes et al. 2021). Indeed, this yeast can grow on acetic acid, a common and abundant inhibitor in lignocellulosic hydrolysates. These results highlight a potential use of *K. phaffii* for EG production through a synthetic pathway using xylonic acid as an intermediate (Bañares et al. 2021; Trichez et al. 2022).

In this work, *K. phaffii* strains were engineered to produce EG from xylose by implementing a synthetic Dahms pathway. The pathway comprises the enzymes XDH, XD, ALDO, and ALDR, which were inserted into *K. phaffii* X-33 by expressing the respective bacterial genes. Since the conversion of xylonate into KDX by an XD enzyme was assumed to be rate-controlling for the synthetic EG production pathway, three new bacterial isogenes encoding XD enzymes were identified and evaluated. In addition, cultivation conditions that influence EG production were determined.

Materials and methods

Strains and growth media

The strain *E. coli* TOP10 (Invitrogen) was employed for cloning experiments, and the strain *K. phaffii* X-33 (Invitrogen) was selected as the host yeast strain for EG production (Table 2). *E. coli* strains were grown in low salt LB (1% tryptone, 0.5% NaCl, 0.5% yeast extract) supplemented with zeocin (25 µg/mL), kanamycin (50 µg/mL) or hygromycin B (50 µg/mL) when appropriate. The yeast strains were grown in YP (1% yeast extract, 2% peptone) supplemented with glucose (YPD) and/or xylose (YPX) at different concentrations. Zeocin (100 µg/mL), Geneticin (200 µg/mL), and hygromycin B (200 µg/mL) were added to the media used to cultivate *K. phaffii* when appropriate. LB or YP solidified with 2% agar was used to grow *E. coli* of *K. phaffii* on agar plates; the respective antibiotic was added as mentioned above. *E. coli* cultures were incubated at 37 °C and 220 rpm, and *K. phaffii* cultures were incubated at 30 °C and 200 rpm.

K. phaffii engineering for EG production

Vector construction

The gene *xylD* from *Halomonas lutea* (*xylB-HL*) encoding the first enzyme of the EG pathway (XDH) was previously cloned under the control of the P_{GAP} promoter (pGAP-HL vector) (Ramos et al. 2021). The three gene sequences encoding putative XDs (*xylD-HL*, *xylD-AM*, and *xylD-BS*) were codon optimized for expression in *K. phaffii*, synthesized by GenOne LTDA/SA, and inserted into the pKLD vector (Betancur et al. 2017). The coding sequences (Supplementary file) were cloned under the control of the P_{PGK1}*K. phaffii* promoter using the restriction sites *Bam*HI and *Not*I. The two previously known XDs, encoded by *yjhG-EC* and *xylD*-CC, were used as controls in the experiments, so the same expression

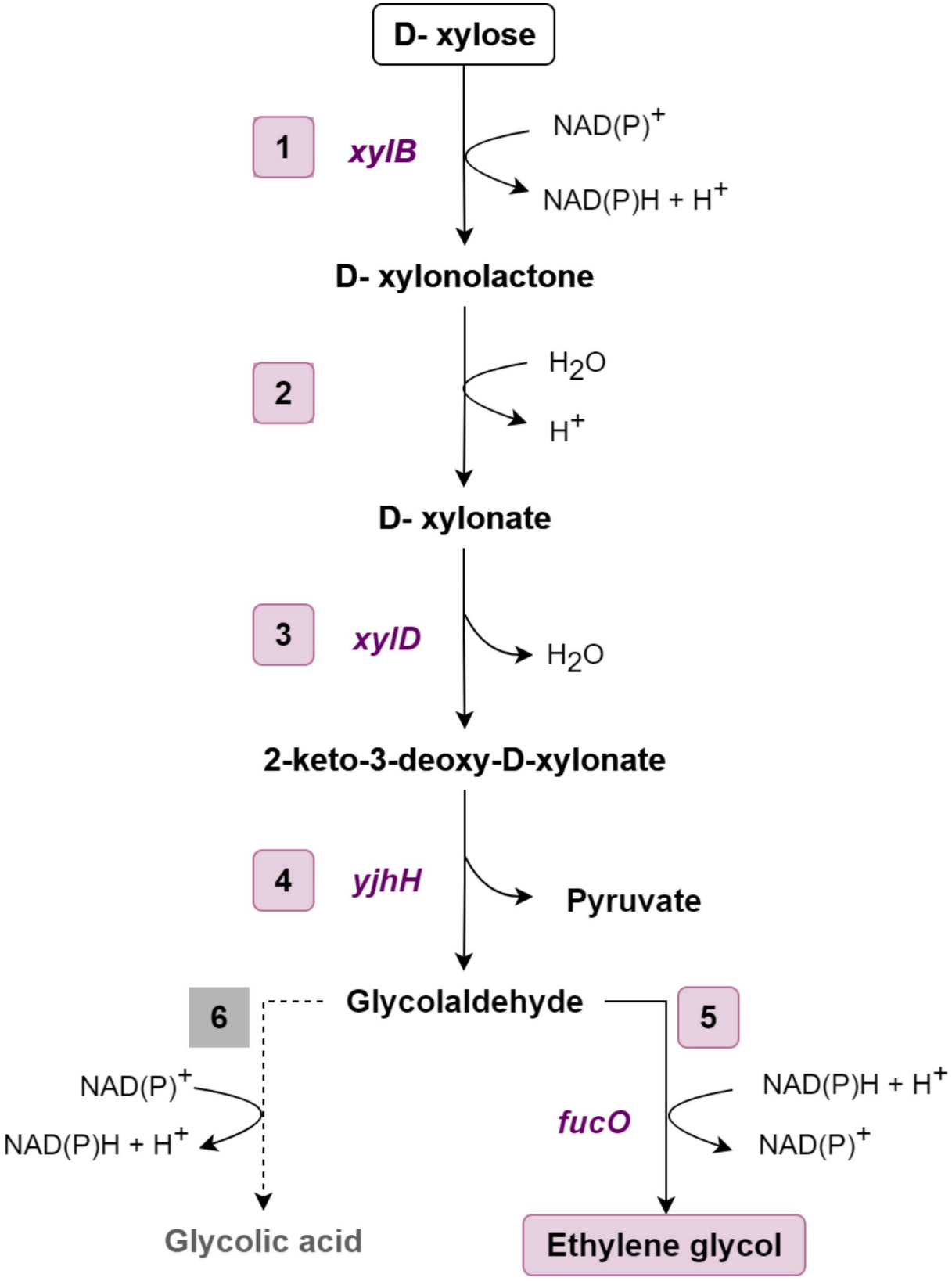

Fig. 1 Metabolic pathway overview for ethylene glycol (EG) production from xylose through the modified Dahms pathway, its enzymatic steps and pathway deviation to produce glycolic acid (GA). The numbers in purple represent each reaction that leads to EG production, and gray to GA: 1—xylose dehydrogenase (XDH); 2—xylonolactonase (XLA); 3—xylonate dehydratase (XD); 4—aldolase (ALDO); 5—aldehyde reductase (ALDR); 6—aldehyde dehydrogenase (ALDH). The names in italics describe the genes overexpressed in *K. phaffii* in this study

procedures were adopted for them. The plasmids were named pKLD-xylD-Xx, where Xx means the microorganism of origin (Table 1).

The two final enzymes of the EG-forming pathway (ALDO and ALDR) were expressed in the B3-036 vector (Prielhofer et al. 2017). For this purpose, two gene expression cassettes (named MOD1) were designed in silico and chemically synthesized (Supplementary file). For MOD1 construction, the *yjhH* gene encoding ALDO from *E. coli* was cloned under the control of the P_{TEF2} and CYC1 terminator using *Xho*I and *Sac*I sites. The same procedure was adopted for *fucO*, encoding ALDR from *E. coli*, but the P_{MDH3} and TDH3 terminator regulated its expression. MOD1 was chemically synthesized, cloned and inserted into the pBSK vector by GenOne LTDA/SA. MOD1 has at the restriction site *Bam*HI on both sides. To be expressed in *K. phaffii*, MOD1 was removed from the plasmid pBSK with the restriction enzyme *Bam*HI and inserted into the B3-036 vector (Prielhofer et al. 2017), previously digested with the same enzyme, resulting in the plasmid pB3Hyg_*YjhH* + *FucO*. To obtain MOD1, the B3-036 vector was first digested with the enzymes *Xho*I and *Sac*I; this reaction generated three nucleotide fragments with different sizes, 1973, 2310, and 4262 base pairs (bp). After the restriction reaction, the fragment with 4262 bp was purified with the Promega DNA purification kit. This fragment contains the cassette with the hygromycin B resistance gene (*hphR*) and the replication origin (PucOri). Then, the purified fragment was digested with *Bam*HI and ligated to MOD1, resulting in the plasmid pB3Hyg_MOD1. All plasmids constructed and employed in this study are listed in Table 1.

Strain construction

The first EG pathway enzyme (XDH) was previously expressed in *K. phaffii* by integrating the pGAP vector carrying the *xylD-HL* gene (Ramos et al. 2021). For the present work, the *K. phaffii* strain P1HL2, already expressing XDH, was chosen to host the enzymes of the complete EG metabolic pathway. The five pKLD-*xylD* plasmids (Table 1) were linearized with *Sac*I and introduced into the *K. phaffii* P1HL2 strain through electroporation (Invitrogen 2010). Transformants were first selected by growth on YPD plates supplemented with sorbitol (18.2%) and geneticin (200 µg/µL) at 30 °C. To further confirm the integration of the plasmid, colony PCR was carried out with the primers pPGK-D7(F) (5′-GTT CTCATCCATGAGTGAGTC-3′) and 3′AOX1(R) (5′-G CAAATGGCATTCTGACATCC-3′), resulting in amplification of the total length of the XD gene (~ 2000 bp). In addition to the 5 strains carrying the same XDH and different XDs, a control strain was also obtained by transformation with pKLD-empty (without *xylD*).

To construct the *K. phaffii* strain with the complete EG pathway, the *K. phaffii* strains that already had the XDH and XD integrated into the genome were transformed with the plasmid pB3Hyg_*YjhH* + *FucO*. For this, the transformation followed the same protocol described for the pKLD. After the completion of the transformation steps, the cells were spread on solid YPD plates supplemented with sorbitol (18.2%) and hygromycin B (200 µg/mL) and incubated at 30 °C until colony appearance. For confirmation, the replicative plasmid pB3Hyg_*YjhH* + *FucO* was extracted from the strains and PCR was performed with the primers PucOri(F) (5′-GA

Table 1 Plasmids used in this study

Plasmids	Characteristics	References
pGAPZB	Promoter GAP, Zeo^R	Invitrogen (2010)
pGAPZB_*xylB-HL*	Express XDH-HL, Zeo^R	Ramos et al. (2021)
pKGFP-ld	Express green fluorescent protein under control of pPGK1, Kan^R	Betancur et al. (2017)
pKLD	pKGFP-ld without the gene *EGFP*, Gen^R,	This work
pKLD_*xylD-AM*	Express XD-AM Gen^R	This work
pKLD_*xylD-BS*	Express XD-BS Gen^R	This work
pKLD_*xylD-HL*	Express XD-HL Gen^R	This work
pKLD_*YhjG-EC*	Express XD-*YhjG* Gen^R	This work
pKLD_*XylD-CC*	Express XD-*XylD* Gen^R	This work
pBSK	Amp^R, derived from pBluescript II SK (+)	Agilent Technologies (2008)
pBSK_MOD_1	Contain the expression cassetes pTEF-ALDO-CYCtt and pMDH3-ALDR-TDH3tt, derived from pBSK carrying the genes *YjhH* and *FucO*, Amp^R	This work
B3-036	Carrying Ca9 gene, Hyg^R	Prielhofer et al. (2017)
pB3Hyg_*YjhH* + *FucO*	Derived from B3-036 carrying the genes *YjhH* and *FucO*, Hyg^R	This work
pB3Hyg_*YjhH*	Derived from B3-036 carrying the genes *YjhH*, Hyg^R	This work

*Promoter for heterologous gene expression in *K. phaffii*

[R]Antibiotic resistance

TCCGGCAAACAAACCACC-3′) and CyctR (5′-CGT ACACGCGTCTGTACAGA-3′). These primers amplified the first gene of MOD1 (ALDO), with a length of approximately 1900 bp. All positive clones from each construct were cryopreserved in a − 80 °C freezer. One clone of each combination of enzymes was selected for the experiments.

To construct *K. phaffii* strains without the heterologous *fucO* encoding the ALDR enzyme, digestion of the plasmid pB3Hyg_*YjhH* + *FucO* was carried out with the restriction enzyme *PvuII*. The enzyme cleaves the *fucO* protein-encoding sequence together with the MDH3 promoter and the TDH3 terminator from the plasmid pB3Hyg_*YjhH* + *FucO*. After digestion, the plasmid backbone without the *fucO* expression cassette was purified and joined using the T4 ligase enzyme. Once the accurate construct configuration was confirmed, the pB3Hyg_*YjhH* plasmid was inserted into *E. coli*, purified and subsequently introduced into *K. phaffii* through cloning. Table 2 lists all the strains selected and used in this work.

Cultivation experiments

To evaluate the functionality of the new XD, the strains expressing XDH and one of the five XDs were incubated in YPDX medium. Initially, strains were recovered from cryogenic glycerol stocks (30%) and transferred to Petri dishes with YPD agar medium supplemented with zeocin (100 µg/mL) and geneticin (200 µg/mL). Precultures were initiated by inoculating each strain grown in solid media in 10 mL YPD medium in 50 mL conical tubes for 48 h in a shaker at 30 °C and 200 rpm. After 48 h, the preculture was transferred to 100 mL of YPD medium in 500 mL flasks for overnight growth under the same conditions of precultures. Then, cultivations were initiated by inoculating each strain to an optical density (OD_{600nm}) equal to 10 in 30 mL YP supplemented with glucose (2 g/L) and xylose (40 g/L) in 250 mL flasks. Samples were collected at 0, 24, and 48 h of cultivation and analyzed for OD_{600nm}, substrate consumption, and metabolite production by high-performance liquid chromatography (HPLC) and liquid chromatography coupled with mass spectrometry (LC-MS).

To evaluate the capacity of engineered *K. phaffii* strains to produce EG under different experimental conditions, these were cultivated as previously mentioned, except for the initial OD600nm, glucose and xylose concentrations, which varied in some experiments. The values employed are specified where appropriate. Cultivations with EG and GA as carbon sources were performed for 72 h in YP supplemented with 5 g/L EG or AG and antibiotics. Samples for OD_{600nm} analysis, substrate consumption, and metabolite production were collected and evaluated as in previous cultivations.

Analytical methods

Quantification of glucose, EG, and glycolic acid was performed as described by Vieira (2018) and Costa et al. (2019). In brief, samples were injected in an Ultra-High-Performance Liquid Chromatography (UHPLC) system (Waters AcQuity UPLC H-Class RID) equipped with an Aminex HPX87H (Bio-Rad) column. The flow rate was adjusted to 0.6 mL/min with a mobile phase of 5 mM H_2SO_4. The injection sample volume was 10 µL, and the

Table 2 List of strains used in this study

Strain	Genotype	Phenotype	Refs.
E. coli TOP10	F-mcrA Δ(mrr-hsdRMS-mcrBC) φ80lacZΔM15 ΔlacX74 recA1 araD139 Δ(ara-leu) 7697 galU galK rpsL (StrR) endA1 nupG λ	–	Invitrogen (2006a)
K. phaffii X-33	Prototrophic	–	Invitrogen (2006b)
K. phaffii P1HL2	pGAP_ *xylB-HL*	Zeo^R + XDH	Ramos et al. (2021)
K. phaffii JA120	pGAP-*xylB-HL* + pKLD-*xylD-AM* + pB3Hyg-MOD1 (*yjhH* + *fuc0*)	Zeo^R + Gen^R + Hyg^R XDH + XD + ALDO + ALDR	This work
K. phaffii JA121	pGAP-*xylB-HL* + pKLD-*xylD-bs* + pB3Hyg-MOD1 (*yjhH* + *fuc0*)	Zeo^R + Gen^R + Hyg^R XDH + XD + ALDO + ALDR	This work
K. phaffii JA122	pGAP-*xylB-HL* + pKLD-*xylD-HL* + pB3Hyg-MOD1 (*yjhH* + *fuc0*)	Zeo^R + Gen^R + Hyg^R XDH + XD + ALDO + ALDR	This work
K, phaffii JA123	pGAP-*xylB-HL* + pKLD-y*hjG-EC* + pB3Hyg-MOD1 (*yjhH* + *fuc0*)	Zeo^R + Gen^R + Hyg^R XDH + XD + ALDO + ALDR	This work
K, phaffii JA124	pGAP-*xylB-HL* + pKLD-*xylD-CC* + pB3Hyg-MOD1 (*yjhH* + *fuc0*)	Zeo^R + Gen^R + Hyg^R XDH + XD + ALDO + ALDR	This work
K, phaffii JA125	pGAP-*xylB-HL* + pKLD-*empty*	Zeo^R + Gen^R XDH	This work
K, phaffii JA126	pGAP-*xylB-HL* + pKLD-*empty* + pB3Hyg-*empty*	Zeo^R + Gen^R + Hyg^R XDH	This work
K, phaffii JA128	pGAP-*xylB-HL* + pKLD-*xylD* + pB3Hyg-*empty*	Zeo^R + Gen^R + Hyg^R XDH + XD	This work
K, phaffii JA131	pGAP-*xylB-HL* + pKLD-*xylD-HL* + pB3Hyg-y*jhH*	Zeo^R + Gen^R + Hyg^R XDH + XD + ALDO	This work
K, phaffii JA133	pGAP-*xylB-HL* + pKLD-*xylD-CC* + pB3Hyg-y*jhH*	Zeo^R + Gen^R + Hyg^R XDH + XD + ALDO	This work

column was kept at 45 °C. The analytical run was carried out for 24 min.

Quantification of xylose, xylonic acid, xylitol, and arabitol was achieved by coupling liquid chromatography (VanquishTM, Thermo FisherTM, USA) with a mass spectrometer (Q ExactiveTM Focus orbitrap, Thermo FisherTM, USA). The samples were prepared by centrifugation at 13.800×g for 10 min. Then, the supernatant was individually filtered with a 0.2 µm filter. Subsequently, samples were injected, and metabolites were separated using a Rezex™ ROA-Organic Acid H+column (300×7.8 mm, 8%, Phenomenex) protected by a SecurityGuard™ Carbo H+precolumn (4×3 mm, Phenomenex). The analytes were eluted with 0.1% formic acid at a constant flow rate of 0.4 mL/min and a column temperature of 80 °C. Monoisotopic mass and retention time of individually injected standards were used for analyte identification and quantification.

The qualitative analysis of KDX was performed in a Nexera X2 UHPLC system (Shimadzu, Kyoto, Japan) equipped with an Acquity UPLC HSS T3 column (2.1 mm × 100 mm, 1.8 µm) (Waters Technologies, Milford, USA) coupled to a high-resolution MaXis 4G™ Q-TOF MS analyzer (Bruker Daltonics, Germany) through an electrospray ionization source. The chromatographic run parameters were isocratic from 0.0 to 1.0 min (0% B), linear gradient from 1.0 to 3.0 min (0–5% B), linear gradient from 3.0 to 10.0 min (5–50% B), linear gradient from 10.0 to 13.0 min (50–100% B), isocratic from 13.0 to 15.0 min (100% B) and isocratic from 15.0 to 20.0 min (0% B). Eluent A was 0.1% v/v formic acid in water, and eluent B was 0.1% v/v formic acid in methanol. The flow rate, oven temperature and injection volume were 0.4 mL/min, 40 °C, and 1 µL of sample, respectively. Fermentation samples were diluted 30x before injection. Standards were injected at 30 µg/mL.

Data were acquired in negative ion mode. The MS parameters were end plate offset=500 V, capillary voltage=4000 V, nebulizer pressure=4.0 Bar, dry gas flow=9.0 L/min, dry temperature=200 °C, spectra rate=3.00 Hz, and detection range=m/z 75–1000. Sodium formate (2 mM) in isopropanol: water (1:1 v/v) was directly injected through a 6-port valve at the beginning of each chromatographic run for external calibration. The MS/MS parameters were collision energy, 20–50 V using a basic stepping program; cycle time of precursor ion acquisition, 3.0 s; mass range, 75–1000 m/z; spectra rate, 3 Hz; pre-pulse storage, 7.0 µs; funnel 1 RF, 300.0 Vpp. UHPLC–MS/MS data were acquired with otofControl (Bruker Daltonics) and HyStar version 3.2 (Shimadzu) and visualized with DataAnalysis 4.2 (Bruker Daltonics). The ion extraction tool was used to manually search the m/z values of different ionization forms of some expected fermentation products in the chromatograms.

Cell growth was monitored by absorbance at OD_{600} with a spectrophotometer (SpectraMax M2, Molecular Devices). 5 mL aliquots were withdrawn from the starting and the last fermentation point for the cell-dry weight measurements. The aliquots were centrifuged at 3000×g for 25 min. The supernatant was discarded, and the pellet was washed in deionized water and centrifuged again. The supernatant was discarded, and the pellet was dried in a glass tube at 60 °C (approximately 96 h). Finally, the cell mass was measured with an analytical balance.

Results

Identification and validation of new xylonate dehydratases (XD)

The synthetic EG production pathway implemented in this work involves four enzymatic steps (Fig. 1). In the first reaction, xylose is converted by XDH to xylonolactone, which can spontaneously open to xylonic acid. Then, XD dehydrates xylonic acid into KDX. The best-characterized known XDs are those encoded by *xylD* from *Caulobacter crescentus* (*xylD*-CC) (Stephens et al. 2007) and *yhjG* from *E. coli* (*YhjG-EC*) (Jiang et al. 2015). Previously, XDH from the bacteria *Halomonas lutea*, *Azospirillum amazonense*, and *Brevundimonas subvibrioides* were successfully identified and successfully expressed in *K. phaffii* (Ramos et al. 2021); thus, the putative XD of the same bacterial species were chosen for evaluation in this study. This strategy was employed to increase the chance of identifying functional enzymes that can be properly expressed by *K. phaffii*. The *xylD* sequences, which putatively encode XD, from *(A) amazonense* (*xylD-AM*—accession number WP_013269499.1), *(B) subvibrioides* (*xylD-bs*—accession number WP_019016838.1), and *H. lutea* (*xylD-HL*—accession number WP_019016838.1) showed identities of 79.9, 61.38, and 64.31% to *xylD* from *(C) crescentus.*

To validate the in vivo functionality of the three selected enzymes, integrative plasmids for the expression of five coding sequences (3 putative XDs and the controls *xylD*-CC and *yhjG-EC* known to encode functional XDs) were constructed and transformed into a *K. phaffii* strain already overexpressing the gene *xylB* from *H. lutea* (Ramos et al. 2021). Transformants were selected on a solid YPD medium with zeocin and geneticin, and plasmid integration was verified by PCR amplification of the XD-encoding gene, generating strains expressing XDH along with one of the five XDs. To evaluate the XD functionality in vivo, the five strains carrying XDH+XD, as well as the two control strains (*K. phaffii* X33 and *K. phaffii*-XDH) were cultivated in a YPDX medium containing 40 g/L of xylose and 2 g/L of glucose. After 48 h of incubation, culture supernatants were collected and qualitatively evaluated by LC-MS for xylonic acid and KDX production. As expected, *K. phaffii* X33 did not produce xylonic acid or KDX, whereas *K. phaffii*-XDH did. The strains expressing XDH and XD produced KDX,

confirming the functional expression of the respective dehydratase (data not shown).

Construction of *K. phaffii* EG-producing strains

To complete the EG synthetic pathway, the enzymes ALDO and ALDR, which are responsible for the cleavage of KDX into glycolaldehyde and pyruvate, and the conversion of glycolaldehyde into EG, respectively (Fig. 1), were also expressed in the yeast. For this, the coding sequences for ALDO (*yjhH)* and ALDR (*fucO)* from *E. coli* (Liu et al. 2013; Alkim et al. 2015) were cloned under the control of the *K. phaffii* promoters TEF2 and MDH3, respectively. The expression cassettes were inserted into the plasmid pB3-036, resulting in the vector pB3Hyg_MOD_1 (Table 2). The plasmid was transformed into the strains of *K. phaffii* harboring the XDH and one of the five XDs. The resulting transformants were confirmed by growth on selective medium and colony PCR for the inserted genes. Thus, five strains were obtained, each expressing a different XD but exhibiting the same expression cassettes for XDH, ALDO, and ALDR. The resulting strains, which are differentiated by XD, were named *K. phaffii* JA120 (*xylD-AM*), JA121 (*xylD-bs*), JA122 (*xylD-HL*), JA123 (*xylD-EC*), and JA124 (*xylD-CC*) (Table 2).

To confirm the functionality of the pathway, the five strains with the complete pathway and two controls (expressing only XDH and XDH + XD, respectively JA125 and JA126 (Table 2) were cultivated in YP supplemented with 40 g/L xylose and 2 g/L glucose for 48 h. All strains expressing the complete pathway were able to produce EG (Fig. 2). The *K. phaffii* strains JA123 and JA124 carrying the genes that encode the XDs already characterized in the literature (*yjhG-EC* and *xylD-CC*, respectively) produced 0.39 and 0.51 g/L EG in the first 24 h, whereas the strains harboring the newly identified XD, JA122 (XD-HL), JA121 (XD-BS), and JA120 (XD-AM) produced 0.84, 0.26 and 0.41 g/L EG, respectively (Fig. 2). It is worth noting that one of the three tested new XD enzymes (i.e., *xylD-HL*) allowed better EG production under the evaluated conditions than the previously studied enzymes in *K. phaffii*.

Glycolic acid (GA) production was also observed in all strains harboring the synthetic Dahms pathway and varied considerably over time (Fig. 2). The GA concentration was lower than that of EG in the first 24 h of cultivation; however, it increased considerably (by more than 100%) after 48 h of cultivation. The highest value detected at this time was 1.69 g/L for strain JA122 (Fig. 2). The amount of GA produced by this strain is two times higher than the

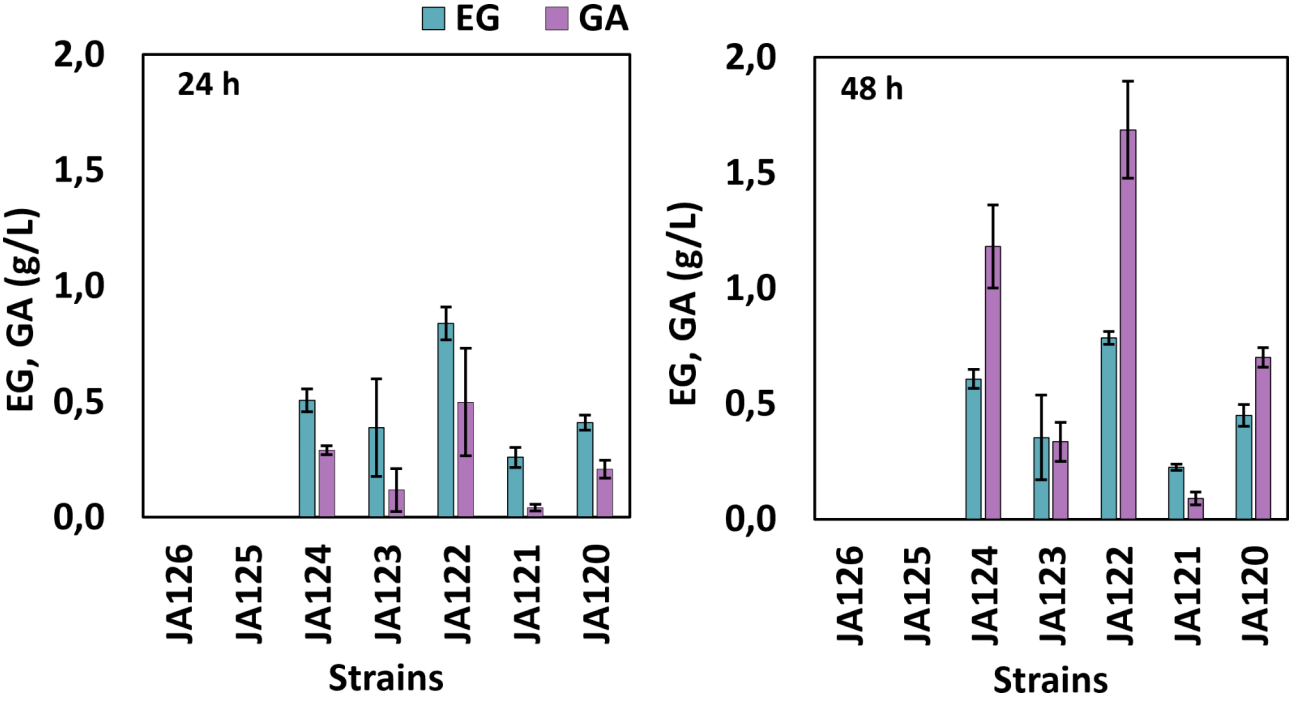

Fig. 2 Ethylene glycol (EG) and glycolic acid (GA) production through co-fermentation of glucose and xylose by the engineered strains. The five strains with the complete pathway express (XDH + XD + ALDO + ALDR) and differ by the XD employed. Positive controls: JA124 (XD-Cc) and JA123 (*XD-EC*); new XD: JA122 (*XD-HL*), JA121 (*XD-Bs*) and JA120 (*XD-AM*). Negative controls: JA126 (expresses only XDH-*HL*) and JA125 (*XDH-HL* + XD-EC + *ALDR*). Strains were inoculated to OD_{600nm} 5 in 250 mL shake flasks filled with 20 mL of YP supplemented with 2 g/L glucose and 20 g/L xylose and antibiotics (zeocin, geneticin and hygromycin). Samples were taken after 24 h and 48 h of cultivation and supernatants were analyzed for EG and GA concentration. The results shown are mean values and standard deviations of three biological replicates. The standard deviation values represent 95% of the confidence interval in a normal curve ($z = 1.96$)

EG concentration at the same time point. The fact that the GA concentration was higher than the EG production after 48 h was also valid for the other strains (Fig. 2).

Native glycolaldehyde reductase (ALDR) activities in *K. phaffii*

The production of GA by the engineered strains indicates that *K. phaffii* possesses endogenous ALDH activity since part of the carbon flux was directed away from the pathway at the level of glycolaldehyde toward GA production (Fig. 2). To confirm the native ALDH and to evaluate possible native ALDR activity, strains expressing only the first three enzymes of the synthetic pathway (XDH+XD+ALDO; without heterologous ALDR were constructed). For this, ALDO was expressed in the strains *K. phaffii*-XDH-XD-*CC*) and *K. phaffii*-XDH-XD-*HL*) to evaluate the effect of different dehydratases. The two resulting *K. phaffii* strains expressing XDH-XD-ALDO (JA133 *xylD*-CC and JA131 *xylD*-HL) were cultivated as previously described in the presence of 40 g/L xylose and 2 g/L glucose. After 48 h of cultivation, the presence of EG and GA was evaluated. The strains could produce both compounds, confirming that *K. phaffii* possesses native ALDR and ALDH activities that can convert glycolaldehyde to the corresponding products (Fig. 3). The strains carrying heterologous ALDR (JA124 and JA122) produced 0.61 and 0.79 and without heterologous ALDR (JA133 and JA 131) produced 0.35 and 0.46 g/L EG, respectively (Fig. 3). EG production increased on average 1.71 times when the heterologous ALDR enzyme was present (Fig. 3). These results indicate that the native ALDR of *K. phaffii* converts glycolaldehyde to EG under the evaluated conditions, although heterologous ALDR expression increases EG production.

Conversion of EG and GA by *K. phaffii*

As *K. phaffii* can reduce and oxidize glycolaldehyde to EG and GA, and the EG concentration is slightly reduced over time (Fig. 2), it was also evaluated whether it possesses EG and GA conversion activities. Thus, *K. phaffii* X-33, *K. phaffii* JA122 (XDH+XD+ALDO+ALDR), and *K. phaffii* JA131 (XDH+XD+ALDO) were incubated in the presence of 5 g/L EG or GA as substrates. The 3 strains could convert more than 50% of the EG into GA (Fig. 4). The expression of the *E. coli* ALDR did not significantly change the conversion rate of EG to GA among strains. In addition, no GA conversion was observed when the yeast strains were incubated in the presence of the acid (data not shown).

Effects of glucose and xylose ratio in EG production

The *K. phaffii* EG-producing strains showed an increased production of GA by the end of cultivation experiments when glucose in the medium was low or finished. In addition, variable glucose concentrations in the medium were shown to inhibit or favor xylose consumption and modify product formation in *S. cerevisiae* (Hahn-Hägerdal et al. 2007; Hou et al. 2017). Thus, the effects of different sugar ratios on *K. phaffii* JA122 metabolism were evaluated to better understand how glucose and xylose can influence EG and GA production. For this, the strain JA122 was incubated in YP supplemented with 20:20 g/L (1:1), 10:20 g/L (1:2), 10:40 g/L (1:4), and 2:40 g/L (1:20) glucose and xylose.

The results demonstrate that the strain JA122 produced EG and GA in all conditions evaluated but at different levels (Fig. 5). GA production has been commonly reported in bacteria and yeast strains engineered to produce EG. For *K. phaffii*, higher glucose concentrations in the medium favor EG production while inhibiting GA production through an unknown mechanism (Fig. 5A–C). Indeed, GA was not produced until glucose was practically depleted from the medium (Fig. 5). The maximum production of EG (1.31 g/L) was achieved when the glucose and xylose were in an initial proportion of 1:2 (Fig. 5B). In contrast, more GA (1.4 g/L) was produced when the medium's initial glucose concentration was below 2 g/L (Fig. 5D).

Fermentative profile of the *K. phaffii* EG-producing strain

The fermentative profile of strains was compared to analyze the functionality of the EG synthetic pathway further and better understand the metabolism of the engineered *K. phaffii*. For this, strains JA122, which expresses the complete EG pathway and showed the best EG production (Fig. 2), and JA128 (control strain that expresses only XDH and XD) were cultivated in YP medium supplemented with 20 g/L glucose and 20 g/L xylose. As expected, EG production was detected in the cultivation with JA122 but not in the control strain (Fig. 6). The fermentative profile of the *K. phaffii* JA122 strain was very similar to that of JA128. The main difference between these strains was the presence of 0.18 g/L EG and 0.45 g/L GA in JA122 and none for the control strain (Fig. 6). The strain *K. phaffii* JA122 assimilated almost 9 g/L of xylose until 48 h of cultivation and then stopped, correlating with the complete assimilation of glucose (Fig. 6). The strain JA128 expressing only XDH and XD consumed 6 g/L xylose and stopped at 48 h. The xylonic acid (4 g/L) and xylitol (~ 1 g/L) production followed almost the exact pattern for the two strains. Both strains also produced arabitol while glucose was present in the medium, reaching its maximum of approximately 5 g/L in 24 h (Fig. 6).

Discussion

This is the first time that a complete pathway enabling EG production is functionally expressed in *K. phaffii*. The ability to engineer *K. phaffii* to metabolize xylose—a

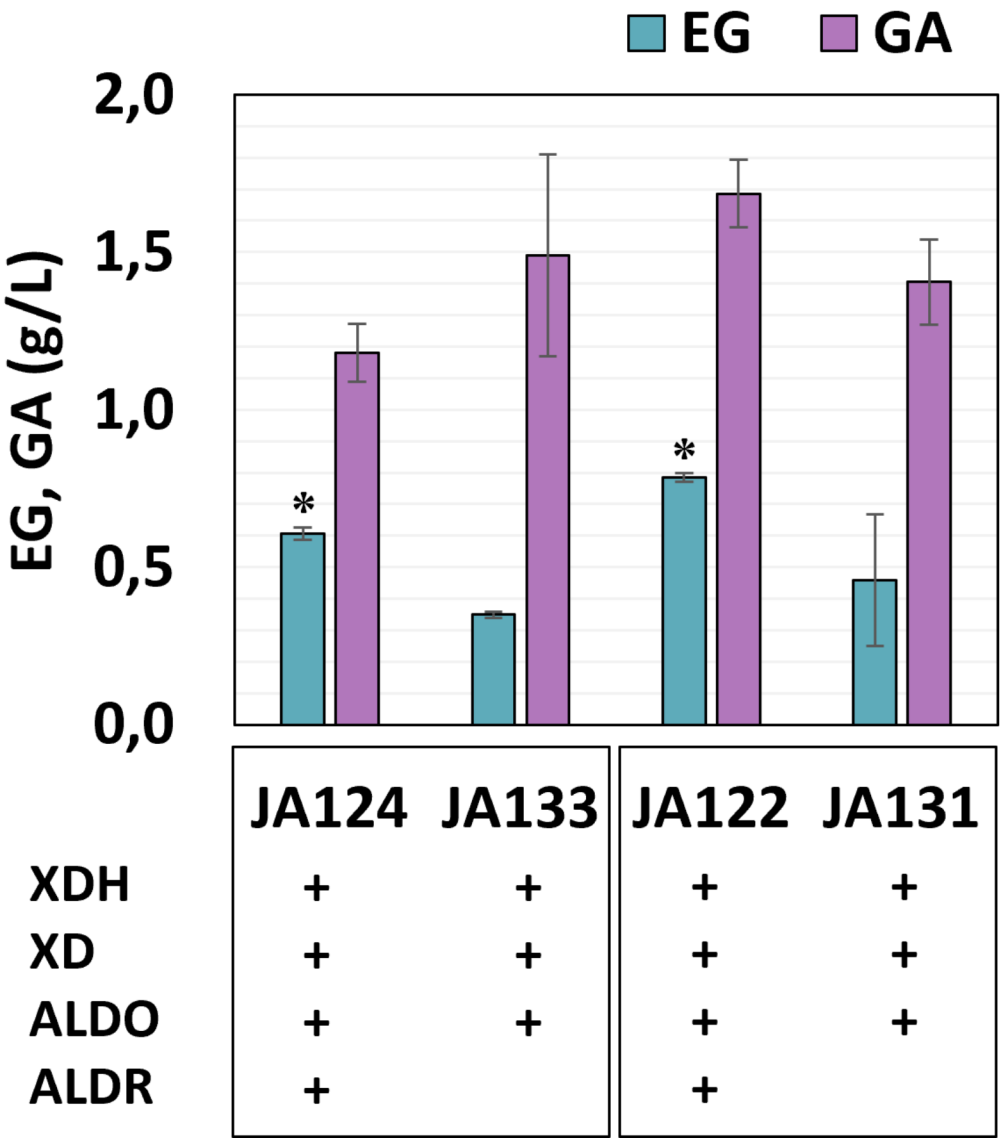

Fig. 3 Ethylene glycol (EG) and glycolic acid (GA) production through co-fermentation of glucose and xylose by the engineered strains, harboring the genes encoding for a heterologous aldehyde reductase (ALDR) (JA124 and JA122) or not (JA133 and JA131). Strains were cultivated in shake flasks with an initial OD_{600nm} of 10 in 250 mL shake flasks with 50 mL of YP supplemented with 40 g/L xylose, 2 g/L glucose and hygromycin. The results shown are mean values and standard deviations of two biological replicates. *Statistical significance in a standard deviation confidence interval of 63% ($z=1$) when comparing the correlation of the EG production values between JA124 and JA133 and JA122 and JA131. Heterologous expression: JA124 (*xylB-HL* + *xylD-CC* + *yjhH* + *fucO*), JA133 (*xylB-HL* + *xylD-CC* + *yjhH*), JA122 (*xylB-HL* + *xylD-HL* + *yjhH* + *fucO*), JA131 (*xylB-HL* + *xylD-HL* + *yjhH*). XDH: xylose dehydrogenase, XD: xylonate dehydratase, ALDO: aldolase, ALDR: aldehyde reductase. "+" indicate the presence of the heterologous enzyme

major component of lignocellulosic biomass - broadens its potential for use in bio-based processes. Previously, EG production in yeasts was only evaluated using *S. cerevisiae* (Table 3). The highest EG concentrations achieved were 14 mg/L through the Dahms pathway and 4 g/L in a strain harboring the xylose 1-P pathway, but only when fed-batch cultivations with glucose and xylose were carried out (Table 3). Thus, the 1.31 g/L of EG produced with *K. phaffii* under unoptimized cultivation conditions demonstrates this yeast's potential. *K. phaffii* is well-established as a robust platform for heterologous protein production due to its ability to grow at high cell densities, tolerate stressful environmental conditions, like low pH; exhibit fast growth rates, has genetic stability and is a Crabtree negative yeast (Carneiro et al. 2022). The successful engineering of *K. phaffii* to produce EG

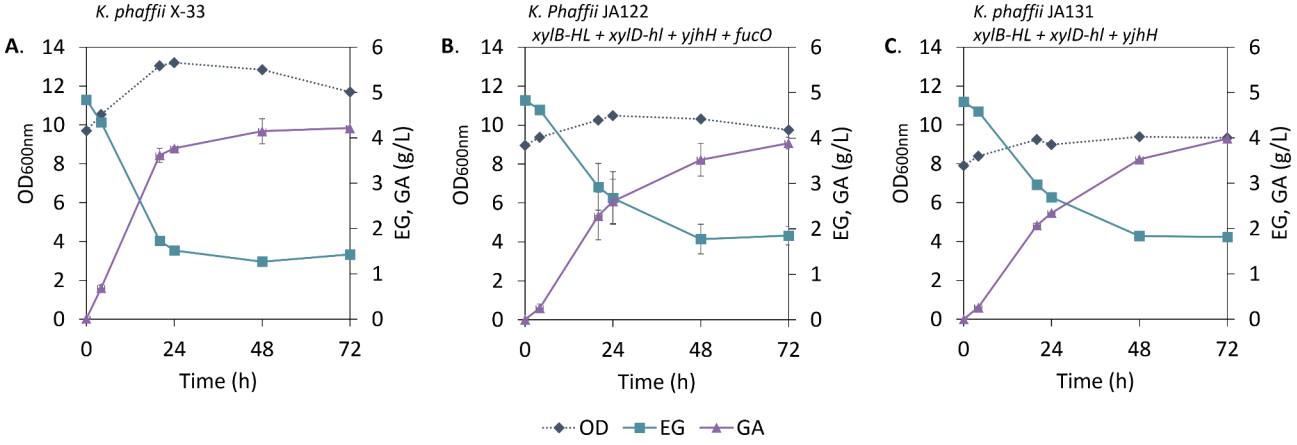

Fig. 4 Impact of the presence of a heterologous aldehyde reductase (ALDR) on the conversion of ethylene glycol (EG) to glycolic acid (GA) in YP medium with EG as substrate. **A** Prototrophic *K. phaffii* X-33 **B** engineered *K. phaffii* JA122 harboring XDH + XD + ALDO + ALDR; **C** engineered *K. phaffii* JA131 harboring XDH + XD + ALDO were inoculated to OD_{600} of 10 in 250 mL shake flasks with 50 mL of YP supplemented with 5 g/L EG. Cultivation was carried out in shake flasks for 72 h. The fermentation was performed in triplicate. The standard deviation values represent 95% of the confidence interval in a normal curve ($z = 1.96$)

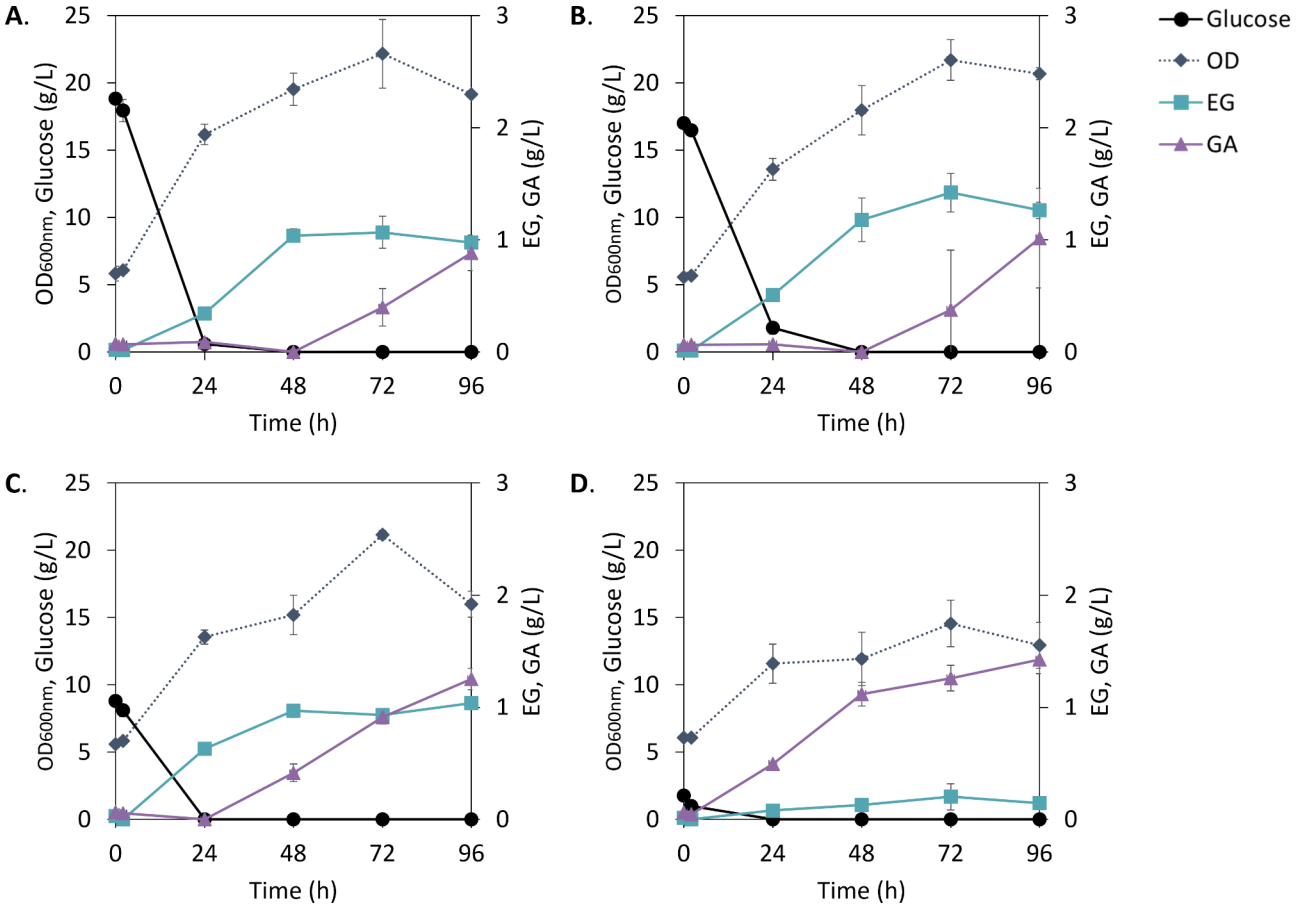

Fig. 5 Effects of glucose and xylose ratio in ethylene glycol (EG) and glycolic acid (GA) production. The strain *K. phaffii* JA122 (*xylB-HL* + *xylD-HL* + *yjhH* + *fucO*) was inoculated to OD_{600nm} 5 in 125 mL flasks containing 20 mL of YP supplemented with glucose and xylose at, respectively: **A** 20:20 g/L (1:1), **B** 10:20 g/L (1:2), **C** 10:40 g/L (1:4), and **D** 2:40 g/L (1:20). The cultivation was carried out in shake flasks for 96 h in biological duplicate

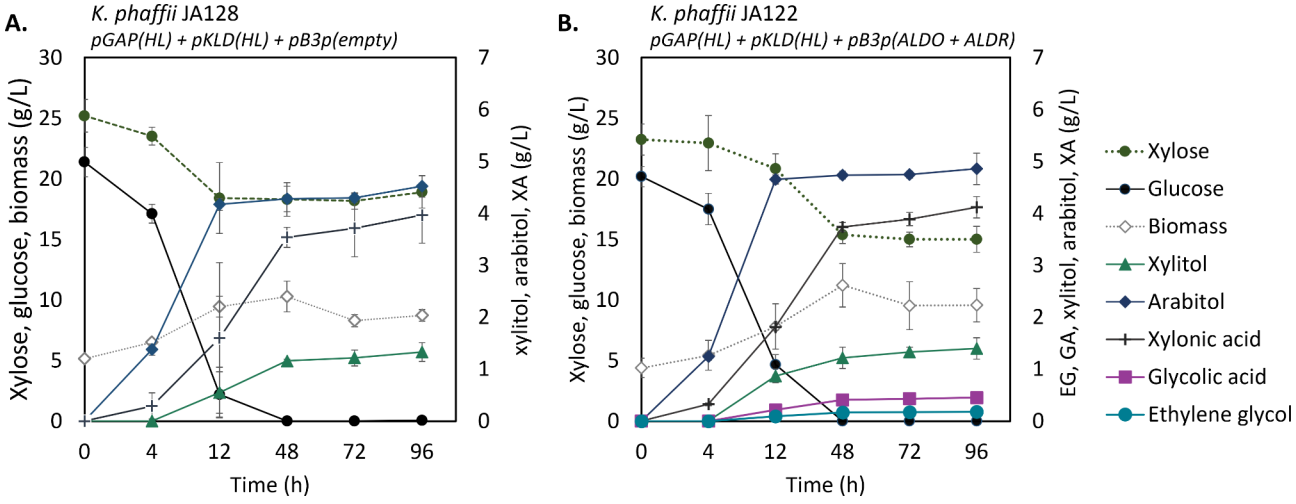

Fig. 6 Metabolite production overview from cultivation in medium with YP + 20 g/L glucose and 20 g/L xylose. The co-fermentation of glucose and xylose was performed in shake flasks for 96 h in biological triplicates. Strains utilized in this experiment **A***K. phaffii* JA128 (*xylB-HL* + *xylD-HL*) and **B***K. phaffii* JA122 (*xylB-HL* + *xylD-HL* + *yjhH* + *fucO*)

Table 3 A brief review of microorganisms able to produce ethylene glycol (EG) from xylose through engineered synthetic pathways

Strain	EG (g/L)	Y (g/g)	Y (mol/mol)	Bioprocess	Substrate	Pathway	Refs.
Bacteria							
E. coli WL3110	108.2	0.36	0.87	Fed batch	Xylose	Dahms	Chae et al. (2018)
E. coli BL21(DE3)	72.0	0.40	0.97	Fed batch	Xylose	Dahms	Wang et al. (2018)
E. coli MG1655	40.0	0.35	0.85	Fed batch	Xylose	Ribulose-1P	Pereira et al. (2016)
E. coli MG1655	20.0	0.38	0.91	Fed batch	Xylose	Xylulose-1P	Alkim et al. (2015)
E. coli W3110(DE3)	11.7	0.29	0.70	Batch	Xylose	Dahms	Liu et al. (2013)
E. coli W3110	7.7	0.39	0.95	Batch	Xylose	Dahms	Cabulong et al. (2017b)
E. cloacae S1	34.1	0.28	0.99	Fed batch	Xylonic acid	Dahms	Zhang et al. (2020)
C. glutamicum SL-1 A pIEARKT	5.8	0.31	0.75	Batch	Xylose	Ru1P	Lee et al. (2019)
Yeasts							
K. phaffii JA122	1.31	–	–	Batch	Xylose + Glucose	Dahms	This work
S. cerevisiae D452-2	0.5	0.01	0.03	Batch	Xylose	Xylulose-1P	Chomvong et al. (2016)
S. cerevisiae H4099	0.014	–	–	Batch	Xylose + Glucose	Dahms	Salusjärvi et al. (2017)
S. cerevisiae H131-A3CS	4.0	–	–	Fed batch	Xylose + Glucose	Xylulose-1P	Uranukul et al. (2019)

demonstrates its potential as a microbial factory for high-value chemical synthesis, adding to the range of molecules that can be produced biotechnologically.

Our results offer three novel functional XD isoenzymes that can be used to construct synthetic pathways for converting xylose to valuable chemicals. The previously characterized XD from *C. crescentus* (*Xyl*D-CC) and *yhjG* from *E. coli* (*YhjG-EC*) showed low enzymatic activity in *S. cerevisiae* and, therefore, represented a bottleneck for EG production via the synthetic Dahms pathway (Salusjärvi et al. 2017). The strain carrying a newly identified XD (XD-HL) showed a 30% increase in EG production compared to the control strains carrying the previously characterized dehydratases under the conditions tested. Even if a direct comparison of enzyme performances cannot be made due to experimental limitations (enzyme kinetics and expression levels were not determined), these results reinforce the observations the XD activity limits xylonate conversion to KDX (Andberg et al. 2016). The next step in optimizing EG production in *K. phaffii* involves the XDs kinetic characterization and optimization of expression levels, as these steps have already been shown to improve EG and/or GA production in *E. coli* (Chae et al. 2018).

The engineered *K. phaffii* strains produced both EG and GA, with maximum EG production of 1.31 g/L and GA reaching 1.69 g/L after 48 h of cultivation. The native ALDR of *K. phaffii* converts glycolaldehyde to EG, although heterologous ALDR expression increases EG production. This dual production of EG and GA occurs because *K. phaffii* possesses native enzymatic activities that convert glycolaldehyde into both compounds (Fig. 4). The reduction of glycolaldehyde to EG by endogenous enzymes was previously observed in *S. cerevisiae* (Salusjarvi et al. 2017). Similar findings have also been reported in engineered *E. coli* strains, where native

enzymatic activities affect product formation (Cabulong et al. 2017a). In addition, *K. phaffii* can convert EG into GA, but not the reverse (Fig. 4), indicating that other enzymes may be involved in EG metabolism in vivo. It was shown that the purified *K. phaffii* alcohol oxidase could oxidize EG to glycolaldehyde in vitro, but the in vivo activity was not evaluated (Isobe et al. 2012). Since *AOX1* expression is repressed by glucose and requires induction by methanol (Wang et al. 2017), the results obtained here indicate that EG could act as an inductor of AOX1 and that other enzymes may be involved in EG conversion to GA via glycolaldehyde in vivo. Indeed, for other yeasts, such as *S. cerevisiae* and *Yarrowia lipolytica*, it has been suggested that the conversion of EG is likely driven by promiscuous enzymes, most probably dehydrogenases, which may be upregulated during the stationary phase. However, details on the reaction directionality are unavailable (Carniel et al. 2024; Senatore et al. 2024). The identification of native glycolaldehyde reduction (ALDR) and oxidation (ALDH) activities and EG oxidation to GA in this yeast is essential to elaborate further strategies of strain improvement, allowing the deletion of competitive pathways or overexpression of particular genes to drive the carbon flux to a specific product.

A higher glucose concentration in the culture medium favors EG production, while GA production increases when glucose is depleted (Fig. 5). This may suggest a regulatory effect of glucose on the pathways involved in GA formation. Glucose may downregulate the activity of *K. phaffii* native ALDH, which converts glycolaldehyde into GA (Senatore et al. 2024). Similarly, it is known that the transcription factors that regulate the *AOX* genes in *K. phaffii* are inhibited by the presence of glucose (Wang et al. 2017). Alternatively, glucose consumption is known to increase glycolytic fluxes in engineered xylose-consuming *S. cerevisiae* strains, indirectly increasing xylose consumption by improving energy and oxidative balance in the cell (Hahn-Hägerdal et al. 2007). Finally, fast glucose consumption may reduce oxygen availability in the medium and modify the cofactor balance in the cell. Thus, optimizing the balance between glucose and xylose in the medium and evaluating different cultivation conditions on EG and GA formation are required to enhance EG production in *K. phaffii* further. Indeed, the best EG production reported in the literature (108 g/L) was obtained in *E. coli* only after extensive metabolic engineering (deletion of competitive pathways and overexpression of several target genes) and bioprocess optimization with fed-batch cultivation (Table 3).

A characteristic shared between the EG-producing and control strains was the production of xylitol and arabitol (Fig. 6), metabolites usually accumulated during yeast cultivations due to redox imbalance. Xylitol is known to be produced from the reduction of xylose through the enzyme xylose reductase (XR), and its production is well-studied in several microorganisms (Qi et al. 2016; Carneiro et al. 2019; Xu et al. 2019). To date, xylitol production in native *K. phaffii* strains has been reported only by Heistinger et al. (2022). Arabitol was produced during the glucose assimilation phase (Fig. 6). Arabitol formation in a *K. phaffii* strain engineered for lactic acid production from glycerol was linked to cofactor regeneration (Melo et al. 2018). The cell requirement for NADH oxidation under limited oxygen situations leads the glycolytic flux to the pentose phosphate pathway, deviating the carbon flux to the ribulose or xylulose reduction into arabitol, resulting in NAD+regeneration (Melo et al. 2018). Similarly, the oxidation/reduction reactions in the EG pathway might lead to disturbances in the cell's redox state, leading to arabitol production.

In conclusion, this work represents a significant advance in the engineering of *K. phaffii* for biotechnological EG production from xylose, marking the first time this non-conventional yeast has been employed for this purpose. The introduction of more efficient xylonate dehydratases and the identification of challenges related to EG production are valuable insights. Future strategies should focus on eliminating competitive pathways and optimizing cultivation conditions to further improve EG yields. This study lays the groundwork for further optimization of *K. phaffii* as a microbial platform for industrial EG production.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13568-024-01795-0.

Supplementary Material 1

Acknowledgements

We thank CAPES and Embrapa for scholarships for C. V. G. C. C. Fundação de Apoio à Pesquisa do Distrito Federal (FAP-DF) under grant number 76/2021, CNPq and INCT Industrial Biotechnology financially supported this work.

Author contributions

C.V.G.C.C. Conceptualization, investigation, data analysis, writing. D.T. Conceptualization, investigation, writing—review and editing. J.C.B. Conceptualization, investigation. N.W. Writing—review and editing. T.W. Supervision, writing—review and editing. J.R.M.A. Conceptualization, supervision, writing—review and editing, funding acquisition. All authors reviewed the manuscript.

Funding

CNPq and FAP-DF.

Data availability

the datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request. Gene sequences employed in this study are listed in the Supplementary file.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication
Not applicable.

Competing interests
J.R.M.A., D.T., J. C. B., and C. V. G. C. C. are coauthors of a patent that describes the present work (BR 10 2023 005615). All remaining authors declare that they have no competing interests.

Received: 20 September 2024 / Accepted: 22 November 2024
Published online: 30 November 2024

References

Agilent Technologies (2008) pBluescript II phagemid vectors. Instruction Manual. www.stratagene.com

Alkim C, Cam Y, Trichez D, Auriol C, Spina L, Vax A, Bartolo F, Besse P, François JM, Walther T (2015) Optimization of ethylene glycol production from (d)-xylose via a synthetic pathway implemented in *Escherichia coli*. Microb Cell Fact 14(1):1–12. https://doi.org/10.1186/s12934-015-0312-7

Andberg M, Aro-Kärkkäinen N, Carlson P, Oja M, Bozonnet S, Toivari M, Hakulinen N, O'Donohue M, Penttilä M, Koivula A (2016) Characterization and mutagenesis of two novel iron–sulphur cluster pentonate dehydratases. Appl Microbiol Biotechnol 100(17):7549–7563. https://doi.org/10.1007/s00253-016-7530-8

Bañares AB, Nisola GM, Valdehuesa KNG, Lee WK, Chung WJ (2021) Understanding D-xylonic acid accumulation: a cornerstone for better metabolic engineering approaches. Appl Microbiol Biotechnol 105(13):5309–5324. https://doi.org/10.1007/s00253-021-11410-y

Betancur MO, Reis VCB, Nicola AM, De Marco JL, de Moraes LMP, Torres FAG (2017) Multicopy plasmid integration in *Komagataella phaffii* mediated by a defective auxotrophic marker. Microb Cell Fact 16(1):1–11. https://doi.org/10.1186/s12934-017-0715-8

Cabulong RB, Valdehuesa KNG, Ramos KRM, Nisola GM, Lee WK, Lee CR, Chung WJ (2017a) Enhanced yield of ethylene glycol production from D-xylose by pathway optimization in *Escherichia coli*. Enzyme Microb Technol 97:11–20. https://doi.org/10.1016/j.enzmictec.2016.10.020

Cam Y, Alkim C, Trichez D, Trebosc V, Vax A, Bartolo F, Besse P, François JM, Walther T (2016) Engineering of a synthetic metabolic pathway for the assimilation of (d)-Xylose into value-added chemicals. ACS Synth Biol 5(7):607–618. https://doi.org/10.1021/acssynbio.5b00103

Carneiro CVGC, E Silva FCP, Almeida JRM (2019) Xylitol production: Identification and comparison of new producing yeasts. Microorganisms 7(11). https://doi.org/10.3390/microorganisms7110484

Carneiro CVGC, Serra LA, Pacheco TF, Ferreira LMM, Brandão LTD, Freitas MNDM, Trichez D, Almeida JRMD (2022) Advances in *Komagataella phaffii* engineering for the production of renewable chemicals and proteins. Fermentation 8(11). https://doi.org/10.3390/fermentation8110575

Carniel A, Santos NF, Buarque FS, Resende JVM, Ribeiro BD, Marrucho IM, Coelho MAZ, Castro AM (2024) From trash to cash: current strategies for bio-upcycling of recaptured monomeric building blocks from poly(ethylene terephthalate) (PET) waste. Green Chem 6:5708–5743. https://doi.org/10.1039/D4GC00528G

Chae TU, Choi SY, Ryu JY, Lee SY (2018) Production of ethylene glycol from xylose by metabolically engineered *Escherichia coli*. American Institute of Chemical Engineers (AIChE). https://doi.org/10.1002/aic.16339

Choi SY, Kim WJ, Yu SJ, Park SJ, Im SG, Lee SY (2017) Engineering the xylose-catabolizing Dahms pathway for production of poly(d-lactate-co-glycolate) and poly(d-lactate-co-glycolate-co-d-2-hydroxybutyrate) in Escherichia coli. Microb Biotechnol 10(6):1353–1364. https://doi.org/10.1111/1751-7915.12721

Chomvong K, Bauer S, Benjamin DI, Li X, Nomura DK, Cate JHD (2016) Bypassing the pentose phosphate pathway: towards modular utilization of xylose. PLoS ONE 11(6):1–16. https://doi.org/10.1371/journal.pone.0158111

Costa PPKG, Mendes TD, Salum TFC, Pacheco TF, Braga SC, de Almeida JRM, Gonçalves SB, Damaso MCT, Rodrigues CM (2019) Development and validation of HILIC-UHPLC-ELSD methods for determination of sugar alcohols stereoisomers and its application for bioconversion processes of crude glycerin. J Chromatogr A 1589:56–64. https://doi.org/10.1016/j.chroma.2018.12.044

de Jong E, Stichnothe H, Bell G, Jørgensen H (2020) IEA bioenergy task 42—biobased chemicals: a 2020 update. http://task42.ieabioenergy.com/wp-content/uploads/2020/02/Bio-based-chemicals-a-2020-update-final-200213.pdf

Gasser B, Mattanovich D (2018) A yeast for all seasons– Is *Pichia pastoris* a suitable chassis organism for future bioproduction. FEMS Microbiol Lett 365(17):1–4. https://doi.org/10.1093/femsle/fny181

Hahn-Hägerdal B, Karhumaa K, Fonseca C, Spencer-Martins I, Gorwa-Grauslund MF (2007) Towards industrial pentose-fermenting yeast strains. Appl Microbiol Biotechnol 74:937–953. https://doi.org/10.1007/s00253-006-0827-2

Heistinger L, Dohm JC, Paes BG, Koizar D, Troyer C, Ata Ö, Steininger-Mairinger T, Mattanovich D (2022) Genotypic and phenotypic diversity among *Komagataella* species reveals a hidden pathway for xylose utilization. Microb Cell Fact 21(1). https://doi.org/10.1186/s12934-022-01796-3

Hou J, Qiu C, Shen Y, Li H, Bao X (2017) Engineering of *Saccharomyces cerevisiae* for the efficient co-utilization of glucose and xylose. FEMS Yeast Res 17(4). https://doi.org/10.1093/femsyr/fox034

Invitrogen (2006a) One shot® TOP10 electrocomp™

Invitrogen (2006b) Pichia pastoris expression system. Methods Mol Biol 72(1997):2–5

Invitrogen (2010) pGAPZ A, B, and C pGAPZ α A, B, and C Pichia expression vectors for constitutive expression and purification of recombinant proteins (Issue 25)

Isobe K, Michihiko Kataoka M, Ogawa J, Hasegawa J, Shimizu S (2012) Microbial oxidases catalyzing conversion of glycolaldehyde into glyoxal. New Biotechnol 29(2):177–182. https://doi.org/10.1016/j.nbt.2011.05.001

Jiang Y, Liu W, Cheng T, Cao Y, Zhang R, Xian M (2015) Characterization of D-xylonate dehydratase yjhG from *Escherichia coli*. Bioengineered 6(4):227–232. https://doi.org/10.1080/21655979.2015.1040208

Lee SS, Choi J, Il, Woo HM (2019) Bioconversion of xylose to ethylene glycol and glycolate in engineered *Corynebacterium glutamicum*. ACS Omega 4(25):21279–21287. https://doi.org/10.1021/acsomega.9b02805

Li P, Sun H, Chen Z, Li Y, Zhu T (2015) Construction of efficient xylose utilizing *Pichia pastoris* for industrial enzyme production. Microb Cell Fact 14(1). https://doi.org/10.1186/s12934-015-0206-8

Liu H, Ramos KRM, Valdehuesa KNG, Nisola GM, Lee WK, Chung WJ (2013a) Biosynthesis of ethylene glycol in *Escherichia coli*. Appl Microbiol Biotechnol 97(8):3409–3417. https://doi.org/10.1007/s00253-012-4618-7

Melo NTM, Mulder KCL, Nicola AM, Carvalho LS, Menino GS, Mulinari E, Parachin NS (2018) Effect of pyruvate decarboxylase knockout on product distribution using *Pichia pastoris* (*Komagataella phaffii*) engineered for lactic acid production. Bioengineering 5(1):1–13. https://doi.org/10.3390/bioengineering5010017

Paes BG, Steindorff AS, Formighieri EF, Pereira IS, Almeida JRM (2021) Physiological characterization and transcriptome analysis of *Pichia pastoris* reveals its response to lignocellulose-derived inhibitors. AMB Express 11(1). https://doi.org/10.1186/s13568-020-01170-9

Pereira B, Li ZJ, De Mey M, Lim CG, Zhang H, Hoeltgen C, Stephanopoulos G (2016) Efficient utilization of pentoses for bioproduction of the renewable two-carbon compounds ethylene glycol and glycolate. Metab Eng 34:80–87. https://doi.org/10.1016/j.ymben.2015.12.004

Prielhofer R, Barrero JJ, Steuer S, Gassler T, Zahrl R, Baumann K, Sauer M, Mattanovich D, Gasser B, Marx H (2017) GoldenPiCS: a Golden Gate-derived modular cloning system for applied synthetic biology in the yeast *Pichia pastoris*. BMC Syst Biol 11(1):1–14. https://doi.org/10.1186/s12918-017-0492-3

Qi XH, Zhu JF, Yun JH, Lin J, Qi YL, Guo Q, Xu H (2016) Enhanced xylitol production: Expression of xylitol dehydrogenase from *Gluconobacter oxydans* and mixed culture of resting cell. J Biosci Bioeng 122(3):257–262. https://doi.org/10.1016/j.jbiosc.2016.02.009

Ramos TGS, Justen F, Carneiro CVGC, Honorato VM, Franco PF, Vieira FS, Trichez D, Rodrigues CM, Almeida JRM (2021) Xylonic acid production by recombinant *Komagataella phaffii* strains engineered with newly identified xylose dehydrogenases. Bioresour Technol Rep 16. https://doi.org/10.1016/j.biteb.2021.100825

Rebnegger C, Vos T, Graf AB, Valli M, Pronk JT, Daran-Lapujade P, Mattanovicha D (2016) *Ppastorisstoris* exhibits high viability and a low maintenance energy requirement at near-zero specific growth rates. Appl Environ Microbiol 82(15):4570–4583. https://doi.org/10.1128/AEM.00638-16

Salusjärvi L, Toivari M, Vehkomäki ML, Koivistoinen O, Mojzita D, Niemelä K, Penttilä M, Ruohonen L (2017) Production of ethylene glycol or glycolic acid from D-xylose in *Saccharomyces cerevisiae*. Appl Microbiol Biotechnol 101(22):8151–8163. https://doi.org/10.1007/s00253-017-8547-3

Salusjärvi L, Havukainen S, Koivistoinen O, Toivari M (2019) Biotechnological production of glycolic acid and ethylene glycol: current state and perspectives. Appl Microbiol Biotechnol 103(6):2525–2535. https://doi.org/10.1007/s00253-019-09640-2

Senatore VG, Milanesi R, Masotti F, Maestroni L, Pagliari S, Cannavacciuolo C, Campone L, Serra I, Branduardi P (2024) Exploring yeast biodiversity and process conditions for optimizing ethylene glycol conversion into glycolic acid. FEMS Yeast Res Jan 9:foae024. https://doi.org/10.1093/femsyr/foae024

Stephen Dahms A (1974) 3-Deoxy-D-pentulosonic acid aldolase and its role in a new pathway of D-xylose degradation. Biochem Biophys Res Commun 60(4):1433–1439. https://doi.org/10.1016/0006-291X(74)90358-1

Stephens C, Christen B, Fuchs T, Sundaram V, Watanabe K, Jenal U (2007) Genetic analysis of a novel pathway for D-xylose metabolism in *Caulobacter crescentus*. J Bacteriol 189(5):2181–2185. https://doi.org/10.1128/JB.01438-06

Trichez D, Carneiro CVGC, Braga M, Almeida JRM (2022) Recent progress in the microbial production of xylonic acid. World J Microbiol Biotechnol 38(7):1–15. https://doi.org/10.1007/s11274-022-03313-5

Uranukul B, Woolston BM, Fink GR, Stephanopoulos G (2019) Biosynthesis of monoethylene glycol in *Saccharomyces cerevisiae* utilizing native glycolytic enzymes. Metab Eng 51:20–31. https://doi.org/10.1016/j.ymben.2018.09.012

Valdehuesa KNG, Ramos KRM, Nisola GM, Bañares AB, Cabulong RB, Lee WK, Liu H, Chung WJ (2018) Everyone loves an underdog: metabolic engineering of the xylose oxidative pathway in recombinant microorganisms. Appl Microbiol Biotechnol 102(18):7703–7716. https://doi.org/10.1007/s00253-018-9186-z

Vieira FS (2018) Desenvolvimento de métodos baseados em espectrometria de massas e cromatografia líquida para análise de compostos químicos produzidos por bioconversão de glicerina. https://repositorio.bc.ufg.br/tede/handle/tede/8883

Wang J, Wang X, Shi L, Qi F, Zhang P, Zhang Y, Zhou X, Song Z, Cai M (2017) Methanol-independent protein expression by AOX1 promoter with trans-acting elements engineering and glucose-glycerol-shift induction in *Pichia pastoris*. Sci Rep 7. https://doi.org/10.1038/srep41850

Wang Y, Xian M, Feng X, Liu M, Zhao G (2018) Biosynthesis of ethylene glycol from d-xylose in recombinant *Escherichia coli*. Bioengineered 9(1):233–241. https://doi.org/10.1080/21655979.2018.1478489

Xu Y, Chi P, Bilal M, Cheng H (2019) Biosynthetic strategies to produce xylitol: an economical venture. Appl Microbiol Biotechnol 103(13):5143–5160. https://doi.org/10.1007/s00253-019-09881-1

Yamada Y, Matsuda M, Maeda K, Mikata K (1995) The phylogenetic relationships of methanol-assimilating yeasts based on the partial sequences of 18S and 26S ribosomal RNAs: the proposal of Komagataella Gen. Nov. (*Saccharomycetaceae*). Biosci Biotechnol Biochem 59(3):439–444

Yang Z, Zhang Z (2018) Engineering strategies for enhanced production of protein and bio-products in *Pichia pastoris:* A review. Biotechnol Adv 36(1):182–195. https://doi.org/10.1016/j.biotechadv.2017.11.002

Zahrl RJ, Peña DA, Mattanovich D, Gasser B (2017) Systems biotechnology for protein production in *Pichia pastoris*. FEMS Yeast Res 17(7):1–15. https://doi.org/10.1093/femsyr/fox068

Zhang Y, Liu D, Chen Z (2017) Production of C2-C4 diols from renewable bioresources: new metabolic pathways and metabolic engineering strategies. Biotechnol Biofuels 10(1):1–20. https://doi.org/10.1186/s13068-017-0992-9

Zhang Z, Yang Y, Wang Y, Gu J, Lu X, Liao X, Shi J, Kim CH, Lye G, Baganz F, Hao J (2020) Ethylene glycol and glycolic acid production from xylonic acid by *Enterobacter cloacae*. Microb Cell Fact 19(1):1–16. https://doi.org/10.1186/s12934-020-01347-8

Publisher's note