



Universidade de Brasília

Instituto de Ciências Biológicas

Programa de Pós-Graduação em Biologia Microbiana

**Engenharia metabólica de *Komagataella phaffii* para produção de ácido glicólico a partir de biomassa lignocelulósica**

Túlio Marcos Godoy de Andrade

Brasília

2024

TÚLIO MARCOS GODOY DE ANDRADE

**Engenharia metabólica de *Komagataella phaffii* para produção de ácido glicólico a partir de biomassa lignocelulósica**

Dissertação apresentada ao Programa de Pós-Graduação em Biologia Microbiana da Universidade de Brasília para obtenção do título de mestre em Biologia Microbiana.

Orientador: Prof. Dr. João Ricardo Moreira de Almeida

Brasília

2024

## RESUMO

O ácido glicólico (AG) é um ácido orgânico produzido pela indústria petroquímica e largamente comercializado como cosmético para esfoliação da pele e rejuvenescimento. Além disso, é empregado na indústria têxtil para curtir e pigmentar tecidos, na indústria alimentícia como conservante e acidulante, e na fabricação de polímeros de relevância médica, como o poli(ácido glicólico) e o poli(ácido láctico-co-ácido glicólico). Atualmente, a produção do AG se dá por via química a partir de derivados do petróleo, o que gera impactos negativos ao meio ambiente e ao ser humano. Dessa maneira, o uso de biomassa lignocelulósica como substrato para a produção de AG em processos biológicos se apresenta como alternativa sustentável à rota química. Nesse contexto, foi construída pelo grupo de pesquisa uma linhagem de *K. phaffii* com capacidade de conversão de xilose a etilenoglicol (EG) e a AG por meio da introdução da via heteróloga de Dahms. Na levedura recombinante, foi constatado que a última etapa da via sintética, a conversão de glicaldeído a AG ou EG, era realizada respectivamente por aldeído redutases (ALDR) e aldeído desidrogenases (ALDH) nativas da levedura. Dessa forma, este trabalho tem por objetivo a otimização do processo de produção de AG por meio da caracterização dos genes envolvidos na última etapa da via, que poderão ser deletados ou superexpressos, de acordo com sua função, na levedura que expressa a via de Dahms, para obter maior produção de AG. Para isso, cinco genes putativos de ALDR's e ALDH's foram selecionados com base na similaridade de sequência a genes de referência para serem avaliados quanto à sua atividade sobre glicaldeído. Foram construídos cassetes de deleção para três desses genes, porém não foi possível identificar linhagens com os genes deletados após vários ciclos de transformação. Além disso, para a otimização dos parâmetros do processo de produção de AG, a linhagem de *K. phaffii* que expressa a via de produção de AG foi cultivada em biorreator em diferentes condições de pH e DO (oxigênio dissolvido), de acordo com um delineamento central composto rotacional (DCCR). A partir do DCCR, foi possível determinar que existe influência do pH e do DO para a produção de AG pela *K. phaffii* engenheirada. Por fim, a levedura engenheirada foi cultivada em biorreator com hidrolisado de biomassa de cana-de-açúcar como meio de alimentação para validar a produção de AG a partir dessa fonte de carbono. Foi possível alcançar a produção de  $8,51 \pm 0,23$  g/L de AG a partir de hidrolisado, valor superior aos descritos na literatura para leveduras. Os resultados demonstram estratégias genéticas e de processo para otimização do processo de produção de AG em *K. phaffii*.

**Palavras-chaves:** Ácido glicólico; Fermentação; *Komagataella phaffii*; *Pichia pastoris*; Engenharia metabólica

## ABSTRACT

Glycolic acid (GA) is an organic acid produced by the petrochemical industry and widely marketed as a cosmetic for skin exfoliation and rejuvenation. Additionally, it is employed in the textile industry for leather tanning and fabric dyeing, in the food industry as a preservative and acidulant, and in the manufacturing of medically relevant polymers, such as poly(glycolic acid) and poly(lactic-co-glycolic acid). Currently, GA is chemically produced from petroleum derivatives, which generates negative impacts on the environment and human health. Thus, the use of lignocellulosic biomass as a substrate for GA production through biological processes is presented as a sustainable alternative to the chemical route. In this context, this research group has constructed a *K. phaffii* strain with the capacity to convert xylose to ethylene glycol (EG) and GA through the introduction of the heterologous Dahms pathway. In the recombinant yeast, it was found that the last step of the synthetic pathway, the conversion of glycolaldehyde to GA or EG, was performed by native yeast aldehyde reductases (ALDR) and aldehyde dehydrogenases (ALDH), respectively. Thus, the objective of this work is to optimize the GA production process through the characterization of the genes involved in the last step of the pathway, that could be deleted or superexpressed, according to its function, in the yeast strain that expresses the Dahms pathway, for GA production optimization. For this purpose, five putative ALDR's and ALDH's genes were selected based on sequence similarity to reference genes to be evaluated for their activity on glycolaldehyde. Deletion cassettes for 3 of these genes were constructed, however it was not possible to identify strains with deleted genes after several transformation cycles. Furthermore, for the optimization of the GA production process parameters, the *K. phaffii* strain expressing the complete GA production pathway was cultured in a bioreactor under different pH and DO (dissolved oxygen) conditions, according to a central composite rotational design (CCRD). The CCRD enabled the determination of the influence of pH and DO conditions for GA production by the engineered *K. phaffii*. Lastly, the engineered yeast was cultivated in a bioreactor fed with sugarcane biomass hydrolysate to validate the production of GA from this carbon source. A production of  $8,51 \pm 0,23$  g/L of GA was achieved from hydrolysate, a value superior to those reported in literature for yeasts. These results highlight genetic and process strategies for the optimization of the GA production process in *K. phaffii*.

**Keywords:** Glycolic acid; Fermentation; *Komagataella phaffii*; *Pichia pastoris*; Metabolic engineering

## Lista de abreviações e símbolos

μ: Micro

Ω: Ohms

°C: Graus Celsius

x g: Força gravitacional

2DH3DX: 2-dehidro-3-deoxi-xilonato

AG: Ácido glicólico

ALDH: Aldeído desidrogenase

ALDR: Aldeído redutase

cm: Centímetros

DCCR: Delineamento composto central rotacional

DHAP: Dihidroxiacetona-fosfato

DNA: Ácido desoxirribonucleico

DR2-MazF-Hyg: Cassetes de seleção e contrasseleção

EG: Etilenoglicol

F: Farads

g: Gramas

h: Horas

HygR: Gene codificante da higromicina fosfotransferase de *Klebsiella pneumoniae*

L: Litros

LX1P: L-xilulose-1-fosfato

m: Metros

M: Molar

MazF: Gene codificante de endonuclease de *Escherichia coli*

min: Minutos

mL: Mililitros

nm: Nanômetros

OD<sub>600</sub>: Densidade ótica a 600 nm

P<sub>AOX1</sub>: Promotor da álcool oxidase 1 de *Komagataella phaffii*

P<sub>GAP</sub>: Promotor da gliceraldeído-3-fosfato desidrogenase de *Komagataella phaffii*

P<sub>LRA3</sub>: Promotor da L-ramnonato desidratase de *Komagataella phaffii*

P<sub>MDH3</sub>: Promotor da malato desidrogenase de *Komagataella phaffii*

p/v: Peso sobre volume

pb: Pares de bases

PCR: Reação em cadeia da polimerase

R1P: Ribulose-1-fosfato

rpm: Rotações por minuto

T<sub>m</sub>: Temperatura de anelamento

v/v: Volume sobre volume

X1P: Xilulose-1-fosfato

## Lista de Figuras

Figura 1. Composição geral da biomassa lignocelulósica .....	2
Figura 2. Diferentes vias metabólicas de assimilação de xilose presentes em microrganismos	3
Figura 3. Exemplos de produtos químicos gerados por via biotecnológica em <i>K. phaffii</i> .....	6
Figura 4. Estrutura química do ácido glicólico .....	7
Figura 5. Vias metabólicas engenheiradas para produção de ácido glicólico.....	11
Figura 6. Via metabólica expressa em <i>K. phaffii</i> para produção de etilenoglicol a partir de xilose .....	15
Figura 7. Representação esquemática da estratégia geral para estabelecimento de processo de produção de ácido glicólico a partir de biomassa lignocelulósica em <i>K. phaffii</i> .....	17
Figura 8. Módulos idealizados para deleção dos genes-alvo e estratégia geral de reciclagem dos cassetes de seleção e contrasseleção .....	27
Figura 9. Representação esquemática do protocolo de montagem dos módulos de deleção...29	
Figura 10. Representação esquemática do protocolo de montagem dos plasmídeos de expressão dos genes-alvo selecionados .....	32
Figura 11. Análise eletroforética em gel de agarose de PCR das sequências amplificadas para construção dos módulos de deleção de genes-alvo.....	41
Figura 12. Análise eletroforética em gel de agarose dos plasmídeos contendo as sequências para deleção dos genes-alvo e os cassetes de seleção e contrasseleção (DR2-MazF-Hyg).....	42
Figura 13. Análise eletroforética em gel de agarose dos plasmídeos contendo os módulos de deleção dos genes ALD2, ADH6 e Succinate ALDH .....	43
Figura 14. Análise eletroforética em gel de agarose de PCR de colônia para confirmação de integração do módulo de deleção no genoma da <i>K. phaffii</i> .....	45
Figura 15. Representação esquemática dos vetores (A) pPICK6 e (B) pMDHyg.....	47
Figura 16. Análise eletroforética em gel de agarose (A) do vetor pPICK6 digerido para liberação do gene KanR e (B) do vetor pPICK6 com o gene KanR removido purificado .....	47
Figura 17. Análise eletroforética em gel de agarose (A) da PCR para amplificação do gene HygR e (B) do vetor pMDHyg digerido por ensaio de restrição com a enzima <i>SacI</i> .....	48
Figura 18. Análise eletroforética em gel de agarose da PCR para amplificação das sequências dos genes-alvo para construção dos plasmídeos de expressão .....	49
Figura 19. Gráficos de superfície de resposta e de Pareto gerados no programa <i>Statistisc</i> a partir dos dados de produção de ácido glicólico obtidos em fermentação em biorreator com a linhagem <i>K. phaffii</i> JA122. ....	52

Figura 20. Gráficos representativos do perfil fermentativo da levedura *K. phaffii* JA122 em batelada alimentada com (A) hidrolisado de biomassa como meio de alimentação, e (B) meio sintético como meio de alimentação .....56



## Lista de Tabelas

Tabela 1. Composição polimérica de lignocelulose derivada de diferentes fontes .....	1
Tabela 2. Rendimento máximo teórico de ácido glicólico a partir de diferentes vias metabólicas e substratos.....	8
Tabela 3. Linhagens microbianas utilizadas neste trabalho.....	18
Tabela 4. Primers utilizados para amplificação de sequências por PCR.....	19
Tabela 5. Genes putativos codificantes de aldeído desidrogenases e aldeído redutases selecionados para deleção em <i>K. phaffii</i> .....	24
Tabela 6. Ciclos utilizados em PCR para amplificação de sequências de DNA .....	25
Tabela 7. Tamanho de fragmentos amplificados a partir do genoma de <i>K. phaffii</i> para construção de módulos de deleção.....	30
Tabela 8. Sítios de restrição utilizados para linearização dos plasmídeos e digestão das extremidades do módulo de deleção .....	31
Tabela 9. Codificação das variáveis pH e DO em níveis para construção do Delineamento Central Composto Rotacional.....	34
Tabela 10. Matriz do planejamento de ensaios para Delineamento Composto Central Rotacional .....	35
Tabela 11. Metodologias utilizadas para quantificação de compostos em solução por HPLC e UPLC .....	37
Tabela 12. Coluna e detector usados na quantificação de cada composto de interesse e seus respectivos tempos de retenção no método utilizado.....	37
Tabela 13. Concentração dos padrões analíticos utilizados para quantificação de compostos em HPLC e UPLC .....	38
Tabela 14. Concentrações de ácido glicólico e etilenoglicol obtidas em cada condição de cultivo estipulada por Delineamento Central Composto Rotacional .....	50

## Sumário

1.	Introdução .....	1
1.1.	Biomassa lignocelulósica.....	1
1.2.	Metabolismo de xilose em microrganismos .....	2
1.3.	<i>Komagataella phaffii</i> .....	4
1.4.	Ácido glicólico.....	7
1.4.1.	Vias biotecnológicas de produção de ácido glicólico .....	8
2.	Objetivo.....	16
2.1.	Objetivo geral.....	16
2.2.	Objetivos específicos .....	16
3.	Estratégia.....	16
4.	Material e métodos.....	17
4.1.	Materiais .....	17
4.1.1.	Linhagens .....	17
4.1.2.	Plasmídeos .....	18
4.1.3.	Oligonucleotídeos .....	18
4.1.4.	Meios de cultura.....	21
4.1.5.	Soluções .....	22
4.2.	Métodos.....	23
4.2.1.	Seleção de genes para ALDH e ALDR putativas para deleção e superexpressão 23	
4.2.2.	Reação em cadeia da polimerase (PCR) .....	24
4.2.3.	Preparo de células de <i>E. coli</i> eletrocompetentes .....	25
4.2.4.	Transformação de <i>E. coli</i> por eletroporação .....	25
4.2.5.	Eletroforese em gel de agarose .....	26
4.2.6.	Sequenciamento de fragmentos de DNA.....	26

4.2.7.	Montagem de módulos de deleção de genes-alvo.....	26
4.2.8.	Montagem de plasmídeos de superexpressão de genes-alvo .....	32
4.2.9.	Transformação de <i>K. phaffii</i> por eletroporação .....	33
4.2.10.	PCR de colônia de <i>K. phaffii</i> .....	34
4.2.11.	Otimização da produção de ácido glicólico pela <i>K. phaffii</i> JA122 em biorreator 34	
4.2.12.	Produção de ácido glicólico por <i>K. phaffii</i> JA122 a partir de hidrolisados de biomassa.....	35
4.2.13.	Análise de metabólitos por HPLC/UPLC .....	36
5.	Resultados e discussão .....	39
5.1.	Seleção de genes-alvo para deleção e superexpressão.....	39
5.2.	Construção dos módulos de deleção de genes-alvo .....	40
5.2.1.	Amplificação de sequências para deleção de genes-alvo .....	40
5.2.2.	Construção dos módulos de deleção de genes-alvo para <i>K. phaffii</i> .....	41
5.3.	Construção de linhagens recombinantes de <i>K. phaffii</i> com genes-alvo deletados 44	
5.4.	Construção de plasmídeos para superexpressão de genes-alvo .....	46
5.4.1.	Construção do vetor pMDHyg.....	46
5.4.2.	Construção de plasmídeos de superexpressão .....	48
5.5.	Otimização do processo de produção de ácido glicólico por <i>K. phaffii</i> JA122 ...	49
5.6.	Produção de ácido glicólico por <i>K. phaffii</i> JA122 a partir de hidrolisados de biomassa.....	54
6.	Conclusão.....	57
7.	Referências bibliográficas.....	59

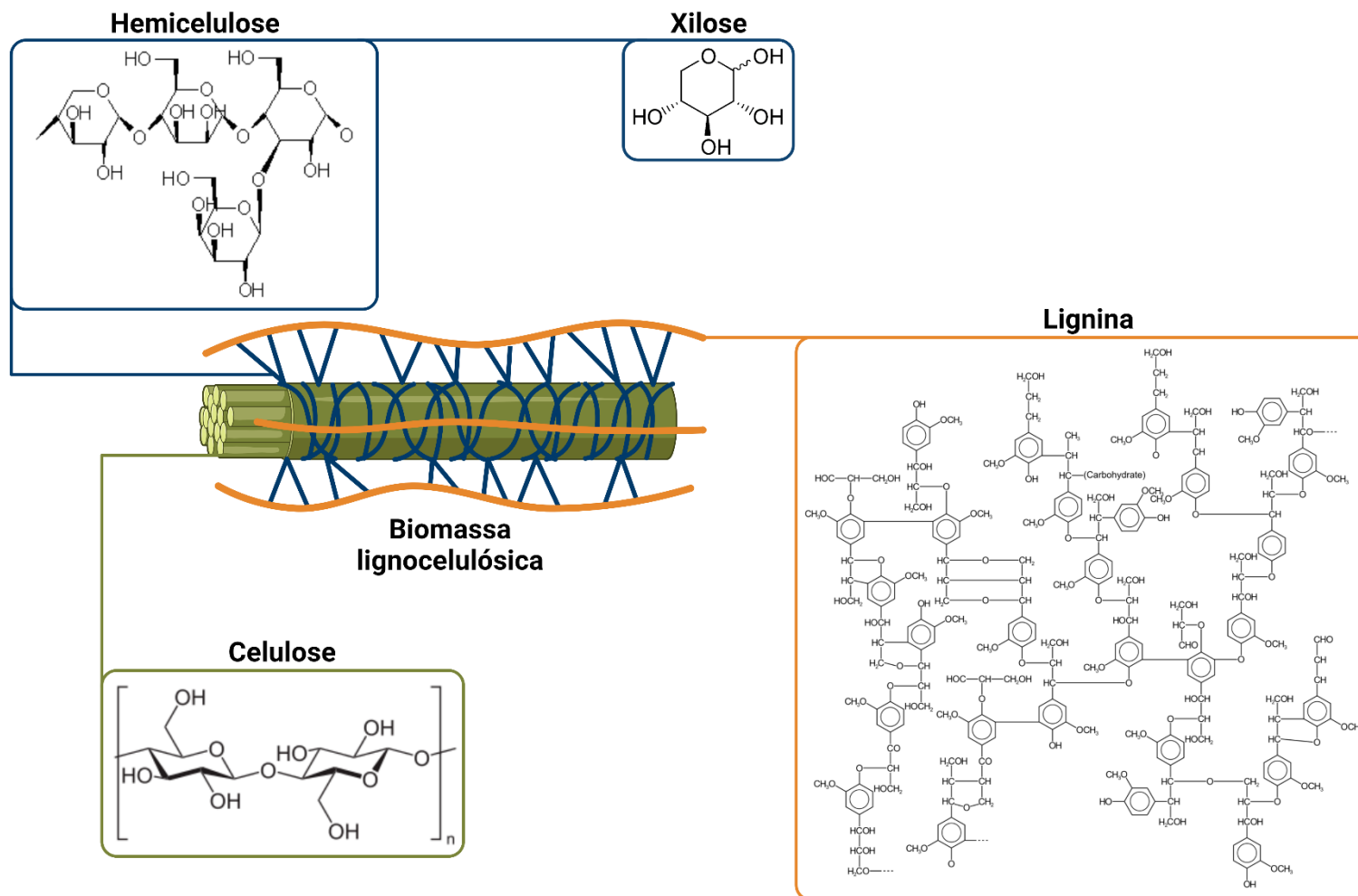
## **1. Introdução**

### **1.1. Biomassa lignocelulósica**

O rápido e constante processo de urbanização observado nas últimas décadas causou um aumento substancial no consumo de derivados da indústria petroquímica. Em 2023, a demanda global de petróleo atingiu 102,2 milhões de barris por dia, e estima-se um aumento de 4,69% nesse valor para o ano de 2024 (OPEC, 2024). No entanto, as reservas globais de fontes fósseis de carbono são limitadas, e por isso torna-se urgente o desenvolvimento de alternativas sustentáveis para a produção de compostos orgânicos derivados do petróleo. Além disso, a produção de diversos compostos orgânicos a partir do petróleo é realizada por processos químicos que utilizam reagentes e geram resíduos potencialmente danosos à saúde humana e ao meio ambiente (SHONZA et al., 2022; LIAO et al., 2022). Assim, mostra-se indispensável o desenvolvimento de alternativas sustentáveis para a produção de compostos orgânicos, que sejam menos prejudiciais ao ser humano e à natureza.

Nesse contexto, a biomassa lignocelulósica tem ganhado espaço enquanto substrato, uma vez que apresenta uma possibilidade de contornar os efeitos lesivos gerados pelo uso de fontes fósseis. Além disso, é a fonte de carbono renovável mais abundante do planeta, com uma produção anual de 200 bilhões de toneladas globalmente. No Brasil, para o ano de 2020, houve uma produção de 196,2 milhões de toneladas apenas de bagaço de cana-de-açúcar, o que demonstra o grande potencial do país para utilização da biomassa como substrato na indústria (ZHAO et al., 2022).

A lignocelulose possui uma complexa hierarquia estrutural constituída majoritariamente de celulose, hemicelulose e lignina, organizadas de maneira não-uniforme na estrutura tridimensional da biomassa (Figura 1) (DEIVAYANAI et al., 2022). A celulose é um homopolímero formado exclusivamente por D-glicose em ligações  $\beta$ -1,4-glicosídicas, e representa a maior parcela da biomassa vegetal (JAYAKUMAR et al., 2023). A hemicelulose é o segundo componente mais abundante da lignocelulose, e é formado por cadeias curtas e heterogêneas de açúcares diversos, como a D-xilose, a D-manose, a D-glicose e a L-arabinose (RAO et al., 2023). A lignina é uma macromolécula irregular e não cristalina formado por monômeros fenólicos, e constitui a menor fração da biomassa vegetal (ASHOKKUMAR et al., 2022). A proporção de cada um desses polímeros na biomassa varia de acordo com a estrutura da lignocelulose em cada espécie (Tabela 1).



**Figura 1.** Composição geral da biomassa lignocelulósica. As fibras em verde claro representam a celulose, recoberta pela hemicelulose, representada em azul. A lignina é representada em laranja. Fonte: Autor.

**Tabela 1.** Composição polimérica de lignocelulose derivada de diferentes fontes. (Adaptado de SARKER et al., 2021)

<b>Biomassa</b>	<b>Hemicelulose (% peso seco)</b>	<b>Celulose (% peso seco)</b>	<b>Lignina (% peso seco)</b>	<b>Referência</b>
Bagaçõ de cana- de-açúcar	15-25	40-50	15-25	PERRONE et al., 2021
Resíduo agrícola	25-50	35-50	15-25	BEIG et al., 2021
Espiga de milho	32	29	29	WINARSIH e SISKAWARDANI, 2020
Casca de milho	32	31	31	MENSAH et al., 2021
Folha de milho	10	41	31	MENSAH et al., 2021
Talo de milho	12	42	31	MENSAH et al., 2021
Palha de milho	21	38	31	MENSAH et al., 2021
Madeira dura	45-50	20-25	20-25	BEIG et al., 2021
Madeira macia	35-40	25-30	25-30	BEIG et al., 2021
Switchgrass	38	42	15	AMARASEKARA e DENG, 2019
Palha de trigo	23	39	12	TIAN et al., 2020

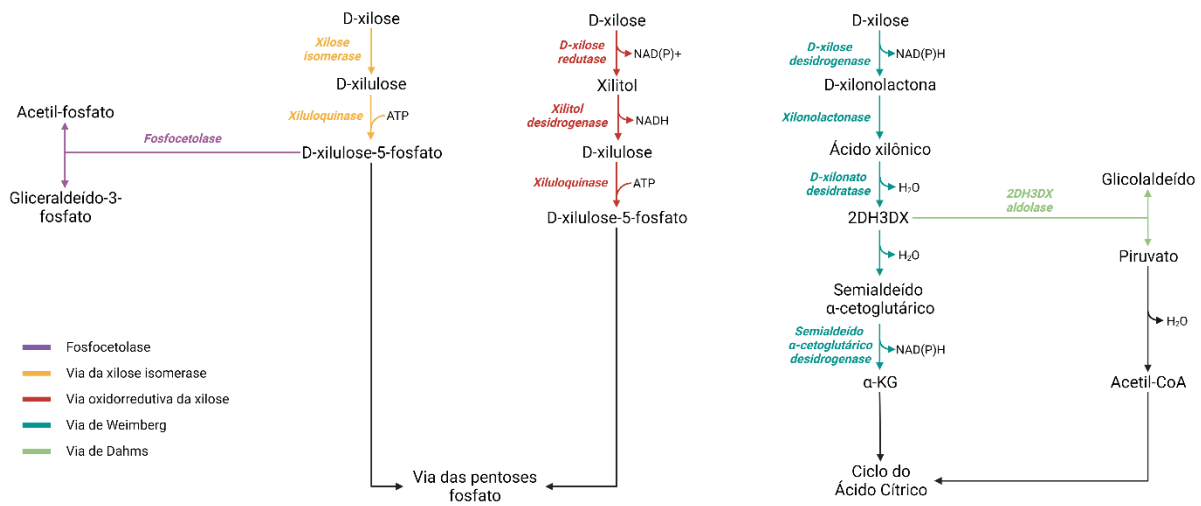
A celulose e a hemicelulose conjuntamente representam a maior parte da biomassa vegetal, o que faz com que os principais componentes monoméricos da lignocelulose sejam carboidratos, especialmente a D-glicose, constituinte de ambos os polímeros, e a D-xilose, presente na hemicelulose. A D-glicose é o açúcar mais abundante na lignocelulose, e compõe de 60 a 70% dos carboidratos presentes na biomassa pré-tratada, enquanto a D-xilose se enquadra como o segundo açúcar mais numeroso na estrutura, e compreende de 30 a 40% dos carboidratos disponíveis nos hidrolisados de biomassa (KIM, HWANG e LEE, 2022). Por isso, no desenvolvimento de processos de valorização da biomassa por via biológica, é importante que a D-xilose, que representa uma fração relevante dos carboidratos presentes na lignocelulose, seja assimilada de maneira satisfatória pelos microrganismos envolvidos no processo.

## **1.2. Metabolismo de xilose em microrganismos**

A D-xilose é uma pentose que tem ganhado cada vez mais atenção no desenvolvimento de processos fermentativos por conta de sua grande disponibilidade na biomassa lignocelulósica. Apesar de ser muito abundante na natureza, a capacidade de metabolizar a xilose não é ubíqua entre os microrganismos, e, quando presente, é comumente ineficiente quando comparada com o metabolismo da glicose. Além disso, a assimilação da xilose usualmente está sujeita à repressão catabólica pela glicose, tal que a hexose é preferencialmente consumida, o que reduz a produtividade de bioprocessos cujo substrato se caracteriza pela presença de ambos os açúcares, como a biomassa lignocelulósica (JAGTAP e RAO, 2018). Dessa forma, é importante explorar o metabolismo da xilose em microrganismos, para que se possa desenvolver novas linhagens robustas com capacidade de assimilação de xilose por métodos biotecnológicos.

Diferentes vias naturais de degradação da D-xilose já foram reportadas em microrganismos. De modo geral, em Eubacteria, como *E. coli*, *Corynebacterium* spp. e *Streptomyces* spp., a xilose é metabolizada pela via da xilose isomerase, em que a pentose é convertida a D-xilulose por uma isomerase. A D-xilulose é então fosforilada a D-xilulose-5-fosfato (X5P), que entra no metabolismo central do carbono (MCC) pela via das pentoses-fosfato (PPP) (Figura 2). Alternativamente, em bactérias como *Clostridium acetobutylicum* e

*Lactococcus lactis*, a X5P é clivada em acetil-fosfato e gliceraldeído-3-fosfato por uma fosfoctolase (ZHAO et al., 2020).



**Figura 2.** Diferentes vias metabólicas de assimilação de xilose presentes em microrganismos. 2DH3DX = 2-dehidro-3-deoxi-xilonato.  $\alpha$ -KG =  $\alpha$ -cetoglutarato. (Adaptado de ZHAO et al., 2020).

Já em fungos filamentosos e leveduras, a rota metabólica comumente encontrada para a assimilação da D-xilose é a via oxidorrredutiva, que consiste na redução do açúcar a xilitol por uma xilose redutase e oxidação do xilitol a D-xilulose-5-fosfato. A X5P, assim como em bactérias, entra no MCC pela PPP, porém já foram descritas espécies de leveduras oleaginosas com capacidade de clivagem da X5P a acetil-fosfato e gliceraldeído-3-fosfato (Figura 2). Algumas espécies, entretanto, possuem os genes necessários para a expressão da via, mas não apresentam taxas de metabolização significativas da D-xilose por conta de outros fatores, como baixa afinidade das enzimas pela D-xilose e baixo número de transportadores capazes de internalizar a pentose. Além disso, apesar da presença dessa via em fungos, as leveduras usualmente utilizadas em processos fermentativos, como *S. cerevisiae* e *Yarrowia lipolytica*, não conseguem crescer ou crescem limitadamente em D-xilose como fonte única de carbono (RUCHALA e SIBIRNY, 2021).

Uma terceira via não fosforilativa de degradação de pentoses ocorre em Archaea e Bacteria (FRANCOIS, ALKIM e MORIN, 2020). Nessa via, chamada de via de Weimberg, a D-xilose é inicialmente oxidada por uma D-xilose desidrogenase a D-xilonolactona, que é então hidratada a ácido xilônico por uma lactonase específica. O ácido xilônico é desidratado a 2-dehidro-3-deoxi-xilonato (2DH3DX) pela D-xilonato desidratase, seguido por uma



segunda desidratação, realizada pela 2-dehidro-3-deoxi-xilonato desidratase, que resulta em semialdeído- $\alpha$ -cetoglutárico. Por fim, esse intermediário é oxidado pela semialdeído- $\alpha$ -cetoglutárico a  $\alpha$ -cetoglutarato, que entra no MCC diretamente pelo ciclo do ácido cítrico (Figura 2). Foi demonstrada a existência de uma quarta via em bactérias, chamada de via de Dahms (Figura 2). Esta se diferencia da via de Weimberg pelas reações finais do metabolismo de xilose. Foi demonstrado em *Sulfolobus* sp. que o 2DH3DX é clivado por uma aldolase a piruvato e glicolaldeído. O piruvato entra diretamente no MCC, enquanto o glicolaldeído é incorporado ao MCC após ser convertido a glioxilato (DAHMS, 1974).

### 1.3. *Komagataella phaffii*

*Komagataella phaffii* é uma levedura aeróbica, não-patogênica e haplóide, que se reproduz por brotamento, e foi inicialmente isolada de exsudatos de plantas norte-americanas. O gênero *Komagataella* foi proposto em 1995, com o objetivo de reclassificar a levedura então conhecida como *Pichia pastoris* como um novo gênero entre as leveduras metilotróficas. Com o avanço da biologia molecular, várias espécies foram classificadas dentro do novo gênero, de acordo com diferenças em sequências de marcadores genéticos. Hoje, o gênero *Komagataella* abarca sete espécies, das quais as espécies *K. phaffii* e *K. pastoris* são melhor caracterizadas (HEISTINGER, GASSER e MATTANOVICH, 2020).

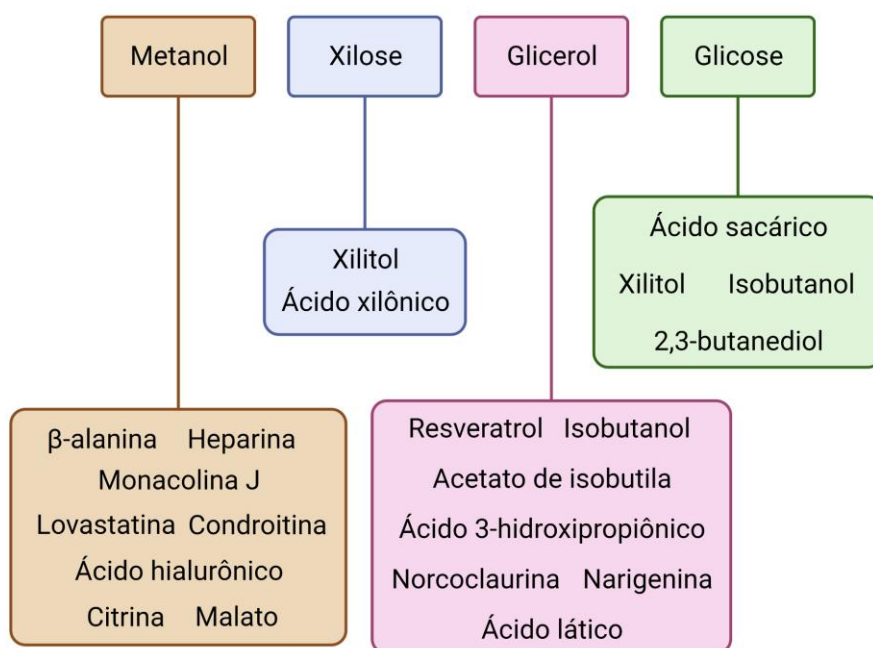
*K. phaffii* é uma levedura conhecidamente metilotrófica, capaz de metabolizar metanol como única fonte de carbono e energia. Também possui capacidade de crescer em diversos outros substratos, como glicose, glicerol, etanol, ácido láctico e ácido acético (ATA et al., 2021). Além de crescer facilmente em D-glicose, a *K. phaffii* não apresenta efeito *Crabtree* como *S. cerevisiae*, em que a levedura preferencia a via fermentativa da D-glicose em vez da respiração celular em altas concentrações do açúcar e na presença de oxigênio. Isso possivelmente ocorre por conta do grande número de transportadores de hexoses presentes nas leveduras *Crabtree*-positivas, o que permite a entrada de açúcar no ambiente celular a uma taxa maior do que a maquinaria respiratória consegue metabolizar. Assim, o excesso de carboidratos é desviado para a via fermentativa, o que diminui os rendimentos de biomassa e resulta em etanol como produto (ATA et al., 2018).

A levedura ainda se destaca pela capacidade de atingir alta densidade celular, de realizar modificações pós-traducionais eucarióticas, como a glicosilação, e de secretar biomoléculas e

proteínas com eficiência. Além disso, construções genéticas estáveis e um número cada vez maior de ferramentas estão disponíveis para engenharia genética da espécie (BARONE et al., 2023). Dentre as ferramentas de engenharia genética desenvolvidas para *K. phaffii*, estão kits para fácil construção de cassetes de expressão (PRIELHOFER et al., 2017), e para uso de sistema CRISPR/Cas9 na levedura (GASSLER et al., 2019). Uma das ferramentas genéticas mais exploradas em *K. phaffii* é o promotor do gene da álcool oxidase 1 ( $P_{AOX1}$ ), que atinge níveis de expressão muito elevados quando induzido por metanol (ERGÜN et al., 2020). Também foram descritos na literatura outros promotores constitutivos e induzidos de ocorrência natural na levedura, como o promotor da 3-fosfoglicerato quinase, o promotor da álcool desidrogenase 2 e o promotor de uma isoenzima da piruvato descarboxilase (DE ALMEIDA, DE MORAES e TORRES, 2005; FISCHER e GLIEDER, 2019; GAO, JIANG e LIAN, 2021).

Além disso, ao longo dos últimos anos, têm-se gerado cada vez mais dados genômicos e fisiológicos sobre *K. phaffii*, como seu genoma completo e vias metabólicas descritas na espécie (DE SCHUTTER et al., 2009; STURMBERGER et al., 2016; ZHA et al., 2023). Uma das vias descritas na levedura é a via de assimilação de xilose, pentose utilizada como fonte de energia por poucas leveduras descritas (HEISTINGER et al., 2022). Em *K. phaffii*, a xilose é assimilada pela via oxidorrredutiva, em que o açúcar é convertido, em três etapas, a xilulose-5-fosfato, que entra no metabolismo central do carbono por meio da via das pentoses-fosfato (Figura 2).

Por essas características, a *K. phaffii* tem sido frequentemente empregada como sistema de expressão de proteínas heterólogas, tal que mais de 5000 diferentes proteínas já foram expressas na espécie (ZHOU et al., 2023). A levedura também tem sido direcionada para a produção de diversos compostos químicos de valor industrial e comercial, como álcoois, ácidos orgânicos e vitaminas, a partir de diferentes substratos, incluindo hidrolisados lignocelulósicos (Figura 3) (CARNEIRO et al., 2022). Recentemente, foi descrita uma linhagem recombinante de *K. phaffii* com capacidade de conversão de xilose a ácido xilônico, ácido orgânico de relevância industrial e intermediário da via de produção de ácido glicólico (AG) (RAMOS et al., 2021).



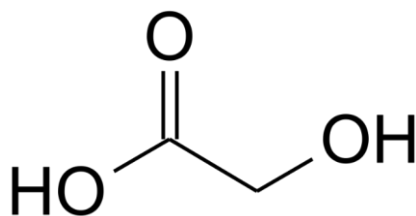
**Figura 3.** Exemplos de produtos químicos gerados por via biotecnológica em *K. phaffii*. (Adaptado de CARNEIRO et al., 2022).

No entanto, ainda é necessário o desenvolvimento de novas ferramentas genéticas, como promotores e plasmídeos, e a otimização de parâmetros críticos de fermentação para que a *K. phaffii* se estabeleça como um sistema de expressão robusto para a produção biotecnológica de compostos orgânicos de alto valor agregado (CARNEIRO et al., 2022). Atualmente, a maior parte do conhecimento fisiológico da levedura ainda se limita aos processos de produção de proteínas heterólogas, sendo necessário investir em esforços para caracterizar a levedura e desenvolver processos produtivos de metabólitos de interesse industrial e comercial.

Apesar disso, leveduras em geral apresentam maior tolerância a compostos inibitórios presentes no hidrolisado de biomassa lignocelulósica e a baixos pH's (TRICHEZ et al., 2022), e, diferentemente de *S. cerevisiae*, a *K. phaffii* é uma levedura de metabolismo respiratório, o que faz desse microrganismo um grande candidato para o desenvolvimento de um processo de produção de AG a partir de biomassa lignocelulósica. Nesse sentido, este trabalho busca o estabelecimento de um processo de produção de AG por *K. phaffii*, por meio do uso de recursos biotecnológicos. Para isso, objetiva-se a elucidação das enzimas nativas da levedura envolvidas na via de produção de AG e de etilenoglicol, e a compreensão de como parâmetros fermentativos têm influência sobre o processo de produção de AG.

#### 1.4. Ácido glicólico

O ácido glicólico (AG), também conhecido como ácido hidroxiacético, é a menor molécula da família dos  $\alpha$ -hidroxiácidos, caracterizados por possuírem um grupo hidroxila ligado ao carbono  $\alpha$  e um grupo carboxila na extremidade de sua estrutura (Figura 4). O AG é largamente utilizado como cosmético, uma vez que possui propriedades esfoliantes e promove o retardamento do envelhecimento e o rejuvenescimento da pele. Além disso, é empregado clinicamente no tratamento de acnes, fotoenvelhecimento, hiperpigmentação e outras condições dermatológicas (HOUSHMAND, 2021). Apesar de sua principal aplicação estar voltada para a indústria cosmética, o AG também é utilizado na indústria alimentícia, como conservante e acidulante, na indústria têxtil, como agente curtidor e pigmentador de tecidos, e na indústria de produtos de limpeza, como ativo higienizante (LACHAUX et al., 2019). Além disso, pode ser polimerizado em poli(ácido glicólico) (PGA) e, em conjunto com o ácido láctico, em poli(ácido láctico-co-ácido glicólico) (PLGA), que possuem diversas aplicações médicas por conta de sua biocompatibilidade e biodegradabilidade (BUDAK, SOGUT e AYEDEMIR SEZER, 2020; LU et al., 2023).



**Figura 4.** Estrutura química do ácido glicólico.

A grande gama de aplicações do AG tem levado a uma crescente demanda desse produto, com um mercado projetado para US\$531,5 milhões em 2027 (ZHOU et al., 2021). O AG é produzido em escala industrial por meio da carbonilação do formaldeído derivado do petróleo, ou pela reação entre o ácido cloroacético e o hidróxido de sódio. Ambas as reações fazem uso de reagentes tóxicos e geram resíduos poluentes, o que provoca efeitos lesivos sobre a saúde humana e o meio ambiente (LIAO et al., 2022). Uma rota biotecnológica de síntese de AG oferece uma alternativa sustentável e menos danosa ao ser humano e à natureza, e, por isso, seu desenvolvimento torna-se extremamente atrativo para a indústria.

### 1.4.1. Vias biotecnológicas de produção de ácido glicólico

O AG é naturalmente sintetizado por diversos microrganismos a partir de diferentes substratos. Várias bactérias e leveduras apresentam capacidade de produção de AG pela oxidação de etilenoglicol (EG) (KATAOKA et al., 2001). Também já foi relatada na literatura a geração de AG por hidrólise de glicolonitrila por um isolado de *Alcaligenes* sp. (HE et al., 2010). O estudo de ZHANG et al., 2020 relatou uma linhagem de *Enterobacter cloacae* capaz de produzir AG a partir de ácido xilônico. Além disso, algumas bactérias quimiolitotróficas oxidantes de ferro e enxofre apresentam capacidade nativa de síntese de AG, por via desconhecida (ÑANCUCHEO e JOHNSON, 2010). Todavia, não existem vias metabólicas naturais para síntese de AG a partir de matérias-primas renováveis e baratas, como carboidratos (SALUSJÄRVI et al., 2019). Por isso, diversas vias metabólicas naturais foram engenheiradas e vias sintéticas foram construídas para a produção de AG, a partir de diferentes substratos (Tabela 2).

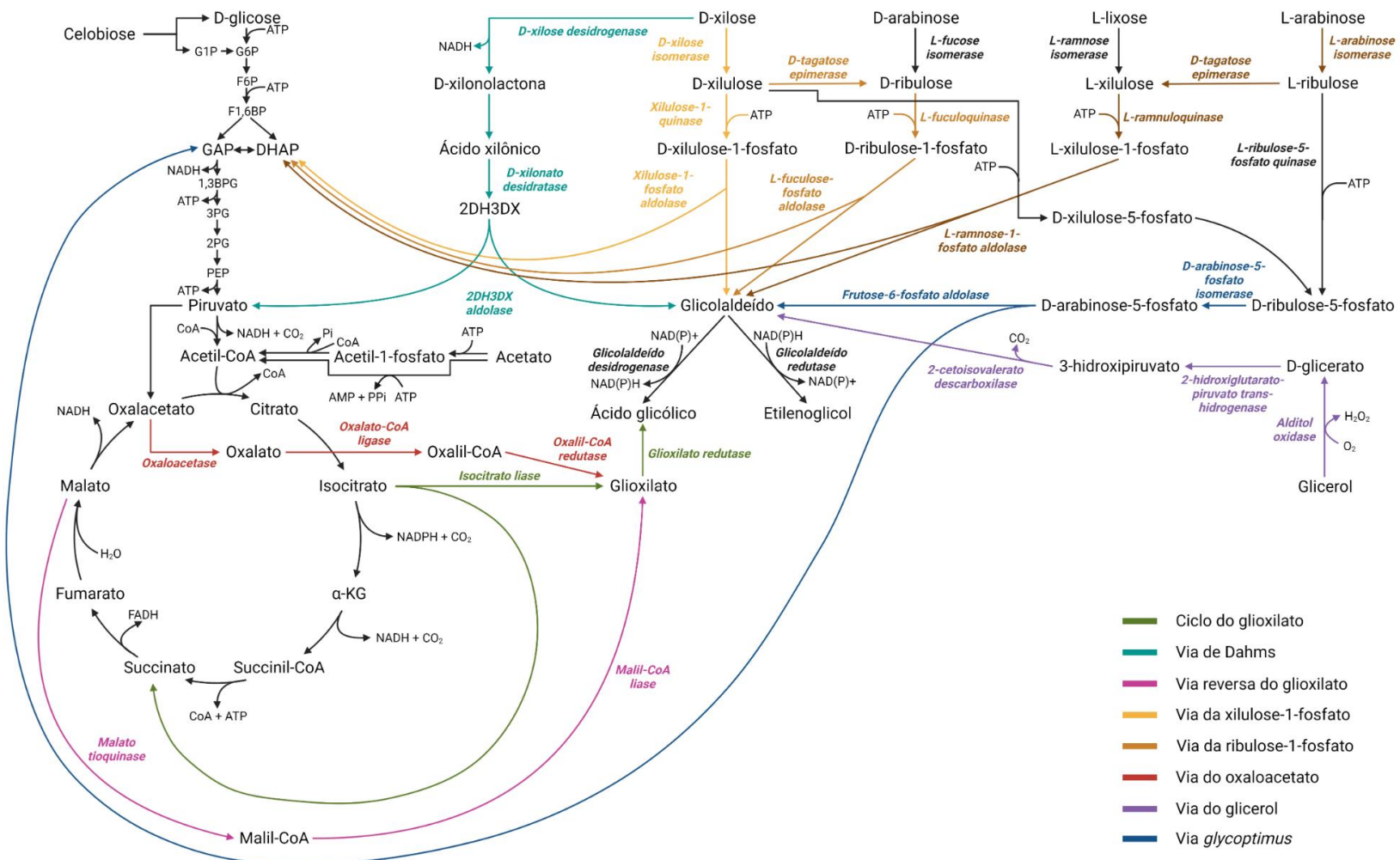
**Tabela 2.** Rendimento máximo teórico de ácido glicólico a partir de diferentes vias metabólicas e substratos. (Adaptado de SALUSJÄRVI et al., 2019 e LACHAUX et al., 2019)

Via metabólica	Organismo	Fonte de carbono	Rendimento máximo teórico		Referência
			mol/mol	g/g	
Ciclo do glioxilato	<i>E. coli</i>	D-glicose	2	0,84	DENG et al., 2018
		D-xilose	1,66	0,84	
	<i>K. lactis</i>	Etanol	1	1,65	KOIVISTOINEN et al., 2013
	<i>E. coli</i>	Acetato	1	1,27	YU et al., 2020
D-xilulose-1-fosfato	<i>E. coli</i>	D-xilose	1	0,51	CAM et al., 2016
D-ribulose-1-fosfato	<i>E. coli</i>	D-xilose	1	0,51	PEREIRA et al., 2016
L-xilulose-1-fosfato	<i>E. coli</i>	L-arabinose	1	0,51	PEREIRA et al., 2016
D-xilulose-1-fosfato + Ciclo do glioxilato	<i>E. coli</i>	D-xilose	2	1	ALKIM et al., 2016
		D-glicose	2	0,84	
D-ribulose-1-fosfato	<i>E. coli</i>	D-xilose	2	1	PEREIRA et al., 2016

+ Ciclo do glioxilato		D-glicose	2	0,84	
L-xilulose-1-fosfato + Ciclo do glioxilato	<i>E. coli</i>	L-arabinose	1	0,51	PEREIRA et al., 2016
Dahms	<i>S. cerevisiae</i>	D-xilose	1	0,51	SALUSJÄRVI et al., 2017
Dahms + Ciclo do glioxilato	<i>E. coli</i>	D-xilose	2	1	CABULONG et al., 2018
Dahms + Ciclo do glioxilato + Via reversa do glioxilato	<i>E. coli</i>	D-xilose	2	1	CABULONG et al., 2018
<i>Glycoptimus</i>	<i>E. coli</i>	D-xilose	2,5	1,27	LACHAUX et al., 2019
		D-glicose	3	1,27	
Oxaloacetato	<i>S. cerevisiae</i>	D-glicose	2	0,84	KAKKO et al., 2023
Glicerol	<i>E. coli</i>	Glicerol	1	0,83	ZHAN et al., 2021

A principal via metabólica engenheirada para a produção de AG é o ciclo do glioxilato, que ocorre naturalmente em vários organismos. Nessa via, o isocitrato do ciclo do ácido cítrico é clivado em glioxilato e succinato por ação da isocitrato liase, e o glioxilato é então reduzido a AG pela glioxilato redutase (Figura 5). A expressão e otimização dessa via para a produção de AG foi reportada para diversos microrganismos, dentre bactérias e leveduras (SALUSJÄRVI et al., 2019). A integração do ciclo do glioxilato ao metabolismo central do carbono torna essa via extremamente versátil, e permite o uso de diversos substratos para conversão ao produto de interesse. Por meio da superexpressão de enzimas-chaves do ciclo do glioxilato e da deleção de vias competitivas em *E. coli*, DENG et al., 2018 atingiram uma produção de 65,5 g/L de AG a partir de glicose, com um rendimento de 92,9% do rendimento teórico máximo da via. No entanto, os estudos que utilizaram o ciclo do glioxilato para a produção de AG demonstraram uma forte inibição da produção do ácido a partir de açúcares, possivelmente por conta da repressão do ciclo do ácido cítrico e do ciclo do glioxilato por carboidratos (KOIVISTOINEN et al., 2013; ZAHOOR, OTTEN e WENDISCH, 2014). Além disso, a preferência da glioxilato redutase por NADPH gera um desequilíbrio do balanço redox celular (LACHAUX et al., 2019).

Por conta das desvantagens do uso do ciclo do glioxilato, foram desenvolvidas vias metabólicas sintéticas para a produção de AG. A primeira via metabólica sintética construída para a produção de AG foi a via de Dahms (Figura 5), expressa em *E. coli*. Foi expressa na bactéria uma D-xilose desidrogenase advinda de *Caulobacter crescentus*, que conferiu ao microrganismo a capacidade de conversão de D-xilose a ácido xilônico. O ácido xilônico produzido é então convertido a 2DH3DX, que é clivado em piruvato e glicolaldeído, etapas realizadas pela maquinaria celular nativa da *E. coli*. A partir do glicolaldeído gerado pela via, é possível produzir EG por redução e AG por oxidação (LU et al., 2021). A via de Dahms é interessante para a produção de AG, já que atinge rendimentos muito próximos ao máximo teórico e possibilita a valorização de pentoses, como a D-xilose (CABULONG et al., 2021). Na levedura *S. cerevisiae*, a expressão heteróloga de uma D-xilose desidrogenase e de uma D-xilonato desidratase advindas de *C. crescentus*, e de uma aldolase e uma aldeído desidrogenase de *E. coli*, além da superexpressão de uma lactato desidrogenase nativa, levaram a levedura a uma produção de 1 g/L de AG. A hipótese levantada pelo grupo de pesquisa é de que a baixa eficiência teria se dado pela baixa atividade da D-xilonato desidratase, que levou a um acúmulo de D-xilonato, e pela perda de carbono na forma de 3-deoxipentonato (SALUSJÄRVI et al., 2017).



**Figura 5.** Vias metabólicas engenheiradas para produção de ácido glicólico. 2DH3DX = 2-dihidro-3-deoxi-xilonato. α-KG = α-cetoglutarato. (Adaptado de SALUSJÄRVI et al., 2019 e LACHAUX et al., 2019).



Outra estratégia explorada para a produção de AG foi a combinação da via de Dahms, do ciclo do glioxilato e de duas enzimas da via reversa do glioxilato (Figura 5), uma malato tioquinase e uma malil-CoA liase, para a síntese de AG em *E. coli* (CABULONG et al., 2018). A malato tioquinase converte malato a malil-CoA, que é clivado em acetil-CoA e glioxilato pela malil-CoA liase, e o glioxilato é então reduzido a AG. Após otimização da linhagem, foi observada uma produção de 4,57 g/L de AG, com um rendimento de 0,46 g/g a partir de D-xilose. Para o aumento da produção de AG, a linhagem foi otimizada para co-utilização de hexoses e pentoses sem que houvesse repressão catabólica da glicose sobre a assimilação de açúcares C5. Quando a glicose entra na célula, ela é fosforilada a glicose-6-fosfato e o sistema de fosfotransferase é desfosforilado, o que inibe a captação de pentoses. A linhagem foi construída por meio da expressão heteróloga de uma celobiose fosforilase de *Saccharophagus degradans*, que permitiu à *E. coli* a metabolização da celobiose. Dessa maneira, o dissacarídeo entra na célula sem ser fosforilado, e só é clivado no ambiente intracelular, o que aumenta a assimilação de açúcares C5. A linhagem produziu 4,27 g/L de AG em batelada alimentada, e foi capaz de gerar o produto de interesse a partir de uma combinação de hexoses e pentoses (CABULONG et al., 2021).

A via da xilulose-1-fosfato (X1P) foi implementada em *E. coli* para a produção de AG a partir de D-xilose (CAM et al., 2016). Nessa via, a D-xilose é convertida a D-xilulose por ação de uma xilose isomerase nativa, e a D-xilulose é então fosforilada a X1P por uma xilulose-1-quinase. A X1P é então clivada por uma aldolase a dihidroxiacetona-fosfato (DHAP) e glicaldeído, que é oxidado a AG, como ocorre com o glicaldeído produzido pela via de Dahms (Figura 5). Tanto a xilulose-1-quinase quanto a aldolase foram expressas heterologicamente na bactéria, e foi realizada a deleção de uma xilulose-5-quinase nativa responsável pela conversão de D-xilose a xilulose-5-fosfato. Assim, foram obtidos de 4,3 g/L de AG a partir de D-xilose, com um rendimento de 0,46 g/g, 90% do rendimento máximo teórico (CAM et al., 2016). A associação entre a via da X1P e o ciclo do glioxilato também foi testada em *E. coli*, porém não houve aumento da síntese de AG quando a linhagem foi cultivada exclusivamente em D-xilose, possivelmente por conta do lento crescimento nessa fonte de carbono. Entretanto, quando cultivada em glicose e xilose conjuntamente, foram obtidos 3,73 g/L de AG, com um rendimento de 0,63 g/g, 75% do rendimento teórico máximo (ALKIM et al., 2016).

A via da D-ribulose-1-fosfato (R1P) é uma via próxima à da X1P, e utiliza as vias catabólicas de L-lixose e D-arabinose nativas de *E. coli* para a produção de AG a partir de

arabinose e xilose (SALUSJÄRVI et al., 2019). Na via nativa, a D-arabinose é convertida a D-ribulose por uma L-fucose isomerase, e a D-ribulose é fosforilada a R1P pela L-fuculoquinase. A L-lixose é convertida a L-xilulose pela L-ramnose isomerase, e a L-xilulose é fosforilada a L-xilulose-1-fosfato (LX1P) pela L-ramnuloquinase. A R1P e a LX1P são então clivadas em DHAP e glicolaldeído por aldolases, como ocorre com a X1P, e o glicolaldeído pode então ser oxidado a AG. A expressão de uma D-tagatose epimerase heteróloga permite a assimilação de D-xilose e L-arabinose pela via, uma vez que a enzima converte a D-xilulose, produzida a partir de D-xilose, a D-ribulose, e a L-ribulose, produzida a partir de L-arabinose, a L-xilulose (Figura 5). Por meio dessa via, foi obtida uma produção de 44 g/L de AG em batelada, com rendimento de 0,44 g/g a partir de D-xilose. A síntese de AG paralelamente a partir da via da R1P e do ciclo do glioxilato com D-xilose como substrato atingiu um rendimento de 0,63 g/g, 63% do rendimento teórico máximo da via (PEREIRA et al., 2016).

Uma via de produção de AG a partir de oxaloacetato também foi descrita para fungos, em que são expressas de forma heteróloga as enzimas oxaloacetase de *Aspergillus niger*, oxalil-CoA redutase de *Methylobacterium extorquens*, e glioxilato redutase de *Arabidopsis thaliana*. Assim, o oxaloacetato produzido pelo ciclo do ácido cítrico é convertido pela oxaloacetase a oxalato, que é então convertido a oxalil-CoA por uma enzima nativa dos fungos. O oxalil-CoA é reduzido a glioxilato pela oxalil-CoA redutase, e o glioxilato pode ser então reduzido a AG pela glioxilato redutase (Figura 5) (TOIVARI et al., 2019). Por meio da expressão heteróloga do repressor do crescimento ClpXP de *E. coli* sob controle de um promotor induzido, foi possível direcionar a fonte de carbono para a produção de AG em uma linhagem de *S. cerevisiae* expressando a via do oxaloacetato. A levedura atingiu uma maior produção quando o crescimento foi reprimido após 8 horas de cultivo, e alcançou uma produção de 0,15 g/L de AG a partir de glicose (KAKKO et al., 2023).

Uma rota de síntese de AG conectada com a via das pentoses-fosfato também foi desenhada e chamada de via *glycoptimus* (Figura 5) (LACHAUX et al., 2019). Nessa via, a D-ribulose-5-fosfato da via das pentoses-fosfato é convertida por uma D-arabinose-5-fosfato isomerase a D-arabinose-5-fosfato, que é então clivada por uma aldolase a gliceraldeído-3-fosfato e glicolaldeído. O glicolaldeído é então oxidado a AG por uma glicolaldeído desidrogenase. As três enzimas-chaves dessa via ocorrem naturalmente em *E. coli*, e foram superexpressas para aumento da produção de AG. Além disso, várias vias competitivas foram deletadas na linhagem para diminuir a perda de carbono e direcionar o metabolismo celular

para a geração do produto de interesse. O rendimento de AG da linhagem a partir de D-glicose, D-xilose e D-arabinose foi de 0,2, 0,6 e 0,68 mol/mol, respectivamente, o que corresponde a 6,3%, 24% e 27% dos rendimentos máximos teóricos para cada açúcar.

Além das vias metabólicas supracitadas, têm sido desenvolvidas novas rotas de síntese de AG a partir de diversos substratos. ZHAN et al., 2020 desenvolveram uma nova rota sintética para produção de AG a partir da oxidação do glicerol em *E. coli*, e obtiveram uma produção de 4,74 g/L de AG em batelada alimentada. Também foi estabelecida uma via de geração de AG a partir de acetato em *E. coli*, baseada na via do glioxilato, porém com rendimento de 45,67% do rendimento teórico máximo (2,75 g/L de AG a partir de 4,39 g/L de acetato) (LI et al., 2019). Por meio da utilização simultânea de glicose e acetato como substratos, YU et al. 2020 conseguiram otimizar a síntese de AG a partir de acetato, e alcançaram uma concentração de 73,3 g/L de AG. Isso foi possível devido ao fato de que a oxidação da glicose supre a demanda de cofator NADPH necessária para a conversão do acetato, que serve de esqueleto de carbono para a síntese de AG.

Nesse contexto, foi desenvolvida previamente pelo grupo de pesquisa na Embrapa Agroenergia uma linhagem de *K. phaffii* que expressa a via de Dahms e uma aldeído redutase capaz de converter o glicaldeído resultante da via de Dahms a etilenoglicol (Figura 6). Para a construção da via, foram avaliadas diferentes xilose desidrogenases e xilose desidratases em combinação para determinar as enzimas com maior atividade de metabolização da D-xilose a ácido xilônico e a 2DH3DX, respectivamente. Uma das linhagens construídas (JA122), que expressa uma xilonato desidratase de atividade ainda não caracterizada, apresentou aumento significativo da produção de EG e AG em relação a linhagens que expressam as xilonato desidratases já descritas na literatura. Em cultivo em biorreator para validação da produção de AG pela linhagem, a levedura apresentou produção de 1,82 g/L de AG a partir de D-xilose, concentração superior às descritas na literatura em leveduras. Essa é a primeira linhagem engenheirada de *K. phaffii* que apresenta capacidade de produção de AG.



**Figura 6.** Via metabólica expressa em *K. phaffii* para produção de etilenoglicol a partir de xilose. Vermelho = enzimas expressas de maneira heteróloga; Azul = enzimas nativas da levedura. Fonte: Autor.

Em experimentos com a linhagem selvagem de *K. phaffii* X33 cultivada em glicolaldeído, foi observado que a levedura possui capacidade nativa de metabolização do glicolaldeído a AG e EG. Dessa forma, esse trabalho tem como objetivo a identificação dos genes nativos de aldeído desidrogenases (ALDH) e aldeído redutases (ALDR) de *K. phaffii* responsáveis pela oxidação e pela redução de glicolaldeído a AG e EG, respectivamente. Os genes de ALDR's que demonstrem atividade de conversão de glicolaldeído a EG serão deletados em uma linhagem de *K. phaffii* que expressa a via de Dahms, para desviar o

glicoaldeído produzido a partir de xilose para a produção de AG. Já os genes de ALDH's relevantes para a oxidação do glicoaldeído a AG poderão ser posteriormente superexpressos na linhagem construída, o que pode ter impacto positivo na produção do AG.

## **2. Objetivo**

### **2.1. Objetivo geral**

O objetivo deste trabalho é otimizar a produção de AG a partir de hidrolisado de biomassa de cana-de-açúcar por linhagens recombinantes de *K. phaffii*.

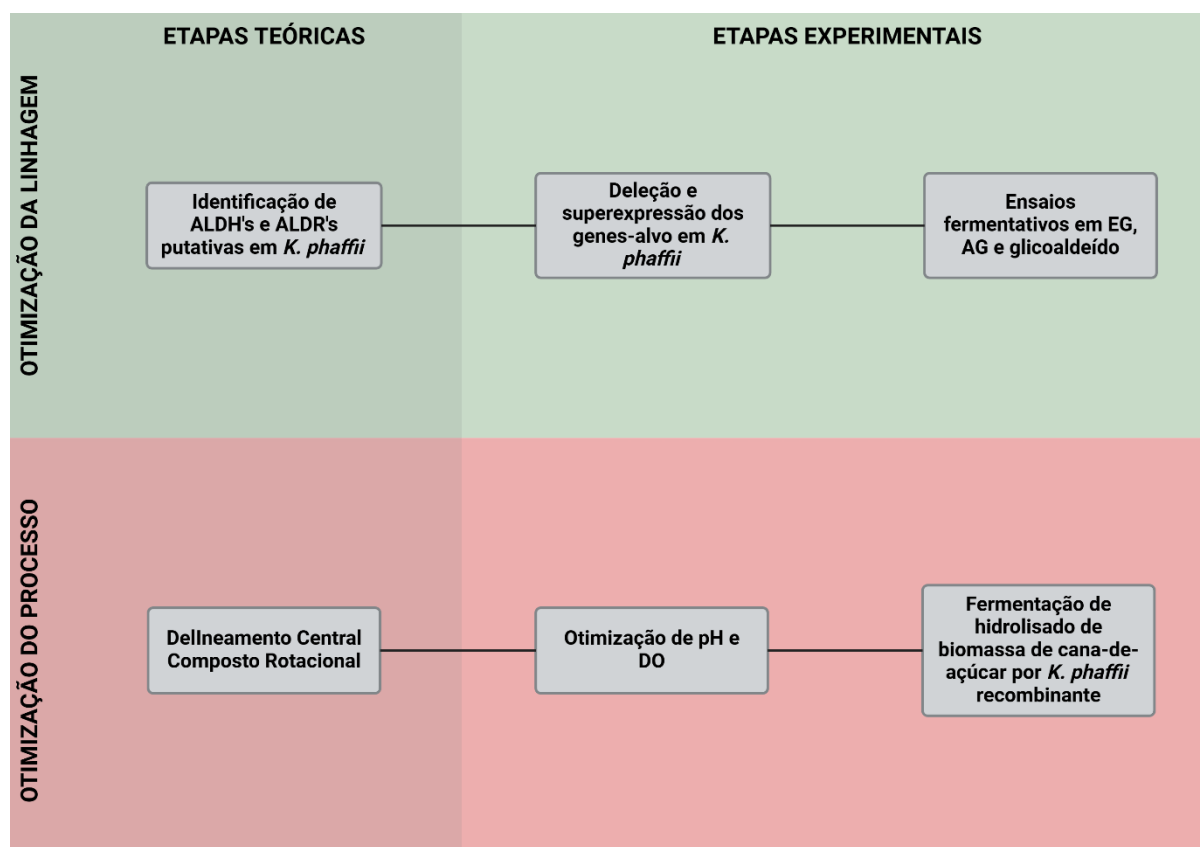
### **2.2. Objetivos específicos**

- Identificar, deletar e superexpressar genes putativos envolvidos com atividades de glicoaldeído desidrogenase (ALDH) e glicoaldeído redutase (ALDR) em *K. phaffii*;
- Avaliar o metabolismo de AG, EG e glicoaldeído em linhagens de *K. phaffii* com os genes-alvo deletados ou superexpressos;
- Avaliar o efeito de pH e oxigênio dissolvido na produção de ácido glicólico pela linhagem produtora de AG *K. phaffii* JA122;
- Estabelecer um processo de produção de ácido glicólico a partir de hidrolisado de biomassa de cana-de-açúcar.

## **3. Estratégia**

A estratégia geral deste trabalho está resumida na Figura 7. Para determinação de seus papéis funcionais, os genes putativos de ALDR's e ALDH's previamente identificados pelo grupo de pesquisa serão deletados do genoma da *K. phaffii* e superexpressos em cassetes integrativos, e a atividade de cada gene sobre a conversão de glicoaldeído a ácido glicólico (AG) ou etilenoglicol (EG) será avaliada por meio de ensaios fermentativos com as linhagens recombinantes na presença de glicoaldeído, AG e EG.

O processo de produção de AG pela linhagem mutante *K. phaffii* JA122 será otimizado por meio de um Delineamento Composto Central Rotacional (DCCR), que avaliará a interação entre pH e oxigênio dissolvido na produção do produto de interesse. Após a construção do modelo estatístico, as variáveis preditas como ótimas serão validadas por fermentação nas condições indicadas, com uso de hidrolisado de biomassa de cana-de-açúcar como fonte de carbono.



**Figura 7.** Representação esquemática da estratégia geral para estabelecimento de processo de produção de ácido glicólico a partir de biomassa lignocelulósica em *K. phaffii*. Fonte: Autor.

## 4. Material e métodos

### 4.1. Materiais

#### 4.1.1. Linhagens

As linhagens utilizadas neste trabalho são listadas na Tabela 3. A levedura *K. phaffii* X33 (Invitrogen) foi utilizada para a construção de linhagens recombinantes. Para clonagem e propagação de DNA plasmidial foram utilizadas as bactérias *E. coli* DH10B T1 (Invitrogen) e XL10-Gold (Agilent). Também foi utilizada a levedura *K. phaffii* JA122, construída

previamente pelo grupo de pesquisa, que expressa a via de Dahms, capaz de produzir glicoadaldeído e piruvato a partir de xilose, e uma aldeído redutase (ALDR), que converte o glicoadaldeído produzido a etilenoglicol (Figura 6).

**Tabela 3.** Linhagens microbianas utilizadas neste trabalho. XDH = Xilose desidrogenase. XD = Xilonato desidratase. ALDO = 2-dehidro-3-deoxi-xilonato aldolase. ALDR = Aldeído redutase.

<b>Linhagem</b>	<b>Genótipo</b>	<b>Fenótipo principal</b>
<i>E. coli</i> DH10B T1	F- <i>mcrA</i> $\Delta$ ( <i>mrr-hsdRMS-mcrBC</i> ) $\phi$ 80 $\Delta$ <i>lacZ</i> $\Delta$ M15 <i>lacX74 recA1 endA1</i> <i>araD139</i> $\Delta$ ( <i>ara-leu</i> )7697 <i>galU galK rpsL nupG</i> $\lambda$ - <i>tonA</i>	-
<i>E. coli</i> XL10-Gold	Tet <sup>r</sup> $\Delta$ ( <i>mcrA</i> )183 $\Delta$ ( <i>mcrCB-hsdSMR-mrr</i> )173 <i>endA1</i> <i>supE44 thi-1 recA1 gyrA96</i> <i>relA1 lac Hte</i> [F' <i>proAB</i> <i>lacI<sup>q</sup>Z</i> $\Delta$ M15 Tn10 (Tet <sup>r</sup> ) Amy Cam <sup>r</sup> ]	-
<i>K. phaffii</i> X33	Selvagem	-
<i>K. phaffii</i> JA122	pGAPZB(XDH) <i>_xylB-HL</i> + pKLD(XD) <i>_xylD-HL</i> + pB3Hyg_MOD1_ <i>YjhH+FucO</i>	Zeo <sup>R</sup> + Gen <sup>R</sup> + Hyg <sup>R</sup> , XDH+XD+ALDO+ALDR (via de produção de EG)

#### 4.1.2. Plasmídeos

Foi utilizado o vetor pGEM®-T Easy (Promega) para a construção dos módulos de deleção descritos neste trabalho. O plasmídeo original contendo os cassetes de seleção e contrasseleção foi doado pelo grupo de pesquisa do professor Hugo Costa Paes, da Faculdade de Medicina da Universidade de Brasília. O vetor pPICK6, construído por nosso grupo de pesquisa, foi utilizado para a construção do vetor pMDHyg, descrito neste trabalho.

#### 4.1.3. Oligonucleotídeos

Os oligonucleotídeos utilizados para amplificação de sequências por PCR neste trabalho foram sintetizados pelas empresas Invitrogen e Exxtend e são descritos na Tabela 4.

**Tabela 4.** Primers utilizados para amplificação de sequências por PCR. As regiões em negrito representam os sítios de restrição presentes nas sequências.

Primer	Sequência (5' → 3')	Temperatura de anelamento (°C)	Sítio de Restrição
DR2-MazF-F	TAG <b>GGATCCCGGGTCGAC</b> CCAC ACGGCATTTCCTCTCTAG	59,4	<i>Bam</i> HI (GGATCC), <i>Sma</i> I (CCCGGG) e <i>Sal</i> I (GTCGAC)
DR2-MazF-R	TAAG <b>GGATCCCGGGTCGAC</b> GTCT GTGGAAGCTAGCGGTACC	59,4	<i>Bam</i> HI (GGATCC), <i>Sma</i> I (CCCGGG) e <i>Sal</i> I (GTCGAC)
ALD2-XhoI-F	TAG <b>CTCGAG</b> TTCTGTCAATGGCC CTTGCA	57,4	<i>Xho</i> I
ALD2-XhoI-R	TAAC <b>CTCGAG</b> TCTGCAGACTCGGA AGTTGC	57,4	<i>Xho</i> I
Succinate ALDH- XhoI-F	CCG <b>CTCGAG</b> GGTACCCTGGCTTG TCTTCC	57,7	<i>Xho</i> I
Succinate ALDH- XhoI-R	TCG <b>CTCGAG</b> TCAGAGTCCTCCAA TCACCATG	56,3	<i>Xho</i> I
ADH2- NotI-F	ATTT <b>GCGGCCG</b> CATGTCTCCAAC TATCCCAACTAC	53,9	<i>Not</i> I
ADH2- NdeIXhoI-R	ATTCC <b>CATATGCTCGAG</b> ACAATCA ATCAGCGGTCCAC	55,5	<i>Nde</i> I (CATATG) e <i>Xho</i> I (CTCGAG)
ADH6- XhoI-F	TGAC <b>CTCGAG</b> CCGCAGCAACCCCA CTATAA	57,4	<i>Xho</i> I
ADH6- XhoI-R	TAAC <b>CTCGAG</b> CCACATTTCCCCCT CCCATC	58	<i>Xho</i> I
ADH900- NotI-F	TTGAG <b>GCGGCCG</b> CACGTTAGCGAC TCTTGCG	59	<i>Not</i> I
ADH900- SacIXhoI-R	TAAG <b>GAGCTCCTCGAG</b> GTCAGAC CACCAACCCTCAC	57,8	<i>Sac</i> I (GAGCTC) e <i>Xho</i> I (CTCGAG)
lnpMDH-F	GAATATACTACCCACCAGTTGAA	51,9	-
lnHyg-R	TTGCAAAGTAACACCTTGAGCT	55	-
M13-F	GTAAAACGACGGCCAGT	52,6	-
M13-R	CAGGAAACAGCTATGAC	47	-



HygR- pPICK6-F	GTATAATACGACAAGGTGAGGAA CTAAACCATGCCTGAATTGACTGC TAC	52,2	-
HygR- pPICK6-R	ATGTAAGCGTGACATAACTAATT ACATGATTTACTCTTTAGCTTAGG TC	47,2	-
ALD2- pMDHyg-F	GAACAATAACAACATGGGAGACC GATCTCGAGATGGCAAGTCCCTT AAGC	51,2	-
ALD2- pMDHyg-R	CAAACCTCAATGATGATGATGATG ATGGTCGACTTAATCTCTGGTCAG GTTAGTG	51,1	-
Succ- pMDHyg-F	GAACAATAACAACATGGGAGACC GATCTCGAGATGAATTCTCTCAGT AAAGCTG	50,1	-
Succ- pMDHyg-R	CAAACCTCAATGATGATGATGATG ATGGTCGACTCAGAGTCCTCCAAT CACC	52,2	-
ADH2- pMDHyg-F	GAACAATAACAACATGGGAGACC GATCTCGAGATGTCTCCAACATC CCAACACTAC	53,4	-
ADH2- pMDHyg-R	CAAACCTCAATGATGATGATGATG ATGGTCGACTTATTTGGAAGTGTC CACAACG	53,4	-
ADH6- pMDHyg-F	GAACAATAACAACATGGGAGACC GATCTCGAGATGGCGTACCCAGA CACC	55,7	-
ADH6- pMDHyg-R	CAAACCTCAATGATGATGATGATG ATGGTCGACCTATTTGAATGCCTT ATCGTAACC	52	-
ADH900- pMDHyg-F	GAACAATAACAACATGGGAGACC GATCTCGAGATGTCTGTGATGAA AGCCCTC	54,3	-
ADH900- pMDHyg-R	CAAACCTCAATGATGATGATGATG ATGGTCGACTTAATCGGCAGATA AGAAAATCTTG	51,9	-

---

#### 4.1.4. Meios de cultura

##### Meio LB (Luria-Bertani) low salt

Cloreto de sódio 0,5% (p/v)

Extrato de levedura 0,5% (p/v)

Peptona bacteriológica 1% (p/v)

Ágar bacteriológico 1,5% (p/v), quando sólido

Suplementação com ampicilina a 100 µg/mL, quando necessário

##### Meio YPD (Yeast Extract-Peptone-Dextrose)

Extrato de levedura 1% (p/v)

Peptona bacteriológica 2% (p/v)

D-(+)-glicose 2% (p/v)

Ágar bacteriológico 1,5% (p/v), quando sólido

Suplementação com higromicina B a 200 µg/mL, zeocina a 100 µg/mL e/ou geneticina a 200 µg/mL quando necessário

##### Meio YPD<sub>X</sub> (Yeast Extract-Peptone-Dextrose-Xylose)

Extrato de levedura 1% (p/v)

Peptona bacteriológica 2% (p/v)

D-(+)-glicose 2% (p/v)

D-(+)-xilose 4% (p/v)

### Meio FM22

D-(+)-glicose 2% (p/v)

D-(+)-xilose 4% (p/v)

Fosfato de potássio monobásico ( $\text{KH}_2\text{PO}_4$ ) 4,29% (p/v)

Sulfato de amônio ( $(\text{NH}_4)_2\text{SO}_4$ ) 0,5% (p/v)

Cloreto de cálcio dihidratado ( $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ ) 0,2% (p/v)

Sulfato de potássio ( $\text{K}_2\text{SO}_4$ ) 1,43% (p/v)

Sulfato de magnésio heptahidratado ( $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ ) 1,17% (p/v)

PTM4 0,435% (v/v)

### **4.1.5. Soluções**

#### PTM4

Sulfato de cobre (II) pentahidratado ( $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ ) 0,2% (p/v)

Iodeto de sódio (NaI) 0,008% (p/v)

Sulfato de manganês (II) monohidratado ( $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ ) 0,3% (p/v)

Molibdato de sódio dihidratado ( $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$ ) 0,02% (p/v)

Ácido bórico ( $\text{H}_3\text{BO}_3$ ) 0,002% (p/v)

Sulfato de cálcio dihidratado ( $\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$ ) 0,05% (p/v)

Cloreto de cobalto (II) ( $\text{CoCl}_2$ ) 0,05% (p/v)

Cloreto de zinco ( $\text{ZnCl}_2$ ) 0,7% (p/v)

Sulfato de ferro (II) heptahidratado ( $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ ) 2,2% (p/v)

Biotina 0,02% (p/v)

Ácido sulfúrico (H<sub>2</sub>SO<sub>4</sub>) 0,1% (v/v)

Solução de sorbitol 1M

D-Sorbitol 18,22% (p/v)

Solução de ácido sulfúrico 5 mM

Ácido sulfúrico (H<sub>2</sub>SO<sub>4</sub>) 0,0272% (v/v)

Solução de glicerol 10%

Glicerol 10% (v/v)

Solução de acetato de sódio 3M

Acetato de Sódio (NaO<sub>2</sub>CCH<sub>3</sub>) 24,61% (p/v)

Solução de etanol 70%

Etanol (CH<sub>3</sub>CH<sub>2</sub>OH) 70% (v/v)

## **4.2. Métodos**

### **4.2.1. Seleção de genes para ALDH e ALDR putativas para deleção e superexpressão**

Os genes-alvo para deleção e superexpressão foram previamente selecionados com base na similaridade das sequências com genes de referência de outros organismos. Foram selecionadas 5 oxidoredutases nativas de *K. phaffii*, nem todas com atividade bioquímica completamente caracterizada. Três dos genes selecionados possuem atividade putativa de

aldeído redutase (ALDR) e dois apresentam atividade putativa de aldeído desidrogenase (ALDH). Para os genes das ALDR's, foi utilizada como referência a sequência da ADH1 de *S. cerevisiae*, enquanto para os genes das ALDH's, foram usadas as sequências da AldA de *E. coli* e da ALD4 de *S. cerevisiae*. Os genes selecionados são descritos na Tabela 5.

**Tabela 5.** Genes putativos codificantes de aldeído desidrogenases e aldeído redutases selecionados para deleção em *K. phaffii*.

Gene ID	Código do gene	Nome do gene	Descrição do gene
8198920	PAS_chr2-1_0453	ALD2	Aldeído desidrogenase citoplasmática, envolvida na oxidação de etanol e na biossíntese de beta-alanina
8197957	PAS_chr1-3_0024	Succinate ALDH	Succinato semialdeído desidrogenase
8198395	PAS_chr2-1_0472	ADH2	Álcool desidrogenase mitocondrial isoenzima III
8200841	PAS_chr4_0576	ADH6	Álcool desidrogenase NADPH-dependente de cadeia média com ampla especificidade de substrato
8199001	PAS_chr2-1_0313	ADH900	Enzima bifuncional com álcool desidrogenase e formaldeído desidrogenase dependente de glutatona

#### 4.2.2. Reação em cadeia da polimerase (PCR)

Para todas as reações de PCR realizadas, foram utilizados os reagentes do kit GoTaq® Hot Start Polymerase (Promega), nas concentrações indicadas pelo fabricante, e os primers correspondentes de cada sequência. As reações de PCR foram feitas em termociclador, de acordo com os ciclos indicados na Tabela 6. Para as amplificações realizadas a partir de DNA genômico, foram utilizados 30 ng de DNA molde, enquanto que para aquelas realizadas a partir de material plasmidial, foi utilizado 1-3 ng de DNA molde.

**Tabela 6.** Ciclos utilizados em PCR para amplificação de sequências de DNA. T<sub>m</sub> = Temperatura de anelamento otimizada para cada reação de amplificação.

Etapa 1		Etapa 2 (x30)		Etapa 3	
95°C	95°C	T <sub>m</sub>	72°C	72°C	10°C
3'	30''	30''	2'	10'	∞

#### 4.2.3. Preparo de células de *E. coli* eletrocompetentes

Foi plaqueada a linhagem de *E. coli* DH10B T1 em meio LB low salt a partir de estoque em freezer -80°C, e as placas foram incubadas a 37°C por 16h. Uma colônia isolada foi inoculada em 10 mL de meio LB low salt e incubada em shaker a 37°C e 200 rpm por 16h. Uma alíquota do pré-inóculo, entre 2 e 5 mL, foi inoculada em 4 erlenmeyers de 1 L contendo 200 mL de meio LB low salt, que foram incubados em shaker a 37°C e 200 rpm até atingir OD<sub>600</sub> ≅ 0,6. Quando atingiram a OD<sub>600</sub> especificada, os inóculos foram transferidos para 4 tubos de ultracentrífuga de 250 mL e foram incubados em gelo por aproximadamente 40 min. As células foram então centrifugadas a 2000 x g e 8°C por 20 min, o sobrenadante foi descartado, e cada pellet foi ressuscitado em 100 mL de água ultrapura estéril gelada. O conteúdo de 2 tubos foi misturado, e as células foram novamente centrifugadas sob as mesmas condições por 15 min. Após a centrifugação, o sobrenadante foi descartado, e as células foram ressuscitadas em 50 mL de glicerol 10% estéril gelado e transferidas para 4 tubos falcon de 50 mL, com 25 mL cada. Os tubos foram centrifugados sob as mesmas condições por 15 min, o sobrenadante foi descartado, e cada pellet foi ressuscitado em 12,5 mL de glicerol 10%. O conteúdo de 2 falcons foi misturado, e foram novamente centrifugados sob as mesmas condições por 15 min. O sobrenadante foi descartado e cada pellet foi ressuscitado em 1 a 2 mL de glicerol 10% gelado, para atingir OD<sub>600</sub> final de 150 a 200. As bactérias eletrocompetentes foram aliqüotadas em tubos de 0,2 mL contendo 50 µL cada, congeladas em nitrogênio líquido, e posteriormente armazenadas em freezer -80°C.

#### 4.2.4. Transformação de *E. coli* por eletroporação

Um volume de 1 a 5 µL dos sistemas de ligação foi adicionado a uma alíquota de 50 µL de células de *E. coli* DH10B T1 eletrocompetentes. As células foram transferidas para cubeta de eletroporação (0,1 cm), em que foram submetidas a descarga elétrica de voltagem

1800 V, capacitância de 25  $\mu$ F e resistência de 200  $\Omega$ . Foi imediatamente adicionado 1 mL de meio LB low salt às células, que foram incubadas por uma hora a 37°C após a eletroporação, e foram então plaqueadas em meio LB low salt sólido suplementado com ampicilina 100  $\mu$ g/mL. As placas foram incubadas a 37°C por 16 horas.

#### **4.2.5. Eletroforese em gel de agarose**

Análises eletroforéticas de sequências de DNA foram realizadas em gel de agarose 0,8%, preparado com tampão TAE 1X e brometo de etídio 0,5  $\mu$ g/mL. O tampão TAE 1X foi preparado por meio de diluição do reagente UltraPure® DNA Typing Grade® 50X TAE Buffer (Invitrogen). As amostras de DNA foram aplicadas ao gel diluídas em tampão Gel Loading Buffer (Sigma-Aldrich) e submetidas a corrente elétrica de voltagem compatível com o volume do gel. Após a aplicação de corrente elétrica, as amostras foram visualizadas em transiluminador de luz ultravioleta.

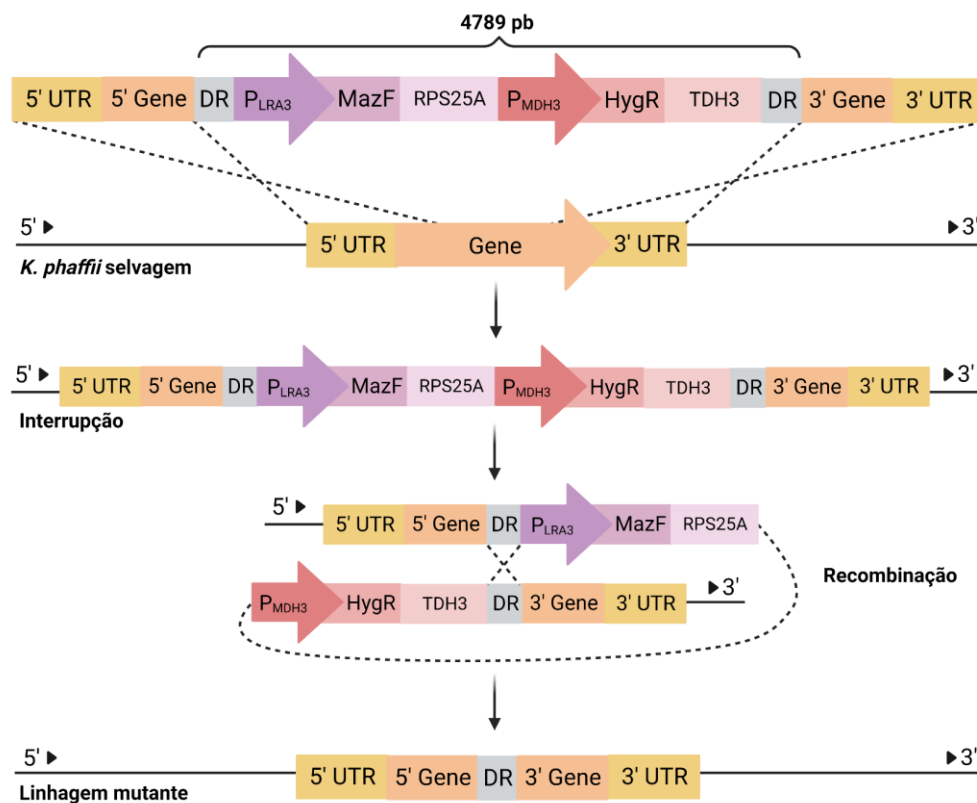
#### **4.2.6. Sequenciamento de fragmentos de DNA**

A correta inserção dos genes-alvo no vetor de clonagem foi confirmada por sequenciamento. As sequências para deleção dos genes-alvo, inseridas no vetor de clonagem pGEM®-T Easy, foram amplificadas por PCR utilizando os primers M13-F e M13-R, que se anelam na sequência do vetor. O sequenciamento foi feito pela empresa Clinilab Laboratório Veterinário.

#### **4.2.7. Montagem de módulos de deleção de genes-alvo**

Os módulos de deleção foram idealizados com base em metodologia que permite reciclagem da marca de seleção, para evitar acúmulo de marcas de seleção na linhagem de *K. phaffii* recombinante final, modificada para a expressão da via de Dahms e para a deleção de genes do genoma. Essa metodologia consiste na construção de um módulo de deleção que apresenta uma região destinada à seleção e à contrasseleção de linhagens recombinantes, flanqueada por braços de homologia de cerca de 600 pb, que permitem a inserção do módulo no genoma por recombinação homóloga. Os braços de homologia foram desenhados com

tamanhos que aumentam a eficiência de recombinação homóloga, uma vez que a *K. phaffii* possui como mecanismo de reparo preferencial a união de extremidade não-homóloga (ITO et al., 2018). Além disso, os cassetes de seleção e contrasseleção são flanqueados por duas regiões homólogas entre si (DR), para permitir recombinação e consequente excisão dos cassetes do genoma, o que permite sua reutilização (Figura 8). O cassete de seleção é constituído pelo gene da higromicina fosfotransferase (HygR) sob controle do promotor constitutivo da malato desidrogenase peroxissomal ( $P_{MDH3}$ ) de *K. phaffii* (PRIELHOFER et al., 2017). O cassete de contrasseleção é constituído pelo gene MazF de *E. coli*, que codifica uma endonuclease tóxica para *K. phaffii*, sob controle do promotor da L-ramnonato desidratase ( $P_{LRA3}$ ) (LIU et al., 2016; JIAO et al., 2019), induzível por ramnose. Assim, as linhagens recombinantes de *K. phaffii* podem ser selecionadas com base no cassete de seleção, que garante resistência à higromicina B, e, após a indução da recombinação entre as regiões DR (conforme tópico 4.2.17), a indução do gene MazF por ramnose pode ser utilizada para contrasselecionar as células da levedura que não tenham realizado o processo de recombinação entre as regiões homólogas DR.

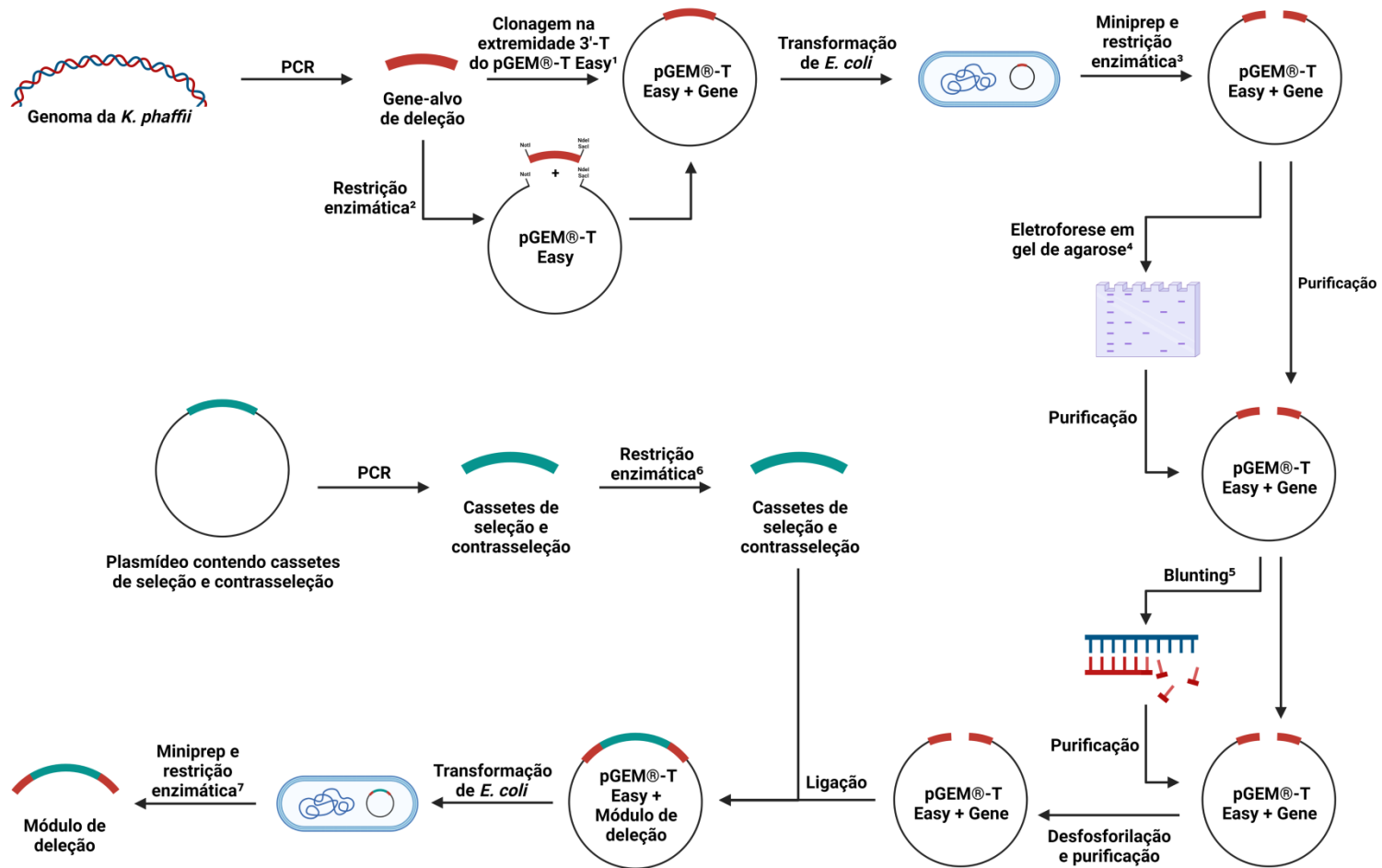


**Figura 8.** Módulos idealizados para deleção dos genes-alvo e estratégia geral de reciclagem dos cassetes de seleção e contrasseleção. DR = Regiões idênticas em repetição direta.  $P_{LRA3}$  = Promotor da L-ramnonato desidratase de *Komagataella phaffii*. MazF = Gene codificante de endonuclease de *Escherichia coli*. RPS25A = Terminador do gene da proteína S25 da subunidade menor do ribossomo de *Komagataella phaffii*.  $P_{MDH3}$  = Promotor da malato desidrogenase de *Komagataella phaffii*. HygR = Gene codificante da higromicina



fosfotransferase de *Klebsiella pneumoniae*. TDH3 = Terminador do gene da isozima 3 da gliceraldeído 3-fosfato de *Komagataella phaffii*. Fonte: Autor.

A estratégia para construção dos módulos de deleção é resumida na Figura 9. Para a construção dos módulos, as sequências para deleção dos genes-alvo foram amplificadas por reação de PCR do genoma da levedura *K. phaffii* X33 e clonados no vetor pGEM®-T Easy (Promega). A sequência dos cassetes de seleção e contrasseleção, denominada DR2-MazF-Hyg, foi amplificada a partir de plasmídeo doado pelo grupo de pesquisa do professor Hugo Paes, e clonada no mesmo vetor. O tamanho das sequências amplificadas se encontra na Tabela 7. Para as clonagens no plasmídeo pGEM®-T Easy, os produtos de PCR foram ligados com T4 DNA ligase ao vetor (Thermo Scientific), de acordo com as orientações do fabricante, e posteriormente transformados em *E. coli*. As sequências para deleção dos genes ALD2, Succinate ALDH e ADH6 foram clonados diretamente na extremidade 3' do vetor. As sequências para deleção dos genes ADH2 e ADH900 foram digeridas por restrição com os pares de enzimas FastDigest (Thermo Scientific) *NotI* e *NdeI*, e *NotI* e *SacI*, respectivamente, e o vetor pGEM®-T Easy foi digerido com os mesmos pares de enzimas. As enzimas foram utilizadas com seus respectivos tampões nas condições indicadas pelo fabricante. A restrição foi realizada para a retirada dos sítios de *SalI* e *NdeI*, respectivamente, do sítio múltiplo de clonagem do vetor, uma vez que esses sítios de restrição, também localizados no interior da sequência dos genes, foram usados para linearização dos plasmídeos e inserção da sequência DR2-MazF-Hyg.



**Figura 9.** Representação esquemática do protocolo de montagem dos módulos de deleção. <sup>1</sup> Realizado para as sequências amplificadas para a deleção dos genes ALD2, Succinate ALDH e ADH6; <sup>2</sup> Realizado para as sequências amplificadas para a deleção dos genes ADH2 e ADH900; <sup>3</sup> Restrição realizada de acordo com a Tabela 8; <sup>4</sup> Realizado para os plasmídeos contendo as sequências amplificadas para a deleção dos genes ALD2, Succinate ALDH e ADH6; <sup>5</sup> Realizado para os plasmídeos contendo as sequências amplificadas para a deleção dos genes ADH6 e ADH900; <sup>6</sup> Restrição realizada de acordo com a Tabela 8; <sup>7</sup> Restrição realizada com a enzima *NotI*. Fonte: Autor.

**Tabela 7.** Tamanho de fragmentos amplificados a partir do genoma de *K. phaffii* para construção de módulos de deleção.

<b>Gene</b>	<b>Tamanho do fragmento amplificado (pb)</b>
ALD2	1896
Succinate ALDH	1821
ADH2	1293
ADH6	1469
ADH900	1398

Transformantes foram selecionados em meio LB contendo ampicilina a 100 µg/mL. Os plasmídeos contendo as sequências para deleção dos genes-alvo foram então extraídos com uso do kit GeneJET Plasmid Miniprep Kit (Thermo Scientific), de acordo com as orientações do fabricante, e a confirmação da construção foi feita por restrição e eletroforese em gel de agarose. Para inserção do cassete DR2-MazF-Hyg entre os fragmentos dos genes putativos de ALDH e ALDR, os plasmídeos construídos (denominados pGEM®-T Easy\_ALDRp or pGEM®-T Easy\_ALDHp) foram linearizados com as enzimas presentes na Tabela 8. As construções contendo os fragmentos dos genes ALD2, Succinate ALDH e ADH6, que liberam fragmentos de até 772 pb (Tabela 8) durante a digestão, foram separadas por eletroforese em gel de agarose, e a banda correspondente ao plasmídeo foi purificada com uso do kit Wizard® SV Gel and PCR Clean-Up System (Promega), de acordo com as instruções do fabricante. Os plasmídeos que contêm os genes da ADH2 e da ADH900 não precisaram ser purificados, já que não liberam fragmentos durante a digestão. Os plasmídeos contendo os genes ADH6 e ADH900 tiveram suas extremidades coesivas convertidas em extremidades abruptas com uso do kit Anza® DNA End Repair Kit (Invitrogen), de acordo com as orientações do fabricante, para ligação da sequência DR2-MazF-Hyg.

Todos os plasmídeos linearizados foram desfosforilados por ação da USB® Shrimp Alkaline Phosphatase (Affymetrix), utilizada de acordo com as instruções do fabricante, e ligados ao cassete de seleção e contrasseleção DR2-MazF-Hyg. Para tanto, os cassetes de seleção e contrasseleção foram digeridos com enzimas de restrição compatíveis com o sítio

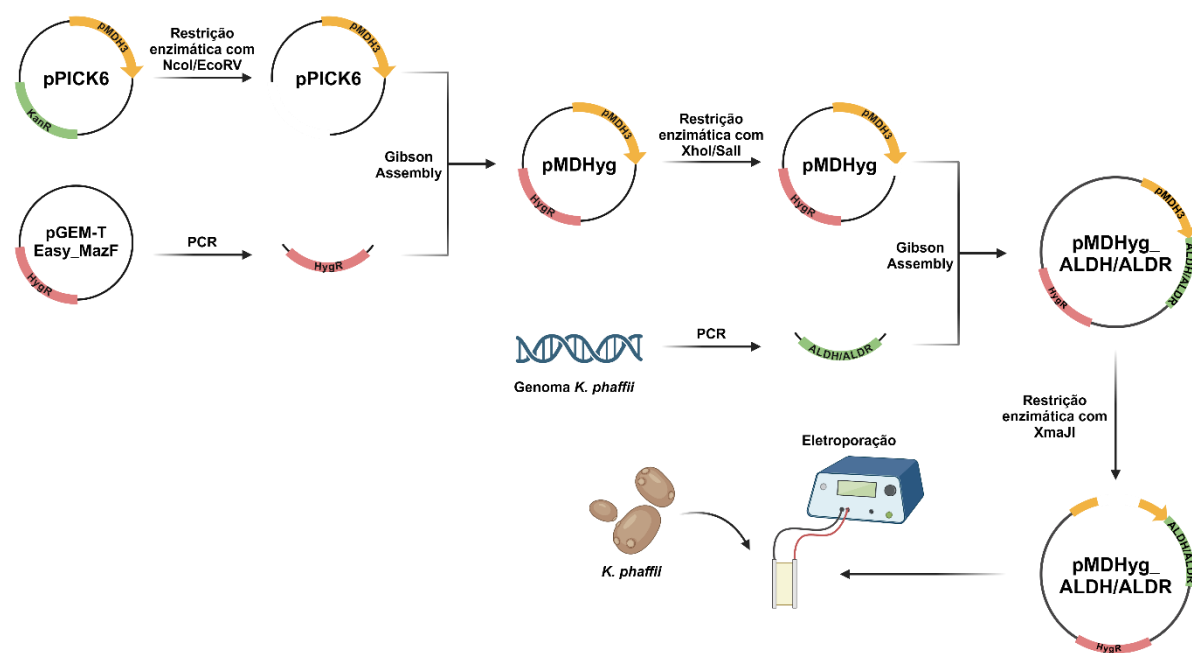
utilizado para linearização dos plasmídeos (Tabela 8). O sistema de ligação foi empregado na transformação de *E. coli*. Transformantes foram selecionados em meio LB contendo ampicilina a 100 µg/mL. Os plasmídeos contendo as sequências para deleção dos genes-alvo foram então extraídos e a confirmação da construção foi feita por restrição e eletroforese em gel de agarose. Os plasmídeos extraídos foram linearizados com uso da enzima FastDigest (Thermo Scientific) *NotI* e transformados em *K. phaffii* X33 por eletroporação.

**Tabela 8.** Sítios de restrição utilizados para linearização dos plasmídeos e digestão das extremidades do módulo de deleção.

Gene	Sítio de restrição utilizado para linearização do plasmídeo pGEM®-T Easy_ALDRp ou ALDHp	Fragmento liberado pela restrição (pb)	Sítio de restrição utilizado para digestão dos cassetes de seleção e contrasseleção
ALD2	<i>Bam</i> HI	772	<i>Bam</i> HI
Succinate ALDH	<i>Eco</i> RV	623	<i>Sma</i> I
ADH2	<i>Sal</i> I	-	<i>Sal</i> I
ADH6	<i>Xba</i> I	328	<i>Sma</i> I
ADH900	<i>Nde</i> I	-	<i>Sma</i> I

#### 4.2.8. Montagem de plasmídeos de superexpressão de genes-alvo

A estratégia para construção dos plasmídeos de superexpressão dos genes-alvo está resumida na Figura 10. Para facilitar a transformação de linhagens previamente construídas para produção de AG, buscou-se a construção de um vetor com promotor e marca de seleção que não haviam sido previamente utilizados. Portanto, para construir um vetor com o promotor constitutivo da malato desidrogenase peroxissomal ( $P_{MDH3}$ ) de *K. phaffii* (PRIELHOFER et al., 2017) para controle da expressão dos genes-alvo com marca de seleção para hidgromicina, foi utilizado o vetor pPICK6, construído anteriormente por nosso grupo de pesquisa.



**Figura 10.** Representação esquemática do protocolo de montagem dos plasmídeos de expressão dos genes-alvo selecionados. Fonte: Autor.

O vetor pPICK6, que apresenta o promotor  $P_{MDH3}$  e marca de seleção para canamicina, foi digerido com as enzimas de restrição FastDigest (Thermo Scientific) *NcoI* e *EcoRV* para excisão do gene de resistência à canamicina (*KanR*) e inserção, no mesmo cassete de expressão, do gene *HygR*, amplificado por PCR a partir do plasmídeo doado pelo professor Hugo Paes, da Faculdade de Medicina da Universidade de Brasília, que contém a sequência DR2-MazF-Hyg. O vetor digerido teve seus fragmentos separados por eletroforese em gel de agarose e o fragmento do vetor linearizado sem o gene *KanR*, com 2899 pb, foi excisado do gel e purificado. A purificação do fragmento foi conferida por eletroforese em gel de agarose. O gene de resistência à higromicina B foi amplificado com uso dos primers pHygR-pPICK6-F e

pHygR-pPICK6-R, que conferiram à sequência do gene braços de homologia com as extremidades 3' e 5' do vetor pPICK6 digerido. Foi utilizada a metodologia de *Gibson Assembly*, com uso do Gibson Assembly® Cloning Kit (New England Biolabs), de acordo com as instruções do fabricante, para clonagem do gene HygR no vetor pPICK6 linearizado. Para a reação, foram utilizados 100 ng do vetor e proporção 3:1 (inserto:vetor), e o material resultante foi transformado em *E. coli*. Transformantes foram selecionados em meio LB contendo higromicina B a 100 µg/mL. O vetor foi então extraído e a confirmação da construção foi feita por restrição e eletroforese em gel de agarose.

Após a obtenção do vetor pMDHyg, as sequências dos genes ALD2, Succinate ALDH, ADH2, ADH6 e ADH900 foram amplificadas por reação de PCR do genoma da levedura *K. phaffii* X33. Foram utilizados primers específicos que conferiram às sequências braços de homologia com o vetor pMDHyg (Tabela 4).

#### **4.2.9. Transformação de *K. phaffii* por eletroporação**

Uma colônia isolada de *K. phaffii* X33 foi inoculada em 10 mL de meio YPD, e foi incubada em shaker a 30°C e 200 rpm por 16h. Após a incubação, um volume do pré-inóculo em meio YPD foi inoculado em 100 mL de meio YPD, até atingir OD<sub>600</sub> entre 1,3 e 1,5. O inóculo foi centrifugado a 1500 x g e 4°C por 5 min, teve seu sobrenadante descartado, e o pellet foi ressuspenso em 100 mL de água estéril gelada. As células foram novamente centrifugadas sob as mesmas condições, e o pellet foi ressuspenso em 50 mL de água estéril gelada. A centrifugação foi repetida, e o pellet foi ressuspenso em 4 mL de sorbitol 1M gelado. Após uma última centrifugação, o pellet foi ressuspenso em 200 µL de sorbitol 1M gelado. Uma alíquota de 80 µL das células de *K. phaffii* foi misturada a cerca de 10 µg de material plasmidial linearizado, e foi transferida a cubeta de eletroporação (0,2 cm). Após incubação por 5 min em gelo, as cubetas receberam descarga elétrica de voltagem 1500 V, capacitância 25 µF e resistência 200 Ω. Imediatamente após a eletroporação, foi adicionado 1 mL de sorbitol 1M gelado às células, que foram então transferidas para tubo de 1,5 mL e incubadas a 30°C por 2h. As células eletroporadas foram plaqueadas em meio YPD suplementado com higromicina B a 200 µg/mL, incubadas por 96h a 30°C.

#### 4.2.10. PCR de colônia de *K. phaffii*

Uma pequena quantidade de uma colônia de interesse foi transferida para microtubo de 0,2 mL contendo 5 µL de água ultrapura. Seguiu-se com o protocolo convencional de PCR, com tempo de 5 min na primeira etapa para lise das leveduras. O material genético presente no lisado celular é usado como molde da PCR.

#### 4.2.11. Otimização da produção de ácido glicólico pela *K. phaffii* JA122 em biorreator

A linhagem de *K. phaffii* JA122, que expressa a via completa de conversão de xilose a ácido glicólico, foi inoculada em tubo do tipo Falcon de 50 mL contendo 10 mL de meio YPD suplementado com geneticina a 200 µg/mL, zeocina a 100 µg/mL e higromicina B a 200 µg/mL, e incubada em shaker a 30°C e 200 rpm por 24h. Após a incubação, o tubo teve seu conteúdo distribuído em 4 erlenmeyers de 1 L contendo 100 mL de meio YPD cada, que foram incubados a 30°C e 200 rpm por 48h. O inóculo em YPD foi então centrifugado a 6000 x g e 4°C por 5 min, e o pellet foi ressuscitado em um pequeno volume de água. As células foram inoculadas em biorreatores de 1 e 0,5 L, contendo 400 mL de meio FM22. Foram mantidos os parâmetros de temperatura a 30°C e de agitação a 200 rpm ao longo da fermentação. As bateladas duraram 96 h e foram planejadas para avaliação da interação entre o pH e o DO (oxigênio dissolvido), de acordo com um Delineamento Composto Central Rotacional (DCCR), contendo 4 pontos axiais e 3 repetições do ponto central. A codificação das variáveis em diferentes níveis se encontra na Tabela 9, e os ensaios foram realizados com a combinação das duas variáveis, de acordo com a matriz da Tabela 10. Foram coletadas amostras a cada 24 h a partir do momento do inóculo, cujos componentes foram quantificados por UPLC.

**Tabela 9.** Codificação das variáveis pH e DO em níveis para construção do Delineamento Central Composto Rotacional.

Nível	pH	DO (%)
-1,41	5,00	2,0

-1	5,44	4,6
0	6,50	11,0
1	7,56	17,4
1,41	8,00	20,0

**Tabela 10.** Matriz do planejamento de ensaios para Delineamento Composto Central Rotacional.

Ensaio	pH	DO
1	-1	-1
2	1	-1
3	-1	1
4	1	1
5	-1,41	0
6	1,41	0
7	0	-1,41
8	0	1,41
9	0	0
10	0	0
11	0	0

#### 4.2.12. Produção de ácido glicólico por *K. phaffii* JA122 a partir de hidrolisados de biomassa

A linhagem de *K. phaffii* JA122, que expressa a via completa de conversão de xilose a ácido glicólico, foi inoculada em tubo do tipo Falcon de 50 mL contendo 10 mL de meio YPD suplementado com geneticina a 200 µg/mL, zeocina a 100 µg/mL e higromicina B a 200 µg/mL, e incubada em shaker a 30°C e 200 rpm por 24h. Após a incubação, o tubo teve seu conteúdo distribuído em 4 Erlenmeyers de 1 L contendo 100 mL de meio YPDX cada, que foram incubados a 30°C e 200 rpm por 48h. O inóculo em YPDX foi então centrifugado a 6000



x g e 4°C por 5 min, e o pellet foi ressuspendido em um pequeno volume de água. As células foram inoculadas em quatro biorreatores de 3,6 L, contendo volume inicial de 900 mL de meio YP suplementado com 70 g/L de glicose e 10 g/L de xilose. Após 24 h, com o consumo total da glicose presente no meio, foi iniciada uma etapa de alimentação, em que dois vasos foram alimentados com 1 L de YP suplementado com 40 g/L de glicose e 86 g/L de xilose e dois vasos foram alimentados com 900 mL de hidrolisado de biomassa de cana-de-açúcar suplementado com glicose e xilose para concentrações finais similares ao do meio sintético. O hidrolisado já apresentava glicose e xilose em concentração inicial de 5,47 g/L e 81,52 g/L, respectivamente. A fase de alimentação durou 48 h, com fluxo de 20 mL/min, e foram mantidos os parâmetros de temperatura a 30°C e de pH a 6, por meio da adição automática de hidróxido de amônio, ao longo da fermentação. O oxigênio dissolvido (DO) foi mantido a 20%, controlado pelos parâmetros de agitação (200 a 600 rpm) e de fluxo de ar (1 a 1,5 L/min) em cascata. Após a fase de alimentação, a fermentação foi continuada por mais 48 h, totalizando 120 h de cultivo. Foram coletadas amostras ao longo do experimento, cujos componentes foram quantificados por cromatografia líquida (HPLC e UPLC).

Para a obtenção do hidrolisado, 8 kg de biomassa de cana-de-açúcar foram processados por explosão a vapor por 8 min a 15 kgf/cm<sup>2</sup> (197,85°C) com 50% de umidade. Em seguida, a porção líquida obtida a partir da explosão a vapor foi submetida a hidrólise ácida para quebra dos oligossacarídeos. Para isso, foi adicionado ácido sulfúrico para concentração final de 0,5% (v/v) e a reação foi incubada a 130°C por 100 min.

#### **4.2.13. Análise de metabólitos por HPLC/UPLC**

As amostras coletadas de ensaios fermentativos foram analisadas por cromatografia líquida de ultra alta eficiência (UPLC) em cromatógrafo Acquity UPLC H-Class Plus (Waters), equipado com uso de diferentes métodos e colunas de acordo com o composto a ser analisado (Tabela 11). O método utilizado para cada composto e seus respectivos tempos de retenção nas colunas estão explicitados na Tabela 12. As fases móveis foram preparadas em capela de exaustão e degaseificada por 10 minutos. Foram utilizados os detectores de Índice de Refração (RID) e de Arranjo de Fotodiodos (PDA), para detecção de xilose, glicose, xilitol, ácido xilônico, glicerol, etilenoglicol e ácido glicólico.

**Tabela 11.** Metodologias utilizadas para quantificação de compostos em solução por HPLC e UPLC.

Coluna	Detector	Temperatura da coluna	Fase móvel	Fluxo da fase móvel	Tempo de análise
Aminex HPX 87H (Bio-Rad)	RID	45°C	H <sub>2</sub> SO <sub>4</sub> 5mM	0,6 mL/min	24 min
C18 HSS T3 2,1 x 150 mm, 1,8 µm (Waters)	PDA	30°C	Solvente A: H <sub>3</sub> PO <sub>4</sub> 0,125% Solvente B: CH <sub>3</sub> OH 100%	0,3 mL/min	15 min
BEH Amide 2,1 x 150 mm, 1,7 µm (Waters)	ELSD	30°C	Solvente A: CH <sub>3</sub> CN 80% + CH <sub>3</sub> COONH <sub>4</sub> 10mM + NH <sub>4</sub> OH 0,2% Solvente B: CH <sub>3</sub> CN 40% + CH <sub>3</sub> COONH <sub>4</sub> 10mM + NH <sub>4</sub> OH 0,2%	0,3 mL/min	11 min

**Tabela 12.** Coluna e detector usados na quantificação de cada composto de interesse e seus respectivos tempos de retenção no método utilizado.

Composto	Coluna	Detector	Tempo de retenção
Glicose	HPX 87H	RID	8.952
Xilose	BEH Amide	ELSD	11.172
Xilitol	HPX 87H	RID	9.600
Glicerol	HPX 87H	RID	13.360
Acetato	HPX 87H	RID	15.263
Etanol	HPX 87H	RID	21.646
Etilenoglicol	HPX 87H	RID	16.153
Ácido glicólico	C18 HSS T3	PDA	12.405
Ácido	BEH Amide	ELSD	3.605

As amostras coletadas durante os cultivos foram centrifugadas (14.000 g x 10 min), e o sobrenadante foi armazenado a -20°C até o momento da análise. A concentração de cada metabólito foi calculada de acordo com o tamanho do pico detectado em cada amostra em comparação com curva-padrão. As curvas-padrão foram montadas a partir dos picos identificados de padrões analíticos de concentração conhecida, preparados por meio de diluição seriada dos metabólitos de interesse (Tabela 13).

**Tabela 13.** Concentração dos padrões analíticos utilizados para quantificação de compostos em HPLC e UPLC.

Composto	Concentração dos padrões analíticos em diluição seriada (g/L)						
	P1	P2	P3	P4	P5	P6	P7
Glicose	20	10	5	2,5	1,25	0,625	0,3125
Xilose	20	10	5	2,5	1,25	0,625	0,3125
Xilitol	10	5	2,5	1,25	0,625	0,3125	0,1563
Glicerol	5	2,5	1,25	0,625	0,3125	0,1563	0,0782
Acetato	5	2,5	1,25	0,625	0,3125	0,1563	0,0782
Etanol	5	2,5	1,25	0,625	0,3125	0,1563	0,0782
Etilenoglicol	10	5	2,5	1,25	0,625	0,3125	0,15625
Ácido glicólico	10	5	2,5	1,25	0,625	0,3125	0,15625
Ácido xilônico	5	2,5	1,25	0,625	0,3125	0,1563	0,0782

## 5. Resultados e discussão

### 5.1. Seleção de genes-alvo para deleção e superexpressão

A construção de linhagem de *K. phaffii* JA122 expressando a via de Dahms para produção de EG demonstrou que essa levedura possui atividades nativas de glicolaldeído desidrogenase (ALDH) e glicolaldeído redutase (ALDR). As enzimas que apresentam essas atividades convertem, respectivamente, o glicolaldeído em ácido glicólico (AG) ou etilenoglicol (EG). Portanto, trabalhou-se com a hipótese de que a construção de uma linhagem de *K. phaffii* capaz de produzir AG com maior eficiência poderia ser obtida pela deleção dos genes de ALDR envolvidos na via competitiva que leva à produção de EG, e por meio da superexpressão de genes de ALDH capazes de converter glicolaldeído a AG. Como as enzimas responsáveis por tais atividades não são conhecidas em *K. phaffii*, a primeira etapa desse trabalho foi identificar por bioinformática possíveis genes envolvidos em tais funções.

Nesse sentido, foram selecionadas duas ALDH's (ALD2, Succinate ALDH) e três ALDR's (ADH2, ADH6, ADH900) putativas para análises. As ALDHs foram escolhidas com base na similaridade com as sequências da AldA de *E. coli* e da ALD4 de *S. cerevisiae*, sequências de aldeído desidrogenases de referência. As enzimas aldeído desidrogenases putativas avaliadas neste trabalho não foram descritas previamente. As ALDR foram escolhidas com base na similaridade à sequência da ADH1 de *S. cerevisiae*, que apresenta atividade de aldeído redutase. Os genes da ADH2 e da ADH900 são os genes codificantes de álcool desidrogenases mais expressos em *K. phaffii*, e apresentam atividade substancial na produção e no consumo de etanol na levedura (ZAVEC et al., 2021). O gene ADH2 é o único responsável pelo consumo de etanol em *K. phaffii*, enquanto o gene ADH900 é o principal produtor de etanol na espécie (KARAOĞLAN et al., 2020). Além disso, foi demonstrado que as enzimas codificadas por esses genes possuem papel na assimilação de metanol em linhagens Mut- ( $\Delta$ ox1,  $\Delta$ ox2) (ZAVEC et al., 2021). O gene ADH6 também está envolvido com a produção de etanol na levedura, porém possui papel mínimo quando comparado à ADH900 (KARAOĞLAN et al., 2020). Dessa forma, presumiu-se que esses genes também podem estar

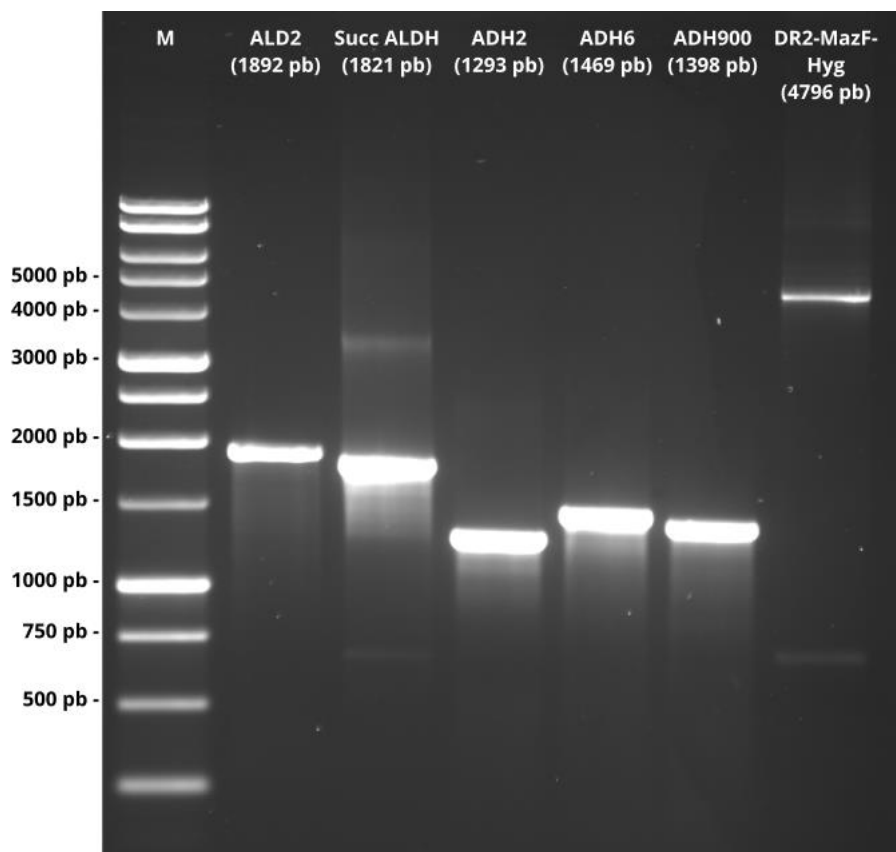
envolvidos no metabolismo de gliceraldeído e EG, uma vez que estão envolvidos com a oxirredução de grupos álcool e aldeído.

## **5.2. Construção dos módulos de deleção de genes-alvo**

A estratégia para construção dos módulos de deleção dos genes-alvo selecionados é resumida na Figura 9. Uma vez construídos os módulos de deleção, esses foram empregados para transformar a levedura *K. phaffii* X33. Os detalhes da construção dos cassetes de deleção e da obtenção das linhagens com genes deletados são apresentados a seguir.

### **5.2.1. Amplificação de sequências para deleção de genes-alvo**

As sequências para deleção dos cinco genes-alvo em *K. phaffii* foram corretamente amplificadas a partir do genoma da levedura. Para todos os genes, foi amplificada a sequência codante mais parte das regiões 5' e 3' não transcritas a partir do DNA genômico da levedura, para garantir um braço de homologia de cerca de 600 pb para a recombinação em *K. phaffii*. A sequência dos cassetes de seleção e contraseleção DR2-MazF-Hyg também foi corretamente amplificada a partir do plasmídeo de origem. Fragmentos correspondentes aos genes ALD2 (1892 pb), Succinate ALDH (1821 pb), ADH2 (1293 pb), ADH6 (1469 pb), ADH900 (1398 pb) e DR2-MazF-Hyg (4796 pb) foram obtidos com sucesso. A confirmação do sucesso da PCR se deu por eletroforese em gel de agarose, em que foi observada banda de DNA do tamanho esperado (Figura 11). A presença de bandas inespecíficas na eletroforese da Succinate ALDH e da DR2-MazF-Hyg fez necessária a excisão e purificação das bandas corretas, porém as bandas incorretas continuaram a aparecer na eletroforese após a purificação.

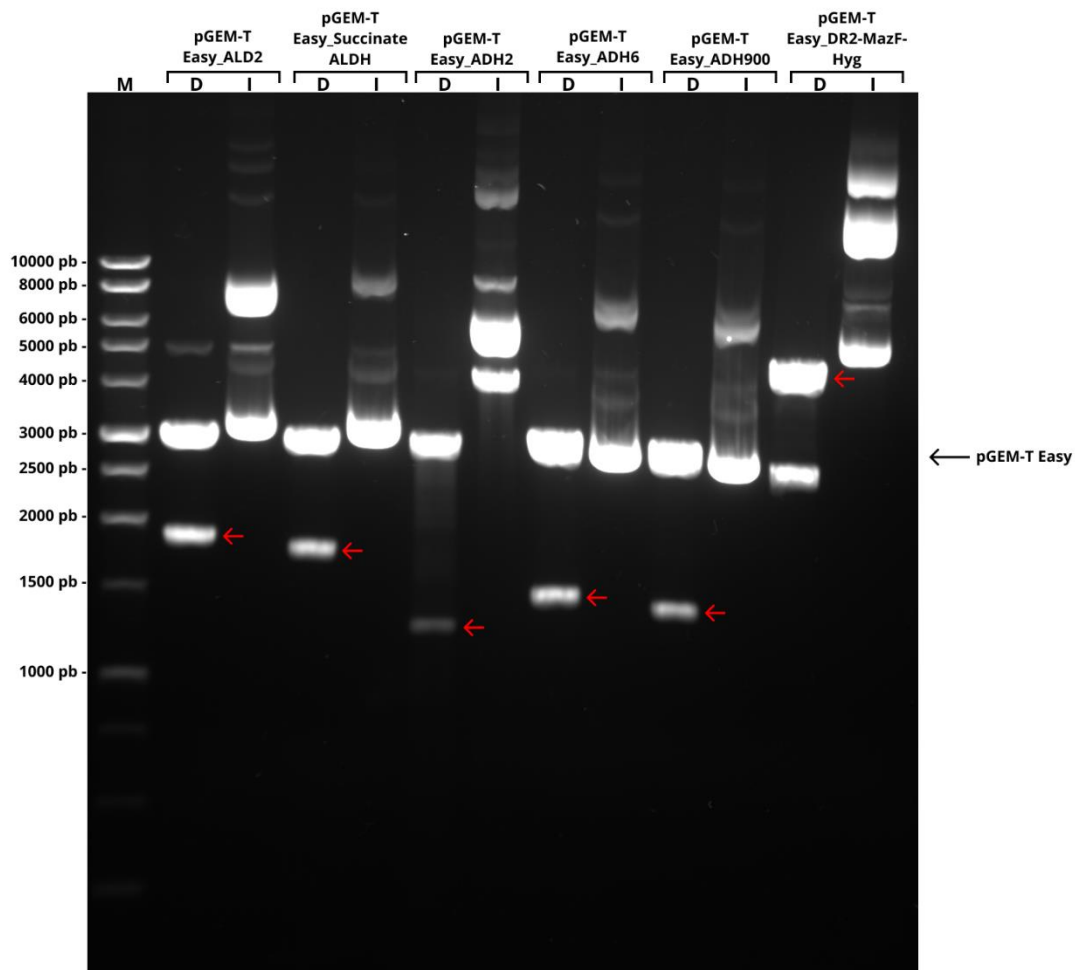


**Figura 11.** Análise eletroforética em gel de agarose de PCR das sequências amplificadas para construção dos módulos de deleção de genes-alvo. M = Marcador de peso molecular Promega 1Kb DNA Ladder. A identificação de cada amostra e tamanho esperado do fragmento estão indicados na figura.

### 5.2.2. Construção dos módulos de deleção de genes-alvo para *K. phaffii*

A primeira etapa para construção dos módulos de deleção de genes-alvo envolveu a clonagem dos produtos de PCR correspondentes aos 5 genes a serem deletados (ALD2, Succinate ALDH, ADH2, ADH6 e ADH900) no vetor pGEM®-T Easy. Além disso, para obter o cassete DR2-MazF-Hyg em grande quantidade e com sítios específicos, a sequência também foi clonada no mesmo vetor. Os produtos de PCR previamente purificados foram ligados ao vetor de clonagem, e os sistemas de ligação foram dialisados e eletroporados em *E. coli* para propagação. Colônias transformadas foram selecionadas por resistência à ampicilina e por seleção azul-branco. As colônias tiveram seu DNA plasmidial extraído, e a confirmação das construções foi feita por digestão dos plasmídeos e eletroforese em gel de agarose dos fragmentos digeridos (Figura 12). A presença de bandas a mais na eletroforese da PCR das sequências Succinate ALDH e DR2-MazF-Hyg (Figura 11) não comprometeu a clonagem dos fragmentos no vetor. Para todas as construções, a digestão foi realizada de forma a liberar um

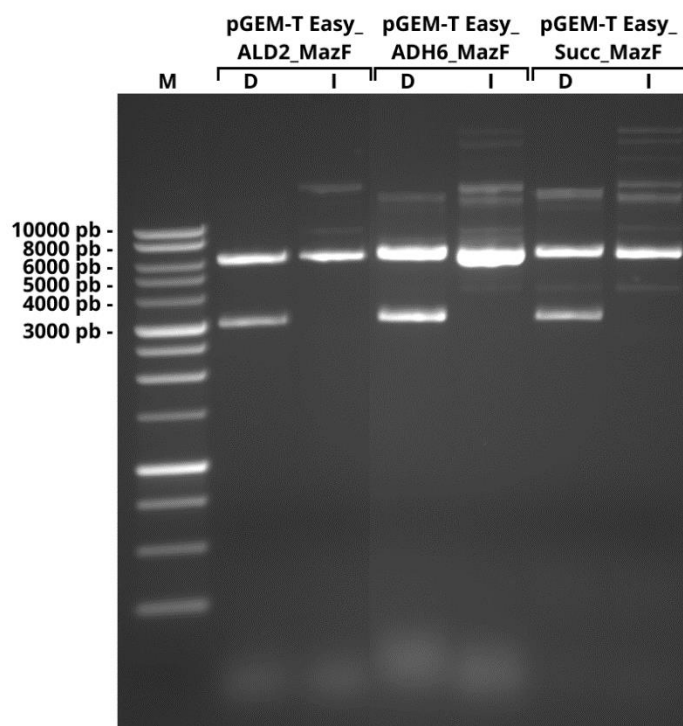
fragmento de cerca de 3 Kb, correspondente ao vetor pGEM®-T Easy, e um fragmento de tamanho próximo ao amplificado anteriormente por PCR. Com a liberação dos fragmentos esperados, também foi realizado o sequenciamento dos insertos. Os resultados demonstraram que todos os fragmentos foram clonados com sucesso e continham as sequências esperadas.



**Figura 12.** Análise eletroforética em gel de agarose dos plasmídeos contendo as sequências para deleção dos genes-alvo e os cassetes de seleção e contrasseleção (DR2-MazF-Hyg). M = Marcador de peso molecular Promega 1Kb DNA Ladder. Setas vermelhas = fragmentos liberados do vetor após a digestão. D = plasmídeo digerido. I = plasmídeo íntegro.

Para a inserção da sequência DR2-MazF-Hyg entre as regiões 5' e 3' dos 5 genes a serem deletados (Figura 8), os plasmídeos pGEM®-T Easy\_ALDRp e pGEM®-T Easy\_ALDHp foram linearizados e passaram por purificação. Os plasmídeos contendo as sequências dos genes ADH6 e ADH900 passaram por processo de geração de extremidades abruptas (*blunting*). Os plasmídeos linearizados foram desfosforilados e ligados aos cassetes de seleção e contrasseleção DR2-MazF-Hyg. Os sistemas de ligação foram eletroporados em

*E. coli* para propagação e as colônias resultantes tiveram seu DNA plasmidial extraído por miniprep. A inserção da DR2-MazF-Hyg no interior da sequência para deleção do gene foi feita com sucesso para os genes ALD2, ADH6 e Succinate ALDH, e foi confirmada por digestão pela enzima de restrição *NotI* e eletroforese em gel de agarose, em que eram esperados fragmentos de aproximadamente 6 Kb e 3 Kb (Figura 13). As outras bandas presentes nas amostras de digestão correspondem a frações não digeridas ou parcialmente digeridas.



**Figura 13.** Análise eletroforética em gel de agarose dos plasmídeos contendo os módulos de deleção dos genes ALD2, ADH6 e Succinate ALDH. M = Marcador de peso molecular Promega 1Kb DNA Ladder. D = plasmídeo digerido; I = plasmídeo íntegro.

Para os genes ADH2 e ADH900, foram realizadas várias ligações e diferentes transformações, porém nenhuma colônia apresentou plasmídeo com o perfil esperado. Foi observada uma alta taxa de recombinação entre as regiões DR2 nos plasmídeos construídos. Com vistas à diminuição de eventos de recombinação, foi empregada transformação por choque térmico e foi utilizada menor temperatura de recuperação e crescimento nas linhagens de *E. coli* transformadas (30°C). Além disso, foi acrescido 1% de glicose ao meio de cultura LB no momento da ressuspensão das células após o choque térmico e nas placas, para melhor recuperação das células transformadas. Apesar das modificações nos protocolos, não foi possível obter as construções para os dois genes citados. Até o momento, os módulos de



deleção dos genes ALD2, ADH6 e Succinate ALDH foram construídos com sucesso e empregados na construção de linhagens recombinantes de *K. phaffii*.

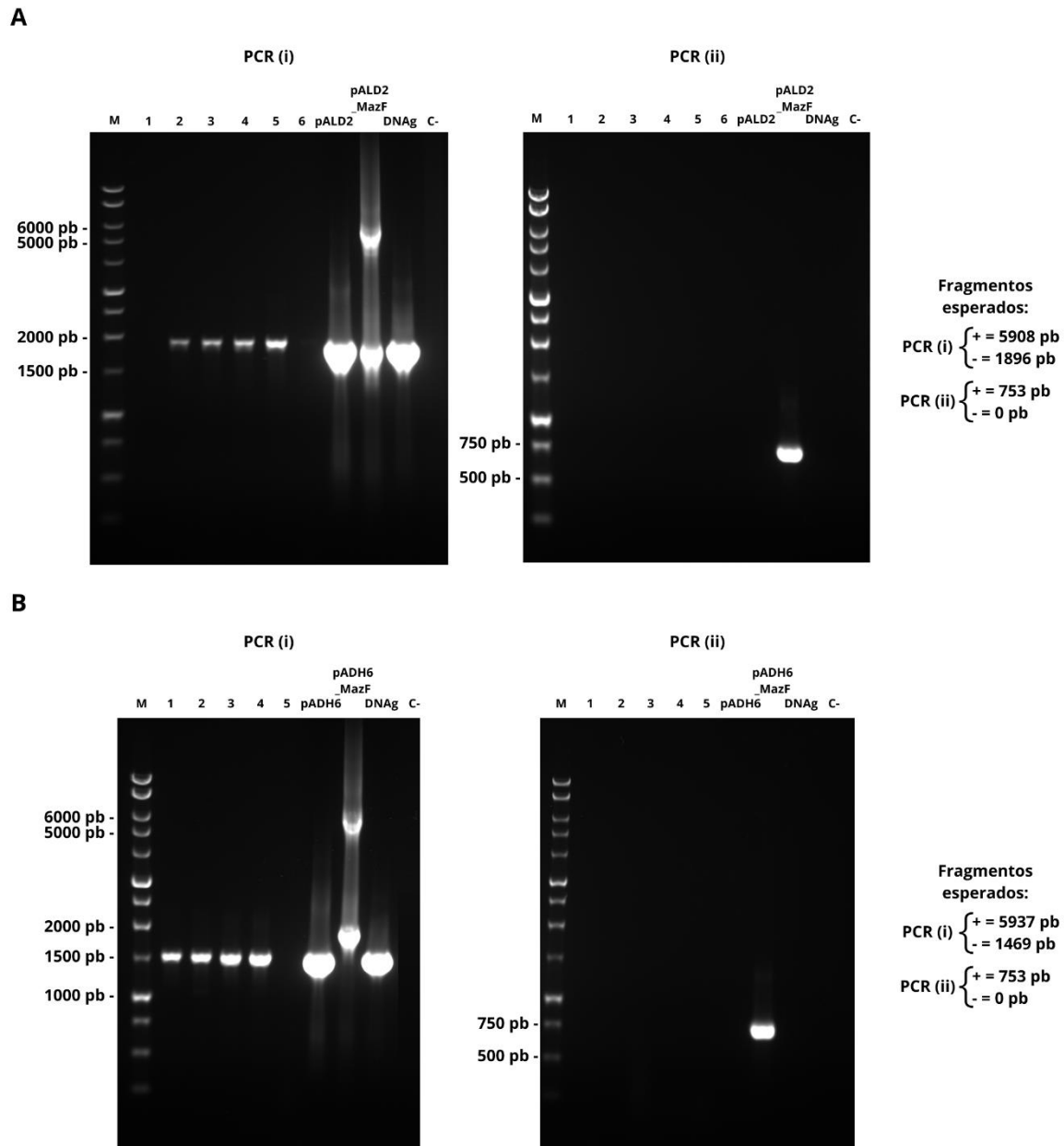
### **5.3. Construção de linhagens recombinantes de *K. phaffii* com genes-alvo deletados**

Para a construção de linhagens de *K. phaffii* contendo os genes ALD2, ADH6 e Succinate ALDH deletados, a linhagem X33 foi transformada com os módulos de deleção construídos. As colônias obtidas a partir da eletroporação foram repicadas para placas de YPD acrescido de higromicina B a 200 µg/mL e utilizadas em PCR de colônia para confirmação da integração do módulo de deleção ao genoma.

Para confirmação da deleção dos genes de interesse, foi realizada PCR de colônia com oligonucleotídeos específicos. Para isso, foram idealizadas duas PCR's para cada gene, com os primers (i) utilizados para amplificação das sequências dos genes-alvo, e (ii) InpMDH-F e InHyg-R, que se anelam no interior da sequência DR2-MazF-Hyg. Dessa forma, caso o módulo tivesse sido inserido corretamente, na PCR (i) seria amplificado todo o módulo de deleção, com cerca de 6 kb, e na PCR (ii) seria amplificado um fragmento de 753 pb. Caso contrário, na PCR (i) seria amplificado o fragmento original para deleção dos genes, e na PCR (ii) não seria amplificado nenhum fragmento. Para todas as reações de PCR, foi feito um controle negativo sem DNA molde para verificar possíveis contaminações da reação, um controle com o plasmídeo contendo a sequência para deleção dos genes, um controle com o plasmídeo contendo o módulo de deleção, e um controle com o DNA genômico da levedura.

O tamanho das sequências amplificadas foi verificado por eletroforese em gel de agarose (Figura 14). Para todos os genes, as colônias obtidas não apresentaram integração correta do módulo de deleção ao genoma. Na PCR (i) foram observadas bandas com o tamanho original dos genes-alvo, enquanto na PCR (ii) não foi amplificado fragmento para nenhuma das reações (Figura 14). No controle do plasmídeo contendo o módulo de deleção, foi observada banda de tamanho inespecífico para todos os genes, entre 1,5 e 2 kb para a ALD2, e próxima a 1,5 kb para a ADH6. Foram realizadas diversas modificações no protocolo de transformação para aumentar sua eficiência, como aumento da voltagem na eletroporação, adição de glicose ao sorbitol na etapa de recuperação após o choque, e adição de sorbitol no meio YPD em que foram plaqueadas as células. Apesar dos esforços, não foram obtidas

colônias positivas para nenhum dos genes. Em grande parte das transformações não houve crescimento de colônias, e as colônias que apresentaram crescimento nas placas da transformação foram todas testadas. Para o módulo de deleção da Succinate ALDH, não houve crescimento de colônias em nenhuma das transformações.



**Figura 14.** Análise eletroforética em gel de agarose de PCR de colônia para confirmação de integração do módulo de deleção no genoma da *K. phaffii*. M = Marcador de peso molecular. C- = Controle negativo da reação de PCR. A = PCR para o gene ALD2. B = PCR para o gene ADH6. PCR (i) = PCR realizada com os primers utilizados para amplificação das sequências dos genes-alvo. PCR (ii) = PCR realizada com os primers InpMDH-F e InHyg-R. + = Integração correta do módulo de deleção ao genoma. - = Módulo de deleção não integrado ao genoma.

Acredita-se que o crescimento de colônias não transformadas e sem resistência à higromicina B nas placas contendo o antibiótico possa ter se dado por conta da quantidade de material celular plaqueado, de tal forma que as colônias cresciam sobre as próprias leveduras mortas, inibindo a ação do antibiótico. Para evitar que isso ocorresse, foram utilizadas placas de petri 150x15 mm, que possuem maior superfície para espalhamento das células. Com o uso dessas placas, não foram mais observadas colônias após as transformações. Há possibilidade de que esses genes, pouco ou não caracterizados, estejam envolvidos em vias metabólicas essenciais para a sobrevivência da levedura, o que justificaria a falta de colônias nas placas da transformação após sua deleção.

Devido às dificuldades para integração dos cassetes, foi pensada uma metodologia alternativa para a avaliação do papel funcional dos genes-alvo selecionados. Para isso, foram desenhados novos primers para construção de cassetes de superexpressão para os genes de interesse, com atividade putativa de ALDH e ALDR. Assim, o papel dos genes na produção de etilenoglicol (EG) e ácido glicólico (AG) pode ser investigado por meio da superexpressão dos genes em *K. phaffii*, seguida de ensaios fermentativos na presença de glicaldeído, EG e AG com as linhagens recombinantes obtidas.

#### **5.4. Construção de plasmídeos para superexpressão de genes-alvo**

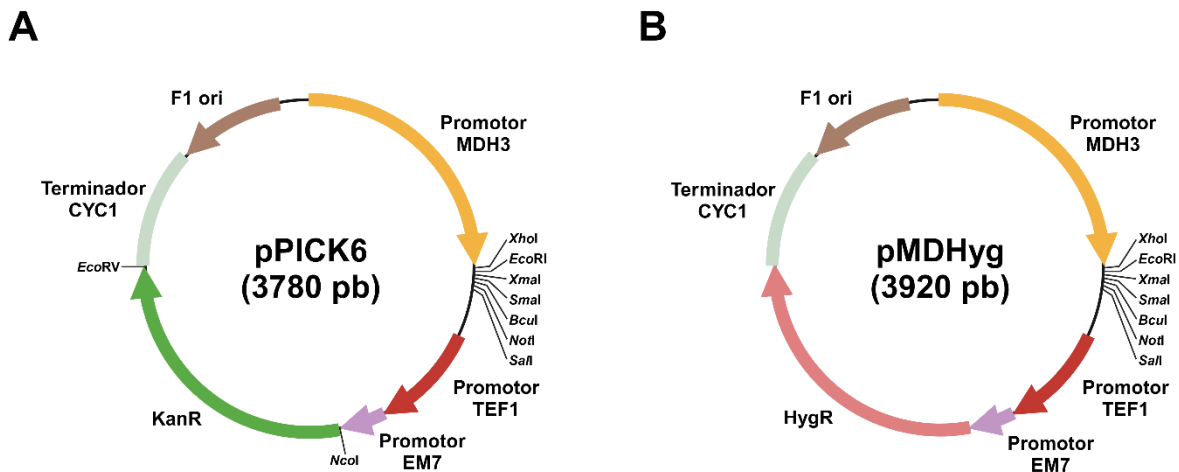
A estratégia para construção dos plasmídeos para superexpressão dos genes-alvo selecionados é apresentado a seguir.

##### **5.4.1. Construção do vetor pMDHyg**

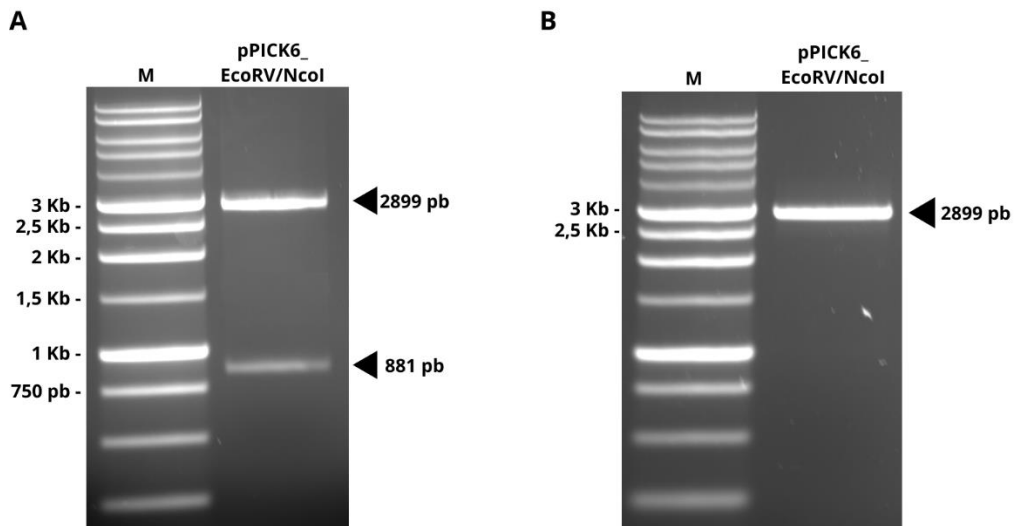
Para a superexpressão dos genes-alvo selecionados, foi idealizado um vetor que contivesse o promotor  $P_{MDH3}$  e marca de resistência à higromicina B, para ser integrado ao genoma de uma linhagem de *K. phaffii* que expressa a via de Dahms. O promotor foi escolhido por ser um promotor constitutivo forte (PRIELHOFER et al., 2017) e porque seu locus ainda não havia sido utilizado para integração dos cassetes de expressão dos genes da via de Dahms. O vetor idealizado foi chamado de pMDHyg. Da mesma forma, a marca de resistência à higromicina B foi escolhida por ainda não ter sido empregada na levedura recombinante.

Para a construção do vetor pMDHyg, primeiramente o vetor pPICK6 foi digerido com as enzimas de restrição *EcoRV* e *NcoI* para excisão do gene KanR, que confere resistência à canamicina. Após a digestão, os fragmentos resultantes foram analisados por eletroforese em gel de agarose (Figura 16a). Os fragmentos esperados, de 881 pb, correspondente ao gene

KanR, e de 2899 pb, correspondente ao restante do vetor, foram obtidos com sucesso. Posteriormente, o fragmento correspondente ao vetor foi excisado do gel e purificado. O sucesso da purificação foi conferido por eletroforese em gel de agarose, em que se observou banda única de 2899 pb, correspondente ao vetor linearizado (Figura 16b).



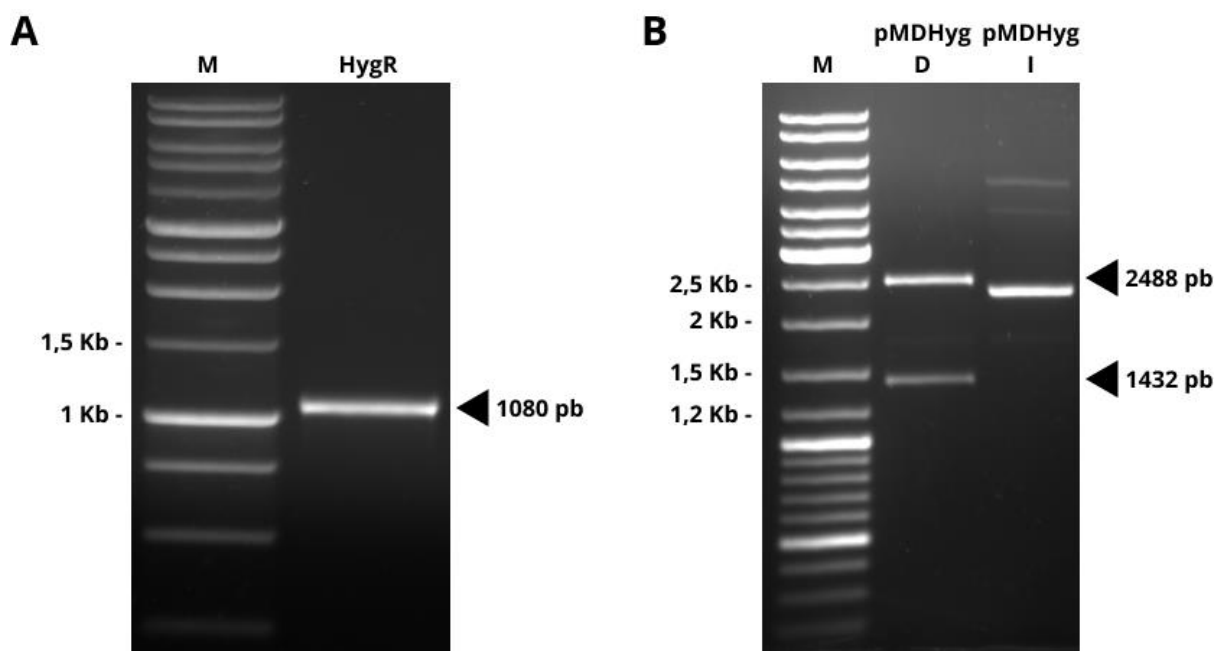
**Figura 15.** Representação esquemática dos vetores (A) pPICK6 e (B) pMDHyg. Fonte: Autor.



**Figura 16.** Análise eletroforética em gel de agarose (A) do vetor pPICK6 digerido para liberação do gene KanR e (B) do vetor pPICK6 com o gene KanR removido purificado. M = Marcador de peso molecular.

Após a obtenção do fragmento do pPICK6 sem o gene KanR, o gene HygR foi amplificado por reação de PCR a partir do vetor doado pelo professor Hugo Paes, da Faculdade de Medicina da Universidade de Brasília, que contém a sequência DR2-MazF-Hyg. A PCR foi realizada com uso dos primers HygR-pPICK6-F e HygR-pPICK6-R, que conferiram à

sequência amplificada braços de homologia com as extremidades 3' e 5' do vetor pPICK6 digerido. O tamanho da sequência amplificada foi conferido por eletroforese em gel de agarose (Figura 17a). Após a confirmação da obtenção do fragmento correto, o produto da PCR foi purificado e clonado no vetor pPICK6 linearizado por *Gibson Assembly*. O material resultante foi transformado em *E. coli*. Colônias transformadas foram selecionadas por resistência à higromicina B e tiveram seu material genético plasmidial extraído por miniprep. A construção do vetor foi confirmada por digestão com a enzima de restrição *SacI* e eletroforese em gel de agarose (Figura 17b). Eram esperados, como resultado da digestão, um fragmento de 2488 pb e um fragmento de 1432 pb. A banda de tamanho entre 1,5 e 2 kb é resultado de digestão parcial do vetor.

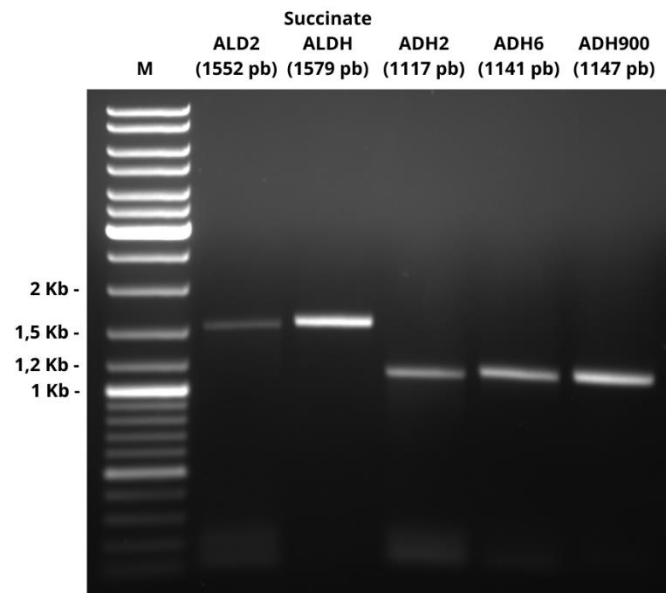


**Figura 17.** Análise eletroforética em gel de agarose (A) da PCR para amplificação do gene *HygR* e (B) do vetor pMDHyg digerido por ensaio de restrição com a enzima *SacI*. D = Digerido; I = Íntegro.

#### 5.4.2. Construção de plasmídeos de superexpressão

Após a obtenção do vetor pMDHyg, os genes-alvo selecionados foram amplificados por PCR a partir do genoma da *K. phaffii* X33, e o tamanho dos fragmentos amplificados foi verificado por eletroforese em gel de agarose (Figura 18). Os genes foram amplificados com uso de primers específicos que conferiram às sequências braços de homologia para sua inserção no vetor pMDHyg por *Gibson Assembly*. Assim, com a construção do vetor pMDHyg e a amplificação das sequências dos genes-alvo, foi viabilizada a construção dos plasmídeos de superexpressão, processo ainda em execução. Os plasmídeos construídos poderão ser

posteriormente utilizados na transformação de linhagem de *K. phaffii* que expresse a via de Dahms, e as linhagens recombinantes geradas poderão ser estudadas a fim de compreender se a superexpressão dos genes putativos de ALDR selecionados contribui ou não interfere na produção de ácido glicólico a partir de xilose e se a superexpressão dos genes putativos de ALDH prejudica ou não interfere nessa produção.



**Figura 18.** Análise eletroforética em gel de agarose da PCR para amplificação das sequências dos genes-alvo para construção dos plasmídeos de expressão. A identificação de cada amostra e o tamanho esperado para cada gene está identificado na imagem.

### 5.5. Otimização do processo de produção de ácido glicólico por *K. phaffii* JA122

Com o objetivo de analisar a interação entre os parâmetros críticos pH e DO (oxigênio dissolvido) na produção de ácido glicólico (AG), foi idealizado um delineamento central composto rotacional (DCCR). O DCCR é uma ferramenta estatística utilizada para estabelecer relações entre múltiplas variáveis independentes e uma variável resposta de um sistema. Assim, é possível determinar as condições ótimas de cultivo para pH e DO de maneira conjunta, buscando a maximização de coeficientes técnicos como produtividade e rendimento no processo de produção do AG. O DCCR foi idealizado com quatro pontos axiais ( $\alpha = 1,41$ ), quatro pontos fatoriais e três repetições do ponto central (Tabela 10), totalizando 11 ensaios. O pH foi escolhido como parâmetro a ser otimizado uma vez que o ácido glicólico e um intermediário da sua via de produção, o ácido xilônico, são ácidos orgânicos, que acidificam o meio de cultura ao longo de processos fermentativos. Além disso, o ácido xilônico se acumula a nível intracelular em pH's mais altos, o que aumenta seu efeito prejudicial sobre a viabilidade

e sobrevivência celulares (TOIVARI et al., 2013). O parâmetro de DO foi escolhido pelo fato de a *K. phaffii* ser uma levedura de metabolismo respiratório, portanto uma maior oxigenação favorece seu crescimento e atividade celular, e porque a via de produção de ácido glicólico faz uso do cofator NAD<sup>+</sup> na primeira e na última reação da via, tal que esse cofator pode ser restaurado pela cadeia transportadora de elétrons.

Para este experimento, a linhagem mutante *K. phaffii* JA122, que expressa a via de produção de etilenoglicol (EG), foi cultivada em meio FM22 em biorreator por 96h nas condições determinadas pelo DCCR. Foram utilizadas como fonte de carbono xilose a 40 g/L e glicose a 20 g/L. A proporção entre os açúcares foi previamente analisada pelo grupo de pesquisa, que determinou as concentrações utilizadas como ótimas para a produção de AG. As amostras, coletadas de 24 em 24 horas a partir do inóculo, foram analisadas em UPLC/HPLC para quantificação dos componentes individuais no meio de cultura ao longo da fermentação. A concentração de AG e EG produzidos em cada condição de cultivo está exposta na Tabela 14.

**Tabela 14.** Concentrações de ácido glicólico e etilenoglicol obtidas em cada condição de cultivo estipulada por Delineamento Central Composto Rotacional.

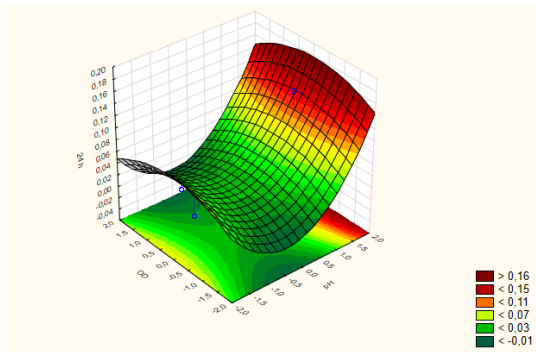
Condição	pH	DO (%)	Ácido Glicólico (g/L)				Etilenoglicol (g/L)			
			24 h	48 h	72 h	96 h	24 h	48 h	72 h	96 h
1	5,44	4,6	0	0,202	0,31	1,595	0,327	2,268	1,102	1,448
2	7,56	4,6	0	0	0,13	0,419	0,703	2,838	3,753	4,399
3	5,44	17,4	0	0,186	0,311	1,48	0,588	3	2,353	0,908
4	7,56	17,4	0	0,205	0,45	0,585	0	0,204	0,37	0,516
5	5	11	0	0,272	0,296	0,313	0,645	2,665	4,386	5,5
6	8	11	0,137	0,207	0,335	0,44	0	0	0,452	0,266
7	6,5	2	0	0	0,21	0,44	0,241	2,245	4,345	5,958
8	6,5	20	0	0,11	0,145	0,165	0,23	1,743	4,14	5,15
9	6,5	11	0	0	0,235	0,465	0,145	1,734	3,143	4,447
10	6,5	11	0	0,22	0,4	0,47	0,745	3,032	4,89	4,945
11	6,5	11	0	0	0,115	0,565	0	0	0	0,536

Os dados de concentração de AG obtidos em cada condição foram utilizados para o desenvolvimento de gráficos de superfície de resposta por meio do software *Statistic*® (Figura 19). O gráfico de superfície de resposta é uma reprodução tridimensional da relação entre as variáveis independentes (pH e DO) e a variável resposta (produção de AG). Nos gráficos de superfície de resposta, a variação de cor do verde ao vermelho indica a direção de aumento máximo da resposta em relação às variáveis independentes. Este modelo seria empregado para prever as condições ideais de pH e DO para maximizar a produção do composto desejado, e, posteriormente, essas condições otimizadas seriam validadas em fermentações em reator.



Gráfico de superfície de resposta

24 h de cultivo



48 h de cultivo

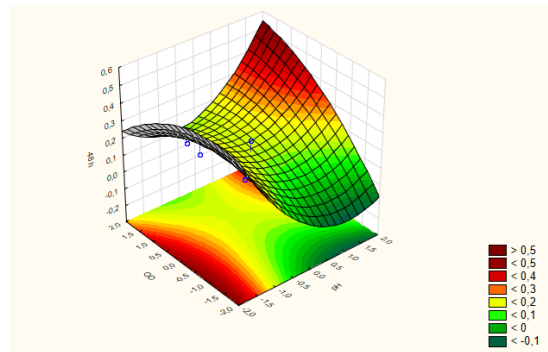


Gráfico de Pareto

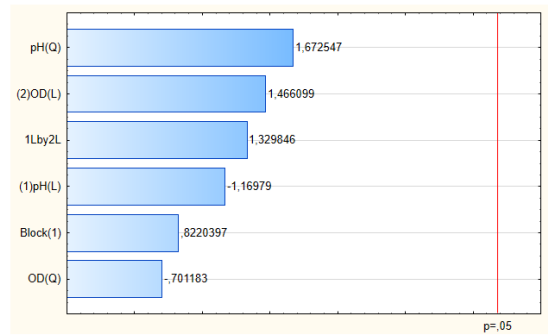
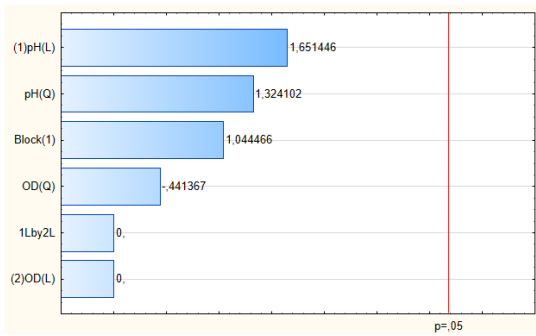
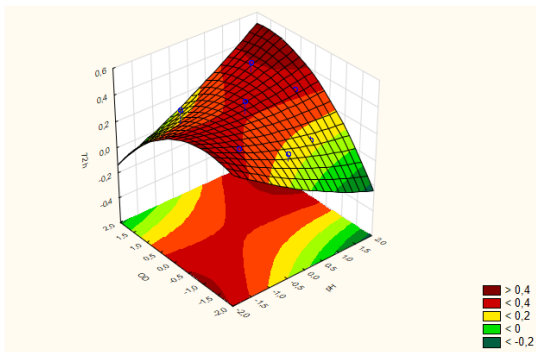


Gráfico de superfície de resposta

72 h de cultivo



96 h de cultivo

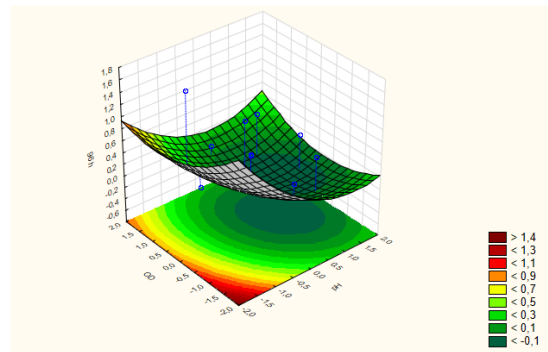


Gráfico de Pareto

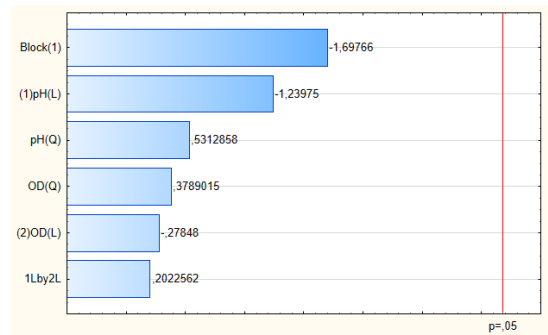
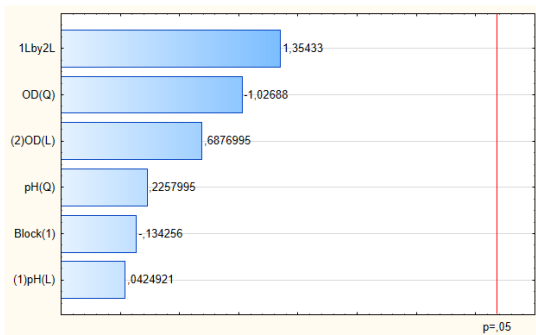


Figura 19. Gráficos de superfície de resposta e de Pareto gerados no programa *Statistisc* a partir dos dados de produção de ácido glicólico obtidos em fermentação em biorreator com a linhagem *K. phaffii* JA122. Fonte: Autor.

Após a construção dos modelos de superfície de resposta, foram feitos diagramas de Pareto no mesmo software utilizado anteriormente, para verificar a significância dos parâmetros pH e DO na produção de AG, com efeitos individuais e de interação (Figura 19). Nesse gráfico, o efeito linear (L) e quadrático (Q) das variáveis e o efeito de sua interação (1Lby2L) é apresentado. No entanto, não se observou uma relação estatisticamente significativa entre as variáveis explicativas (pH e DO) e a variável resposta (AG produzido) em nenhum dos tempos analisados, de maneira linear ou quadrática, considerando significativos os parâmetros cujos  $p$ -valores foram inferiores a 0,05 ( $p < 0,05$ ). A falta de significância dos modelos pode ser atribuída a uma falta de consistência dos dados obtidos, possivelmente devida a erros nos sensores e controladores de DO, que modificaram as condições dos cultivos de tal forma que não foi possível manter os parâmetros estáveis ao longo das 96h de fermentação. Ao longo das bateladas, foram realizados contínuos esforços para o correto funcionamento dos sensores, porém sem sucesso.

Apesar disso, foi observada uma produção de 1,595 g/L e 1,48 g/L de AG nas condições 1 e 3, respectivamente (Tabela 14), valores que superam a maior produção descrita na literatura para leveduras geneticamente modificadas para expressão da via de Dahms, que alcançou concentração de 1 g/L em *S. cerevisiae* (SALUSJÄRVI et al., 2017). Nessas condições, a concentração de AG também foi superior à de EG, coproduto da via de produção de AG, o que indica o potencial da linhagem *K. phaffii* JA122 de atingir maiores concentrações do composto de interesse. No entanto, a falta de significância estatística do modelo preditivo não permite que se tenha conclusões sobre a influência dos parâmetros de pH e DO na produção de AG. Portanto, uma repetição dos ensaios em reator com um ajuste fino dos controladores de DO e a reconstrução dos modelos estatísticos são ainda promissores para a otimização do processo de produção de AG.

Com base nesses resultados, decidiu-se avaliar a produção de AG pela linhagem *K. phaffii* JA122 em batelada alimentada, que permite à levedura atingir maior densidade celular antes da introdução de altas quantidades de xilose no meio de cultura, e previne o acúmulo de subprodutos tóxicos que podem inibir a produção de AG. Além disso, buscou-se avaliar a produção de AG a partir de hidrolisado de biomassa de cana-de-açúcar e comparar os resultados com a produção a partir de meio sintético.

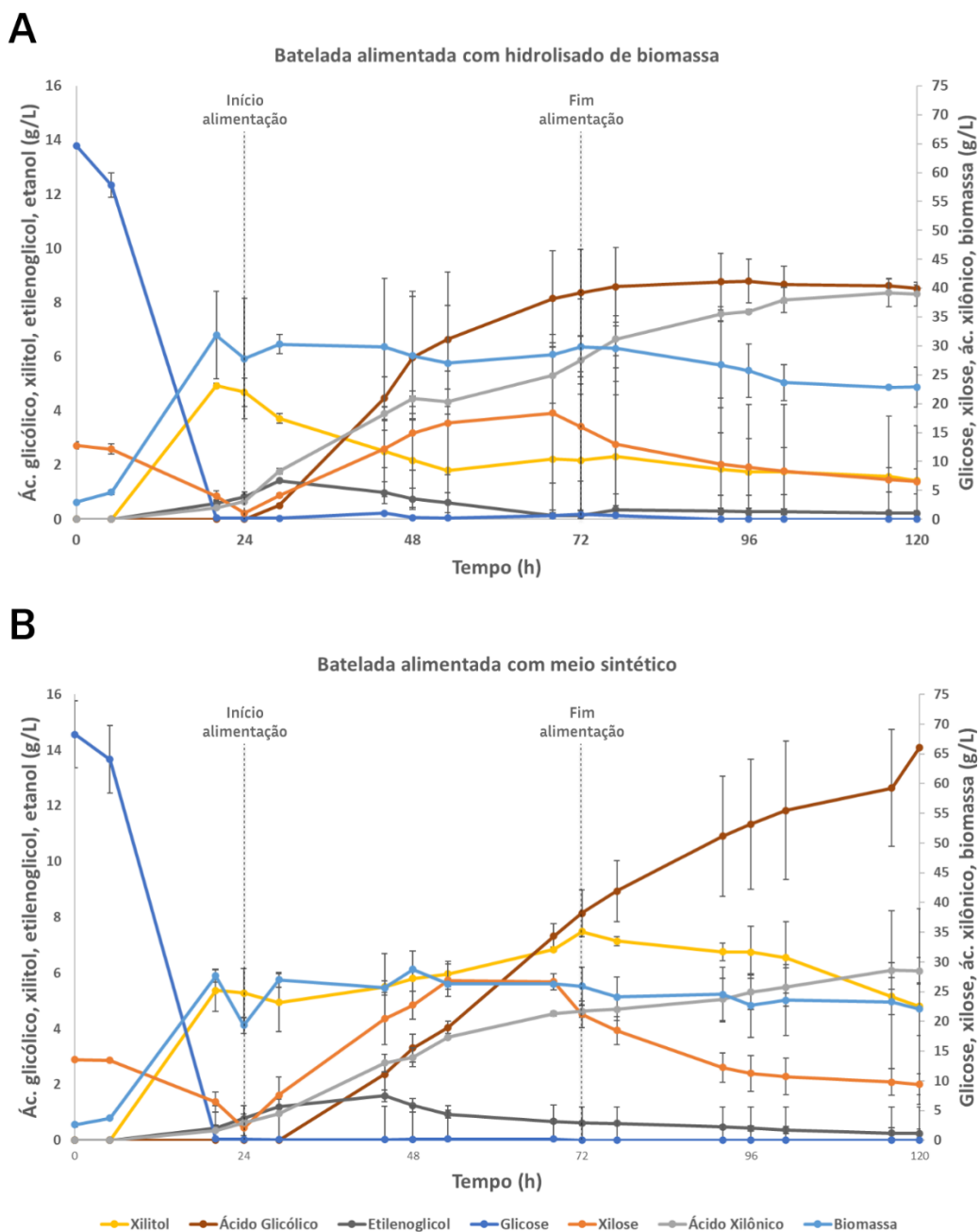
## 5.6. Produção de ácido glicólico por *K. phaffii* JA122 a partir de hidrolisados de biomassa

A fim de otimizar a produção de ácido glicólico (AG) em meio sintético e obter pela primeira vez a produção de AG a partir de hidrolisado de biomassa de cana-de-açúcar, foram idealizadas bateladas alimentadas com a linhagem *K. phaffii* JA122. O modelo de batelada alimentada foi selecionado para que a levedura pudesse alcançar alta densidade celular antes da introdução do hidrolisado, substrato inibitório para o crescimento microbiano, com alta concentração de acetato (15,12 g/L) e presença de compostos tóxicos para o microrganismo, como furfural e hidroximetilfurfural (HMF). Além disso, esse modelo de fermentação permite a adição lenta e gradual desse substrato, o que possibilita à levedura a metabolização dos açúcares presentes no meio de alimentação com menor acúmulo de compostos inibitórios. O pH foi mantido em 6 ao longo da fermentação, conforme indicado na literatura como pH ótimo para o crescimento de *K. phaffii* (INVITROGEN, 2002). O nível de oxigênio dissolvido (DO) foi mantido em 20%, uma vez que experimentos paralelos realizados pelo grupo de pesquisa, que avaliaram a influência desse parâmetro de maneira isolada, demonstraram que altos níveis de oxigenação contribuem positivamente para a produção de AG em detrimento da produção etilenoglicol (EG).

Dessa forma, foram realizadas bateladas alimentadas, em que a levedura foi inicialmente cultivada por 24 h em 900 mL de meio YP suplementado com 70 g/L de glicose e 10 g/L de xilose, garantindo a formação de biomassa celular. Após a primeira fase de batelada, os vasos de fermentação foram alimentados com 1 L de meio YP suplementado com 40 g/L de glicose e 86 g/L de xilose, ou com 1L de hidrolisado suplementado com YP e com glicose e xilose para as mesmas concentrações que o meio definido. Dessa forma, pode-se comparar o efeito da toxicidade do hidrolisado na produção de AG por *K. phaffii*. A fase de alimentação durou 48 h com fluxo de aproximadamente 21 mL/min, e, após essa etapa, a fermentação foi continuada por mais 48 h, com um total de 120 h de cultivo. Amostras foram coletadas a partir do momento do inóculo, e seus componentes foram quantificados em UPLC.

O perfil fermentativo da levedura nas duas condições de alimentação está representado na Figura 20. Em ambas as condições de cultivo, a glicose inicial presente no meio de cultura foi rapidamente consumida pela levedura, antes das primeiras 19 h de fermentação, para atingir alta densidade celular (20-35 g/L). Após o início da etapa de alimentação, não houve acúmulo de glicose no meio de cultura, o que indica o rápido consumo do açúcar pela levedura à medida

que era alimentado. A xilose inicial foi quase completamente consumida durante a etapa inicial de batelada, e se acumulou no meio de cultura durante a fase de alimentação. Após o fim da alimentação, a xilose foi parcialmente consumida, mas foi observada concentração final de  $6,53 \pm 2,14$  g/L de xilose no meio alimentado com hidrolisado, e de  $9,41 \pm 1,77$  g/L no meio alimentado com meio sintético. A grande diferença pode ser parcialmente explicada pelo fato de a alimentação com hidrolisado não ter sido completamente finalizada por conta de precipitados que se acumularam no meio de alimentação e obstruíram as linhas de alimentação. Entre 50 e 60 mL do hidrolisado não foi alimentado aos biorreatores, o que corresponde a  $4,86 \pm 0,67$  g de xilose a menos nesses reatores.



**Figura 20.** Gráficos representativos do perfil fermentativo da levedura *K. phaffii* JA122 em batelada alimentada com (A) hidrolisado de biomassa como meio de alimentação, e (B) meio sintético como meio de alimentação. Fonte: Autor.

A linhagem *K. phaffii* JA122 atingiu produção média final de  $8,51 \pm 0,23$  g/L de ácido glicólico (AG) quando alimentada com hidrolisado de bagaço de cana-de-açúcar, e de  $14,1 \pm 0,04$  g/L quando alimentada com meio sintético. O grande aumento de produção em comparação com as bateladas realizadas de acordo com o modelo do DCCR se deve ao modo de fermentação de batelada alimentada e ao uso de um meio de cultura complexo (YP) em comparação com o meio simples utilizado nas bateladas anteriores (FM22). Os valores de produção atingidos superam em várias vezes a produção máxima de AG relatada na literatura para leveduras, cerca de 1 g/L (SALUSJÄRVI et al., 2017), o que indica o potencial da levedura para o estabelecimento de um processo de produção de AG eficiente a partir de xilose. Apesar disso, esses valores ainda são bastante inferiores àqueles obtidos em *E. coli*, que atingem 65,5 g/L (DENG et al., 2018).

É possível observar que a produção de AG se interrompe no meio alimentado com hidrolisado a partir de 72 horas de fermentação, quando se deu o fim da alimentação, o que não ocorreu na fermentação alimentada com meio sintético. Acredita-se que isso se deve ao acúmulo de compostos tóxicos presentes no hidrolisado, que podem ter tido efeito inibitório sobre a via de produção de AG. No entanto, é possível observar que o ácido xilônico, intermediário da via, continua a ser produzido após o fim da alimentação, o que indica que a inibição por parte dos compostos presentes no hidrolisado se deu sobre as últimas etapas da via.

Mesmo com a alta concentração de acetato e com a presença de compostos inibitórios, como o furfural e o HMF, no hidrolisado de biomassa de cana-de-açúcar, a linhagem JA122 atingiu rendimento produto-substrato ( $Y_{p/s}$ ) de  $0,4 \pm 0$  (mol/mol) nesse substrato, o que representa 39,8% do rendimento teórico máximo ( $Y_{máx}$ ). Quando alimentada com meio sintético, o  $Y_{p/s}$  foi de  $0,66 \pm 0,03$  (mol/mol), o que simboliza 65,9% do  $Y_{máx}$ . Apesar do desempenho da levedura em ambas as fermentações, o  $Y_{p/s}$  reduzido em relação ao  $Y_{máx}$  revela espaço para melhoria da produção de AG, seja por meio do melhoramento genético da linhagem, seja por meio do aperfeiçoamento do processo de produção.

A redução do  $Y_{p/s}$  pode ser explicada por meio da formação de coprodutos ao longo da fermentação, assim como por um acúmulo de um intermediário da via, o ácido xilônico (Figura 6). Uma fração da xilose consumida foi convertida a xilitol, produto de uma via alternativa à via de Dahms, especialmente quando a glicose foi totalmente consumida nos meios de cultura. Foram atingidas concentrações finais de xilitol de  $1,43 \pm 2,02$  g/L e  $4,8 \pm 3,51$  g/L, com hidrolisado e meio sintético como meios de alimentação, respectivamente. Além disso, não foi possível quantificar acetato no meio de cultura, o que indica a rápida assimilação do composto pela levedura.

O acúmulo de ácido xilônico é o maior gargalo de produção de AG, possivelmente causado pela baixa afinidade da xilonato desidratase (XD) heteróloga a esse substrato. Até o presente momento, existem poucas XD's descritas, e os trabalhos publicados reportam baixa atividade para esse grupo de enzimas (JIANG et al., 2015; SALUSJÄRVI et al., 2017; CHAE et al., 2018). Foi observada concentração final de ácido xilônico de  $38,98 \pm 2,03$  g/L com hidrolisado de biomassa como substrato, e de  $28,51 \pm 2,25$  g/L com meio sintético como substrato. A maior concentração de ácido xilônico presente no hidrolisado pode ser explicada pela inibição das últimas etapas da via de produção de AG, o que gerou um acúmulo do intermediário. A produção de etilenoglicol, produto competitivo da via de produção de AG, não foi significativa em nenhuma das duas condições de cultivo. Foi demonstrado, em estudos paralelos realizados pelo grupo de pesquisa, que maiores níveis de oxigenação são positivos para a produção de AG, em detrimento da produção de etilenoglicol, o que pôde ser observado nas bateladas alimentadas realizadas.

Portanto, foi validada a produção de AG a partir de hidrolisado de biomassa como substrato pela linhagem *K. phaffii* JA122. A levedura atingiu concentrações finais de AG superiores às reportadas na literatura para leveduras. Ainda assim, existe espaço para otimização da linhagem produtora e do processo de produção de AG, a fim de estabelecer uma alternativa sustentável à síntese química atualmente empregada para a produção industrial do ácido.

## 6. Conclusão

Em conclusão, este trabalho abre caminho para o estabelecimento de um processo de produção de ácido glicólico (AG) por *K. phaffii* a partir de biomassa lignocelulósica por meio do uso de ferramentas de engenharia genética e otimização de processo.

Dessa forma, pretende-se futuramente construir uma linhagem de levedura que contribua para uma bioeconomia circular com base na valorização de um resíduo abundante, a biomassa vegetal, por meio de sua conversão a um produto de alto valor agregado. Para isso, se faz necessária, como etapa posterior, a avaliação do papel metabólico dos genes-alvo selecionados por meio de sua superexpressão, viabilizada pela construção do vetor pMDHyg e pela amplificação por PCR dos genes-alvo a partir do genoma da *K. phaffii*. Após a obtenção de linhagens de levedura recombinantes, suas capacidades metabólicas serão avaliadas, o que poderá elucidar a via de produção de ácido glicólico (AG) e etilenoglicol (EG).

Também foi avaliada a influência dos parâmetros de pH e oxigênio dissolvido sobre a produção de AG, porém a falta de significância estatística dos modelos construídos faz necessária a repetição dos ensaios fermentativos em biorreator com a linhagem que expressa a via de produção de EG.

Além disso, foi validada a produção de AG a partir de hidrolisado de biomassa de cana-de-açúcar, tal que, nesse substrato, foi atingida concentração de AG superior às relatadas na literatura em outros microrganismos. Mesmo que a produção não tenha alcançado os níveis reportados para algumas bactérias, a levedura demonstrou potencial para atingir esses valores por meio de futura otimização da linhagem e do processo. Assim, o presente trabalho buscou o desenvolvimento de um processo de produção de AG eficiente e sustentável por meio do melhoramento genético de uma levedura recombinante que expresse a via heteróloga de Dahms e por meio da otimização dos parâmetros do processo.

## 7. Referências bibliográficas

ALKIM, Ceren et al. The synthetic xylulose-1 phosphate pathway increases production of glycolic acid from xylose-rich sugar mixtures. **Biotechnology for Biofuels**, v. 9, n. 1, p. 1-11, 2016.

AMARASEKARA, Ananda S.; DENG, Fang. Acidic ionic liquid catalyzed liquefaction of untreated switchgrass biomass in acetone and Pd-La (OTf)<sub>3</sub> catalyzed reduction of the products. **Biomass and Bioenergy**, v. 127, p. 105260, 2019.

ASHOKKUMAR, Veeramuthu et al. Recent advances in lignocellulosic biomass for biofuels and value-added bioproducts-A critical review. **Bioresource technology**, v. 344, p. 126195, 2022.

ATA, Özge et al. A single Gal4-like transcription factor activates the Crabtree effect in *Komagataella phaffii*. **Nature Communications**, v. 9, n. 1, p. 4911, 2018.

ATA, Özge et al. What makes *Komagataella phaffii* non-conventional?. **FEMS Yeast Research**, v. 21, n. 8, p. foab059, 2021.

BARONE, Giovanni Davide et al. Industrial production of proteins with *Pichia pastoris*—*Komagataella phaffii*. **Biomolecules**, v. 13, n. 3, p. 441, 2023.

BEIG, Bilal et al. Current challenges and innovative developments in pretreatment of lignocellulosic residues for biofuel production: A review. **Fuel**, v. 287, p. 119670, 2021.

BUDAK, Kamil; SOGUT, Oguz; AYDEMIR SEZER, Umran. A review on synthesis and biomedical applications of polyglycolic acid. **Journal of polymer research**, v. 27, p. 1-19, 2020.

CABULONG, Rhudith B. et al. Engineering *Escherichia coli* for glycolic acid production from D-xylose through the Dahms pathway and glyoxylate bypass. **Applied microbiology and biotechnology**, v. 102, p. 2179-2189, 2018.

CABULONG, Rhudith B. et al. Enhanced glycolic acid yield through xylose and cellobiose utilization by metabolically engineered *Escherichia coli*. **Bioprocess and Biosystems Engineering**, v. 44, p. 1081-1091, 2021.



CAM, Yvan et al. Engineering of a synthetic metabolic pathway for the assimilation of (d)-xylose into value-added chemicals. **ACS synthetic biology**, v. 5, n. 7, p. 607-618, 2016.

CARNEIRO, Clara Vida Galvão Corrêa et al. Advances in *Komagataella phaffii* Engineering for the Production of Renewable Chemicals and Proteins. **Fermentation**, v. 8, n. 11, p. 575, 2022.

CHAE, Tong Un et al. Production of ethylene glycol from xylose by metabolically engineered *Escherichia coli*. **AIChE Journal**, v. 64, n. 12, p. 4193-4200, 2018.

DAHMS, A. Stephen. 3-Deoxy-D-pentulosonic acid aldolase and its role in a new pathway of D-xylose degradation. **Biochemical and biophysical research communications**, v. 60, n. 4, p. 1433-1439, 1974.

DE ALMEIDA, João Ricardo Moreira; DE MORAES, Lídia Maria Pepe; TORRES, Fernando Araripe Gonçalves. Molecular characterization of the 3-phosphoglycerate kinase gene (PGK1) from the methylotrophic yeast *Pichia pastoris*. **Yeast**, v. 22, n. 9, p. 725-737, 2005.

DE SCHUTTER, Kristof et al. Genome sequence of the recombinant protein production host *Pichia pastoris*. **Nature biotechnology**, v. 27, n. 6, p. 561-566, 2009.

DEIVAYANAI, V. C. et al. A comprehensive review on the biological conversion of lignocellulosic biomass into hydrogen: Pretreatment strategy, technology advances and perspectives. **Bioresource Technology**, p. 128166, 2022.

DENG, Yu et al. Balancing the carbon flux distributions between the TCA cycle and glyoxylate shunt to produce glycolate at high yield and titer in *Escherichia coli*. **Metabolic engineering**, v. 46, p. 28-34, 2018.

ERGÜN, Burcu Gündüz et al. Engineered Dereglulation of Expression in Yeast with Designed Hybrid-Promoter Architectures in Coordination with Discovered Master Regulator Transcription Factor. **Advanced Biosystems**, v. 4, n. 4, p. 1900172, 2020.

FISCHER, Jasmin E.; GLIEDER, Anton. Current advances in engineering tools for *Pichia pastoris*. **Current opinion in biotechnology**, v. 59, p. 175-181, 2019.

FRANCOIS, Jean Marie; ALKIM, Ceren; MORIN, Nicolas. Engineering microbial pathways for production of bio-based chemicals from lignocellulosic sugars: current status and perspectives. **Biotechnology for biofuels**, v. 13, p. 1-23, 2020.

GAO, Jucan; JIANG, Lihong; LIAN, Jiazhang. Development of synthetic biology tools to engineer *Pichia pastoris* as a chassis for the production of natural products. **Synthetic and systems biotechnology**, v. 6, n. 2, p. 110-119, 2021.

GASSLER, Thomas et al. CRISPR/Cas9-mediated homology-directed genome editing in *Pichia pastoris*. **Recombinant protein production in yeast**, p. 211-225, 2019.

HE, Yu-Cai et al. Bioproduction of glycolic acid from glycolonitrile with a new bacterial isolate of *Alcaligenes* sp. ECU0401. **Applied Biochemistry and Biotechnology**, v. 160, p. 1428-1440, 2010.

HEISTINGER, Lina; GASSER, Brigitte; MATTANOVICH, Diethard. Microbe Profile: *Komagataella phaffii*: A methanol devouring biotech yeast formerly known as *Pichia pastoris*. **Microbiology**, v. 166, n. 7, p. 614-616, 2020.

HOUSHMAND, Elizabeth Bahar. Effect of glycolic acid, phytic acid, soothing complex containing Emulsion on Hyperpigmentation and skin luminosity: A clinical evaluation. **Journal of Cosmetic Dermatology**, v. 20, n. 3, p. 776-780, 2021.

INVITROGEN, C. O. *Pichia* fermentation process guidelines [em linha]. 2002.

ISOBE, Kimiyasu et al. Microbial oxidases catalyzing conversion of glycolaldehyde into glyoxal. **New biotechnology**, v. 29, n. 2, p. 177-182, 2012.

ITO, Yoichiro et al. Deletion of DNA ligase IV homolog confers higher gene targeting efficiency on homologous recombination in *Komagataella phaffii*. **FEMS yeast research**, v. 18, n. 7, p. foy074, 2018.

JAGTAP, Sujit Sadashiv; RAO, Christopher V. Microbial conversion of xylose into useful bioproducts. **Applied microbiology and biotechnology**, v. 102, p. 9015-9036, 2018.

JAYAKUMAR, Mani et al. Bioethanol production from agricultural residues as lignocellulosic biomass feedstock's waste valorization approach: A comprehensive review. **Science of The Total Environment**, v. 879, p. 163158, 2023.

JIANG, Yudong et al. Characterization of D-xylonate dehydratase YjhG from *Escherichia coli*. **Bioengineered**, v. 6, n. 4, p. 227-232, 2015.

JIAO, J.; WANG, S.; LIANG, M.; ZHANG, Y.; XU, X.; ZHANG, W.; LIU, B. Basal transcription profiles of the rhamnose-inducible promoter PLRA3 and the development of efficient PLRA3-based systems for markerless gene deletion and a mutant library in *Pichia pastoris*. **Current Genetics**, [s. l.], v. 65, n. 3, p. 785–798, 2019.

KAKKO, Natalia et al. Inducible Synthetic Growth Regulation Using the ClpXP Proteasome Enhances cis, cis-Muconic Acid and Glycolic Acid Yields in *Saccharomyces cerevisiae*. **ACS Synthetic Biology**, v. 12, n. 4, p. 1021-1033, 2023.

KARAOĞLAN, Mert et al. Identification of major ADH genes in ethanol metabolism of *Pichia pastoris*. **Yeast**, v. 37, n. 2, p. 227-236, 2020.

KATAOKA, Michihiko et al. Glycolic acid production using ethylene glycol-oxidizing microorganisms. **Bioscience, biotechnology, and biochemistry**, v. 65, n. 10, p. 2265-2270, 2001.

KIM, Jiwon; HWANG, Sungmin; LEE, Sun-Mi. Metabolic engineering for the utilization of carbohydrate portions of lignocellulosic biomass. **Metabolic Engineering**, v. 71, p. 2-12, 2022.

KOIVISTOINEN, Outi M. et al. Glycolic acid production in the engineered yeasts *Saccharomyces cerevisiae* and *Kluyveromyces lactis*. **Microbial cell factories**, v. 12, p. 1-16, 2013.

LACHAUX, Cléa et al. A new synthetic pathway for the bioproduction of glycolic acid from lignocellulosic sugars aimed at maximal carbon conservation. **Frontiers in bioengineering and biotechnology**, v. 7, p. 359, 2019.

LI, Wei et al. Microbial production of glycolate from acetate by metabolically engineered *Escherichia coli*. **Journal of biotechnology**, v. 291, p. 41-45, 2019.

LIAO, S.; TIAN, Q.; XIAO, Y.; QIN, D.; LI, J.; HU, C. Glycerol Valorization Towards Glycolic Acid Production Over Cu-Based Biochar Catalyst. **ChemSusChem**, [s. l.], v. 15, n. 23, p. e202201537, 2022.

LIU, B.; ZHANG, Yuwei; ZHANG, X.; YAN, C.; ZHANG, Yuhong; XU, X.; ZHANG, W. Discovery of a rhamnose utilization pathway and rhamnose-inducible promoters in *Pichia pastoris*. **Scientific Reports**, [s. l.], v. 6, n. 1, p. 27352, 2016.

LU, Xiyang et al. Ethylene glycol and glycolic acid production by wild-type *Escherichia coli*. **Biotechnology and Applied Biochemistry**, v. 68, n. 4, p. 744-755, 2021.

LU, Yue et al. Properties of Poly (Lactic-co-Glycolic Acid) and Progress of Poly (Lactic-co-Glycolic Acid)-Based Biodegradable Materials in Biomedical Research. **Pharmaceuticals**, v. 16, n. 3, p. 454, 2023.

MENSAH, Michael B. et al. Assessment of quantities and composition of corn stover in Ghana and their conversion into bioethanol. **Scientific African**, v. 12, p. e00731, 2021.

ÑANCUCHEO, Ivan; JOHNSON, D. Barrie. Production of glycolic acid by chemolithotrophic iron-and sulfur-oxidizing bacteria and its role in delineating and sustaining acidophilic sulfide mineral-oxidizing consortia. **Applied and Environmental Microbiology**, v. 76, n. 2, p. 461-467, 2010.

OPEC. **OPEC 2023 Annual Report**. [S. l.: s. n.], 2024.

PEREIRA, Brian et al. Efficient utilization of pentoses for bioproduction of the renewable two-carbon compounds ethylene glycol and glycolate. **Metabolic engineering**, v. 34, p. 80-87, 2016.

PERRONE, Olavo Micali et al. Improving cellulosic ethanol production using ozonolysis and acid as a sugarcane biomass pretreatment in mild conditions. **Bioresource Technology Reports**, v. 13, p. 100628, 2021.

PRIELHOFER, R.; BARRERO, J. J.; STEUER, S.; GASSLER, T.; ZAHRL, R.; BAUMANN, K.; SAUER, M.; MATTANOVICH, D.; GASSER, B.; MARX, H. GoldenPiCS: a Golden Gate-derived modular cloning system for applied synthetic biology in the yeast *Pichia pastoris*. **BMC Systems Biology**, [s. l.], v. 11, n. 1, p. 123, 2017.

RAMOS, Talita GS et al. Xylonic acid production by recombinant *Komagataella phaffii* strains engineered with newly identified xylose dehydrogenases. **Bioresource Technology Reports**, v. 16, p. 100825, 2021.

RAO, Jun et al. Hemicellulose: Structure, chemical modification, and application. **Progress in Polymer Science**, p. 101675, 2023.

RUCHALA, Justyna; SIBIRNY, Andriy A. Pentose metabolism and conversion to biofuels and high-value chemicals in yeasts. **FEMS Microbiology Reviews**, v. 45, n. 4, p. fuaa069, 2021.

SALUSJÄRVI, Laura et al. Production of ethylene glycol or glycolic acid from D-xylose in *Saccharomyces cerevisiae*. **Applied microbiology and biotechnology**, v. 101, p. 8151-8163, 2017.

SALUSJÄRVI, Laura et al. Biotechnological production of glycolic acid and ethylene glycol: current state and perspectives. **Applied microbiology and biotechnology**, v. 103, p. 2525-2535, 2019.

SARKER, Tumpa R. et al. Hydrothermal pretreatment technologies for lignocellulosic biomass: A review of steam explosion and subcritical water hydrolysis. **Chemosphere**, v. 284, p. 131372, 2021.

SHONZA, N. S.; ANDREATTA, D.; MUNIZ, E. P.; DALMASCHIO, C. J.; DE FREITAS, R. R.; PORTO, P. S. D. S. Crude oil wastewater treatment by electrocoagulation in a continuous process with polarity switch. **Environmental Technology**, [s. l.], v. 43, n. 11, p. 1700–1708, 2022.

STURMBERGER, Lukas et al. Refined *Pichia pastoris* reference genome sequence. **Journal of biotechnology**, v. 235, p. 121-131, 2016.

TIAN, Wenjing et al. Effects of hydrothermal pretreatment on the mono-and co-digestion of waste activated sludge and wheat straw. **Science of The Total Environment**, v. 732, p. 139312, 2020.

TOIVARI, Mervi et al. Low pH D-xylonate production with *Pichia kudriavzevii*. **Bioresource technology**, v. 133, p. 555-562, 2013.

TOIVARI, Mervi *et al.* **Improved production of oxalyl-coa, glyoxylate and/or glycolic acid.** Depositante: Valtion Teknillinen Tutkimuskeskus. EP3658677A1. Depósito: 27 jul. 2018.

TRICHEZ, Débora et al. Recent progress in the microbial production of xylonic acid. **World Journal of Microbiology and Biotechnology**, v. 38, n. 7, p. 127, 2022.

WINARSIH, Sri; SISKAWARDANI, Devi Dwi. Hydrolysis of corncobs using a mixture of crude enzymes from *Trichoderma reesei* and *Aspergillus niger* for bioethanol production. **Energy Reports**, v. 6, p. 256-262, 2020.

YU, Yong et al. Construction of a carbon-conserving pathway for glycolate production by synergetic utilization of acetate and glucose in *Escherichia coli*. **Metabolic Engineering**, v. 61, p. 152-159, 2020.

ZAHOOR, Ahmed; OTTEN, Andreas; WENDISCH, Volker F. Metabolic engineering of *Corynebacterium glutamicum* for glycolate production. **Journal of biotechnology**, v. 192, p. 366-375, 2014.

ZAVEC, Domen et al. Beyond alcohol oxidase: the methylotrophic yeast *Komagataella phaffii* utilizes methanol also with its native alcohol dehydrogenase Adh2. **FEMS yeast research**, v. 21, n. 2, p. foab009, 2021.

ZHA, Jian et al. Advances in Metabolic Engineering of *Pichia pastoris* Strains as Powerful Cell Factories. **Journal of Fungi**, v. 9, n. 10, p. 1027, 2023.

ZHAN, Tao et al. Constructing a novel biosynthetic pathway for the production of glycolate from glycerol in *Escherichia coli*. **ACS Synthetic Biology**, v. 9, n. 9, p. 2600-2609, 2020.

ZHANG, Zhongxi et al. Ethylene glycol and glycolic acid production from xylonic acid by *Enterobacter cloacae*. **Microbial Cell Factories**, v. 19, n. 1, p. 1-16, 2020.

ZHAO, Lei et al. Advances in pretreatment of lignocellulosic biomass for bioenergy production: Challenges and perspectives. **Bioresource Technology**, v. 343, p. 126123, 2022.

ZHAO, Zhe et al. Biochemical routes for uptake and conversion of xylose by microorganisms. **Biotechnology for biofuels**, v. 13, p. 1-12, 2020.

ZHOU, Xin et al. Glycolic acid production from ethylene glycol via sustainable biomass energy: integrated conceptual process design and comparative techno-economic–society–environment analysis. **ACS Sustainable Chemistry & Engineering**, v. 9, n. 32, p. 10948-10962, 2021.

ZHOU, Hualan et al. Efficient expression of heterologous protein by engineered *Komagataella phaffii* by harnessing a bioelectrical CO<sub>2</sub> reduction system. **Biochemical Engineering Journal**, v. 191, p. 108762, 2023.