



Universidade de Brasília  
Instituto de Ciências Biológicas  
Departamento de Biologia Celular  
Grupo de Genética Molecular e Biotecnologia de Leveduras

**Desenvolvimento de linhagens de *Komagataella phaffii*  
contendo novos sistemas de controle da expressão gênica  
livres de metanol**

**Igor Patrick Vasconcelos Vieira**

**Orientador:** Prof. Fernando Araripe Gonçalves Torres

Brasília, 2024.

Universidade de Brasília  
Instituto de Ciências Biológicas  
Departamento de Biologia Celular  
Grupo de Genética e Biotecnologia de Leveduras

**Desenvolvimento de linhagens de *Komagataella phaffii*  
contendo novos sistema de controle da expressão gênica  
livres de metanol**

Tese de doutorado apresentada  
ao Departamento de Biologia  
Celular do Instituto de Ciências  
Biológicas da Universidade de  
Brasília como requisito parcial  
para a obtenção do título de  
Doutor em Biologia Molecular.

**Igor Patrick Vasconcelos Vieira**

**Orientador:** Prof. Fernando Araripe Gonçalves Torres

Trabalho realizado no Laboratório de Biologia Molecular, Departamento de Biologia Celular, Instituto de Ciências Biológicas, Universidade de Brasília, sob a orientação do professor Fernando Araripe Gonçalves Torres e com apoio financeiro do Conselho Nacional de Desenvolvimento científico e Tecnológico (CNPq) e Fundação de Apoio à Pesquisa do Distrito Federal (FAPDF).

## **Banca Examinadora**

---

Fernando Araripe Gonçalves Torres  
Universidade de Brasília

---

Cintia Marques Coelho  
Universidade de Brasília

---

Milca Rachel da Costa Ribeiro Lins  
Universidade Federal do ABC

---

João Ricardo Moreira de Almeida  
Embrapa Agroenergia

---

Lidia Maria Pepe de Moraes  
Universidade de Brasília  
Membro Suplente

*“You make your own luck.”*

## AGRADECIMENTOS

À minha família, que me deu apoio financeiro e emocional imensurável durante toda a minha jornada científica. Em especial à minha mãe, que durante toda a vida se sacrificou para que seus filhos pudessem galgar os voos mais altos que a vida lhes permitisse; à minha avó, que sempre sonhou em ver seu neto se formar “dotô”; ao meu pai e meu irmão, que por tantas vezes estiveram ao meu lado e apoiaram todas as minhas decisões. Fica aqui perpetuado meu mais sincero **Obrigado!**

Ao meu orientador Fernando Araripe, que por vários momentos foi também amigo e conselheiro. Este momento não poderia ter acontecido sem você ter topado seguir comigo quando lhe procurei via e-mail em 2016.

À prof. Lidia Pepe, que esteve comigo para uma conversa, um lanche, uma discussão científica, assistir séries, falar sobre filmes ou qualquer outro assunto. Uma verdadeira amiga e pessoa de alma ímpar.

Aos demais amigos de laboratório que estão e estiveram ao meu lado desde que iniciei minha vida acadêmica em Brasília. Todos os momentos de risada, de estresse e mau humor; de discussões técnicas e/ou filosóficas e de comilança certamente fizeram toda a diferença para que eu me mantivesse são durante todo o processo.

À profa. Ildinete que desde o primeiro semestre acompanhou meu projeto e opinou ativamente. Seu afincamento em aprender sobre o meu projeto me impressiona até hoje. Existem poucos professores com tão genuíno interesse em ajudar um aluno sem nenhuma obrigação para qual.

Aos colegas de corredor e outros laboratórios que sempre me forneceram reagentes e tiraram dúvidas nos momentos de necessidade. Sem vocês muitas coisas não seriam possíveis.

À minha esposa Ray, que compartilhou dos momentos de elaboração deste trabalho e me deu total apoio para perseguir meu sonho e transpassar os obstáculos.

## Sumário

Lista de figuras .....	8
Lista de tabelas .....	9
Lista de abreviatura e símbolos .....	9
Abstract .....	12
1. Introdução.....	13
1.1 A crescente demanda por inovação e produtos biotecnológicos.....	13
1.2 Biologia sintética e engenharia metabólica .....	13
1.3 A levedura <i>Komagataella phaffii</i> .....	14
1.4 Ferramentas para a manipulação genética de <i>K. phaffii</i> .....	16
1.5 Controle de expressão gênica e promotores utilizados em <i>K. phaffii</i> .	19
1.6 O sistema de expressão gênica regulado por galactose em <i>S. cerevisiae</i> .....	20
1.7 Aptâmeros como sequências controladoras de expressão gênica e proteica .....	22
2. Objetivos.....	25
a) Desenvolver uma linhagem de <i>K. phaffii</i> com um sistema de controle de expressão gênica baseado no regulon GAL de <i>S. cerevisiae</i> .....	25
Objetivos específicos: .....	25
b) Desenvolver uma linhagem de <i>K. phaffii</i> com um sistema de controle de expressão proteica baseado no <i>riboswitch</i> sensível à tetraciclina.....	25
Objetivos específicos: .....	25
2. Estratégia .....	1
3. Materiais e Métodos .....	1
4.2. Meios de cultivo .....	1
4.3 Soluções .....	2
4.4 Preparo de células bacterianas competentes por choque térmico .....	4
4.5 Preparo de células bacterianas competentes por eletroporação.....	5
4.6 Extração de plasmídeos de <i>E. coli</i> por lise alcalina .....	6
4.7 Análise de sequências de DNA por homologia .....	7
4.8 Amplificação por PCR.....	7
4.9 Construção dos plasmídeos pV2A-T para o cassete policistrônico GAL .....	13
4.10 Transformação de leveduras por eletroporação .....	14
4.11 Microscopia de fluorescência e triagem de transformantes positivos.	15
4.12 Extração de proteínas de <i>K. phaffii</i> .....	15

4.13	Dosagem de proteínas solúveis em extrato total de <i>K. phaffii</i> .....	16
4.14	Extração rápida de DNA genômico para confirmação de transformantes .....	17
4.15	Construção dos vetores de expressão de <i>KU70</i> e <i>lacZ</i> .....	17
4.16	Atividade de $\beta$ -galactosidase.....	19
4.17	Análise estatística .....	20
4.18	Eletroforese de proteínas em gel de poliacrilamida (PAGE) e Western Blotting .....	21
5	Resultados e Discussão .....	22
5.1	Desenvolvimento da linhagem de <i>K. phaffii</i> contendo o sistema de controle de expressão gênica baseado no regulon GAL de <i>S. cerevisiae</i> .....	22
5.1.1	Construção do sistema repórter ativado por galactose .....	22
5.1.2	Construção e inserção do regulon GAL em <i>K. phaffii</i> .....	27
5.2	Desenvolvimento da linhagem de <i>K. phaffii</i> responsiva à tetraciclina	43
6	Conclusões e perspectivas.....	50
7	Referências bibliográficas .....	50

## Lista de figuras

Figura 1: O aumento do interesse pelo uso comercial de <i>K. phaffii</i> gerou o crescimento do número de artigos científicos publicados. ....	16
Figura 2: Representação gráfica do circuito genético ativado por galactose. ....	22
Figura 3: Mecanismo de ação de um <i>riboswitch</i> . ....	24
Figura 4: Estratégias utilizadas na montagem das linhagens de controle da expressão gênica e proteica em <i>K. phaffii</i> . ....	1
Figura 5: Estratégia de montagem do plasmídeo epissomal pPICH-Pgal1GFP. ....	25
Figura 6: Digestão do plasmídeo pYES2 para retirada do gene eGFP. ....	26
Figura 7: Confirmação da correta montagem do plasmídeo pPICH-GFP. ....	26
Figura 8: Representação do cassete policistrônico contendo os genes da família GAL. ....	29
Figura 9: Amplificação dos elementos do regulon GAL. ....	29
Figura 10: Fragmentos de PCR obtidos a partir do DNA genômico dos clones transformantes de <i>K. phaffii</i> com o regulon GAL. ....	31
Figura 11: Amplificação por PCR dos genes do regulon GAL com o conjunto de oligonucleotídeos ampRegal. ....	31
Figura 12: Mapa do vetor pPIC9 utilizado para montagem do regulon GAL em <i>E. coli</i> . ....	32
Figura 13: Fragmentos de PCR amplificados com os oligonucleotídeos ampRegal após purificação. ....	34
Figura 14: Amplificação do vetor pPIC9 por PCR. ....	35
Figura 15: Montagem do regulon GAL no vetor pPIC9m em <i>E. coli</i> . ....	35
Figura 16: Fragmento de PCR amplificado do vetor pPIC9Gal contendo o promotor PTEF1 e o gene <i>GAL2</i> . ....	36
Figura 17: Descrição visual do vetor pV2A-T. ....	37
Figura 18: Esquema de montagem do cassete policistrônico com auxílio do vetor pV2A-T. ....	39
Figura 19: Eletroforese em gel de agarose dos fragmentos de DNA utilizados para construção do regulon GAL no plasmídeo pV2A-T. ....	40
Figura 20: Confirmação da construção do vetor pV2A_regal contendo o cassete policistrônico para expressão do regulon GAL em <i>K. phaffii</i> . ....	41
Figura 21: Prova de conceito do funcionamento do regulon GAL em <i>K. phaffii</i> . ....	43
Figura 22: Mapa dos plasmídeos contendo o aptâmero de tetraciclina (tc) controlando a produção de proteínas. ....	44
Figura 23: Mecanismo de funcionamento do <i>riboswitch</i> durante a tradução do mRNA. ....	45
Figura 24: O aptâmero de tetraciclina reduz a tradução de $\beta$ -galactosidase em <i>K. phaffii</i> . ....	46
Figura 25: Atividade de $\beta$ -galactosidase é reduzida após exposição da célula a tetraciclina. ....	47



Figura 26: Redução da tradução de Ku70p após ativação do *riboswitch* de tetraciclina. .... 49

## Lista de tabelas

Tabela 1: Oligonucleotídeos utilizados neste trabalho.. .... 8

## Lista de abreviatura e símbolos

$\Omega$	Ohm
$^{\circ}\text{C}$	Graus celsius
$\mu\text{F}$	Micro Faraday
$\mu\text{g}$	Micrograma
$\mu\text{g}$	Micrograma
$\mu\text{L}$	Microlitro
A	Ampere
ARS	<i>Autonomously replicating sequence</i>
BSA	Bovine Serum Albumin
DNA	Ácido desoxirribonucleico
$\text{DO}_{600}$	Densidade celular óptica medida no comprimento de onda 600 nm
g	Gramma
$g$	Aceleração gravitacional
G418	Geneticina
h	Hora
kb	Quilobase
kV	Quilovolts
L	Litro
LB	Luria-Bertani
M	Molar
MD	<i>Minimal Dextrose</i>
min	Minuto
mL	Mililitro
mM	Milimolar
MOPS	Ácido 3-(n-morfolino) propanosulfônico
ng	Nanograma
nm	Nanômetro
ORF	<i>Open Reading frame</i>
pb	Pares de bases
PCR	Reação em cadeia da polimerase

pH	Potencial hidrogeniônico
RNA	Ácido ribonucleico
rpm	Rotações por minuto
UTR	Untranslated region
UAS	Upstream activation sequence
YPD	Yeast Peptone Dextrose

## Resumo

*Komagataella phaffii* (anteriormente conhecida como *Pichia pastoris*) é uma levedura considerada como uma das mais importantes plataformas disponíveis para produção de proteínas e metabólitos em larga escala da atualidade, devida a sua alta produtividade e fácil controle dos parâmetros fermentativos. Sua característica fisiológica mais distinta é a capacidade de usar metanol como fonte única de carbono graças à hiperexpressão do gene *AOX1* que codifica para a enzima peroxissomal álcool oxidase 1. O promotor  $P_{AOX1}$  permite a alta expressão com controle finamente regulado pela concentração de metanol, o que o tornou o sistema de expressão recombinante mais utilizado neste organismo. No entanto, a maior limitação deste sistema no contexto de bioprocessos industriais é o fato do metanol ser altamente tóxico e inflamável. Faz-se necessário o desenvolvimento de um sistema de indução da expressão gênica baseado em um agente indutor atóxico, preferencialmente não consumido pela célula, a fim de diminuir os custos de produção em larga escala.

Neste trabalho foram testados dois mecanismos de controle de expressão: à nível transcricional, pelo sistema GAL proveniente de *S. cerevisiae* buscando a indução da produção da proteína recombinante fluorescente verde (GFP) sob controle do promotor  $P_{GAL1}$  regulado por galactose e o sistema de controle de expressão a nível traducional utilizando um *riboswitch* controlado por tetraciclina buscando inibir a produção das proteínas  $\beta$ -galactosidase e Ku70p. Os resultados demonstraram sucesso na utilização do *riboswitch* como mecanismo de controle da produção de proteína em *K. phaffii* com fácil aplicação e potencial grande impacto na aplicação deste sistema em larga escala.

Palavras chave: *Komagataella phaffii*; riboswitch; tetraciclina; aptâmero; controle de expressão gênica;

## Abstract

*Komagataella phaffii* (formerly *Pichia pastoris*) is considered one of the most important platforms available for large-scale production of proteins and metabolites today, due to its high productivity and easy control of fermentative parameters. Its most distinctive physiological feature is the ability to use methanol as a sole carbon source, achieved through the hyperexpression of the *AOX1* gene encoding the peroxisomal alcohol oxidase enzyme. The  $P_{AOX1}$  promoter allows high expression with finely regulated control by methanol concentration, making it the most widely used recombinant expression system in this organism. However, the main limitation of this system in the context of industrial bioprocesses is the highly toxic and flammable nature of methanol. The development of a gene expression induction system based on a non-toxic inducer agent, preferably not consumed by the cell, is necessary to reduce production costs on a large scale.

In this work, two expression control mechanisms were tested: at the transcriptional level we used the GAL regulation system from *S. cerevisiae* to control the transcription of the green fluorescent protein (GFP) under the control of the  $P_{GAL1}$  promoter regulated by galactose. At the translational level, expression control was achieved using a tetracycline-controlled riboswitch aiming to inhibit the production of  $\beta$ -galactosidase and Ku70p proteins. The results demonstrated success in using the riboswitch as a mechanism for controlling protein production in *K. phaffii* with easy application and significant potential impact on the application of this system on a large scale.

Keywords: *Komagataella phaffii*; riboswitch; tetracycline; aptamer; translational control of gene expression.

## **1. Introdução**

### **1.1 A crescente demanda por inovação e produtos biotecnológicos**

Biotecnologia é definida como a exploração de processos biológicos para propósitos industriais, utilizando manipulações genéticas de microrganismos para produção de produtos de interesse variados. (Dicionário Oxford, 3ª edição, 2018). Em especial, considera-se que o desenvolvimento de tecnologias de produção de biocombustíveis e biofármacos estão entre as prioridades humanitárias modernas. A crescente demanda energética, necessidade de transição de matriz energética, os altos custos de desenvolvimento e produção de novas drogas, acesso universal à saúde e elevação dos custos de produção tradicionais impulsionam os investimentos nesta área (ACHARYA 2004). Neste contexto, a biologia sintética e a biotecnologia industrial ganham destaque e fazem parte do ecossistema de Bioeconomia, que promove a criação de produtos, serviços, empresas e marcos regulatórios necessários para que estas tecnologias sejam utilizadas de forma estratégica e possam alcançar a população (CLARKE 2020).

### **1.2 Biologia sintética e engenharia metabólica**

A biologia sintética compreende um conjunto de tecnologias que combina partes biológicas de maneira a construir genes, proteínas, vias metabólicas e/ou organismos completamente novos (CHURCH *et al.*, 2014). É uma área multidisciplinar, abrangendo a biologia computacional e molecular, a engenharia metabólica e a biologia de sistemas com o objetivo de criar ou conferir uma nova função a um organismo, desenhando circuitos com elementos que se combinam para atingir a resposta desejada. Esta resposta pode ser uma adaptação de vias metabólicas existentes ou até a criação de circuitos e organismos completamente novos (JENSEN; KEASLING, 2015; PICATAGGIO, 2009).

As aplicações deste tipo de tecnologia incluem a construção de organismos capazes de sintetizar moléculas de interesse comercial, como biocombustíveis, fármacos ou intermediários de indústrias químicas e alimentícias (D'ESPAUX *et al.*, 2015; DHAMANKAR; PRATHER, 2011; GALANIE, 2015). Como exemplos de aplicações da biologia sintética em *K.*

*phaffii*, pode-se citar a construção de enzimas semissintéticas que modificam o padrão de glicosilação da levedura e construções policistrônicas de até 9 genes (GEIER *et al.*, 2015a; VOGL; HARTNER; GLIEDER, 2013). O sistema CRISPR-Cas9 também foi utilizado no nocaute eficiente de genes em *K. phaffii* (WENINGER *et al.*, 2016, 2018).

Construções de DNA usadas em biologia sintética podem incluir diversas enzimas, sítios de ligação e proteínas regulatórias, fatores de transcrição e até polimerases sintéticas, compondo grandes extensões de DNA (NIELSEN; SEGALL-SHAPIRO; VOIGT, 2013; PINHEIRO *et al.*, 2012). Para permitir a correta montagem destes grandes cassetes de expressão heteróloga, foi necessário um grande esforço no desenvolvimento de novas tecnologias. Classicamente, diferentes fragmentos de DNA eram unidos através de reações mediadas por enzimas de restrição e ligases. No entanto, esta metodologia pode ser bastante demorada e relativamente difícil dependendo dos sítios de restrição utilizados. Sendo assim, na última década, novos sistemas de clonagem foram disponibilizados, podendo ser divididos em quatro classes principais: baseados em enzimas de restrição (ex.: Bio-Brick e Golden Gate); reações com sequências de homologia *in vivo* (ex.: recombinação associada à transformação em leveduras e bactérias) e *in vitro* (ex.: Gibson Assembly); e com anelamento de oligonucleotídeos (ex.: LCR – Ligase Chain Reaction e CPEC - reação de polimerase circular) (QUAN; TIAN, 2011 ; GOMES *et al.*, 2018).

Dessa forma, para adaptar *K. phaffii* ao uso como uma plataforma de biologia sintética são necessárias ferramentas que permitam a rápida montagem de grandes e complexas sequências de DNA, controle específico da expressão gênica e facilidade na obtenção de clones geneticamente modificados.

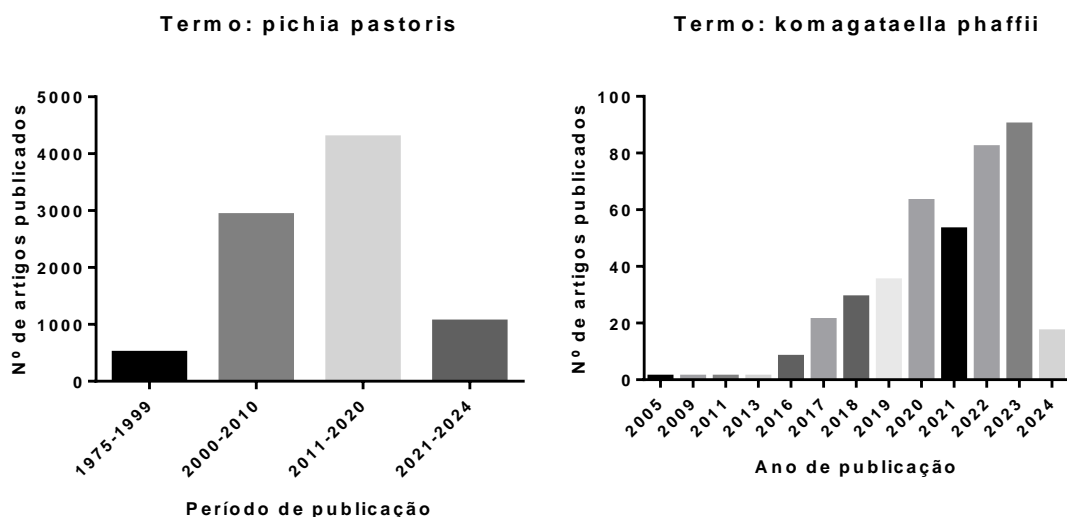
### **1.3 A levedura *Komagataella phaffii***

*Komagataella phaffii* é uma levedura metilotrófica utilizada comercialmente desde a década de 70, primeiramente servindo como fonte de proteína para ração animal. Anteriormente era chamada de *Pichia pastoris* até sua reclassificação ao gênero *Komagataella* em 1995. O genoma de três cepas de *K. phaffii* já foram sequenciados e dados sobre genes identificados e putativos estão disponíveis nas bases de dados do site [www.pichiagenome.org](http://www.pichiagenome.org) e no banco

de dados GenBank sob o número de acesso GCA\_000027005.1 (KÜBERL *et al.*, 2011; MATTANOVICH *et al.*, 2009; DE SCHUTTER *et al.*, 2009).

Desde então, com o avanço da tecnologia do DNA recombinante, esta levedura tem sido cada vez mais utilizada como plataforma eucariótica para a produção de proteínas heterólogas (CEREGHINO, 2000; LI *et al.*, 2007). Atualmente duas linhagens (X-33 e GS115) são comercializadas como plataformas de expressão de proteínas pela empresa Invitrogen (Madison, EUA). O crescimento a altas densidades celulares, a capacidade de realizar modificações pós traducionais similares à células de mamíferos e seu desenvolvido sistema de secreção de proteínas tornam esta levedura uma boa opção neste tipo de aplicação (AHMAD *et al.*, 2014). *K. phaffii* já foi utilizada como plataforma de expressão de diversas enzimas, peptídeos, proteínas bacterianas, proteínas virais e anticorpos humanos (ARAUJO *et al.*, 2015; BAGHBAN *et al.*, 2019; ÇELIK; ÇALIK, 2012; MACAULEY-PATRICK *et al.*, 2005).

Além de sua aplicação na produção de proteínas heterólogas, *K. phaffii* também tem sido utilizada na produção de outras moléculas de interesse comercial, como a riboflavina, S-adenosil-metionina, licopeno e  $\beta$ -caroteno (ARAYA-GARAY *et al.*, 2012; HE *et al.*, 2006; MARX; MATTANOVICH; SAUER, 2008). *K. phaffii* também possui potencial para a produção de metabólitos secundários graças ao forte fluxo na via das pentoses fosfato e ao seu metabolismo fortemente aeróbico, que ocorre mesmo em grandes concentrações de substrato pois não sofre influência do efeito Crabtree (ATA *et al.* 2021). Isso a favorece como plataforma de produção pois não diverge seu fluxo metabólico à produção de etanol como no caso da levedura *Saccharomyces cerevisiae* (POTVIN; AHMAD; ZHANG, 2012; SOLÀ *et al.*, 2004). Sendo assim, o número de estudos publicados com este microrganismo ganhou grande impulso desde a década de 1990 (Figura 1), fazendo com que se intensificasse o desenvolvimento de novas ferramentas para manipulação genética, técnicas de aumento da eficiência no cultivo e novas frentes comerciais utilizando *K. phaffii*.



**Figura 1: O aumento do interesse pelo uso comercial de *K. phaffii* gerou o crescimento do número de artigos científicos publicados. Período de pesquisa: 1975 a 2024. Os termos de pesquisa utilizados estão indicados acima dos respectivos gráficos. Fonte: PubMed (NCBI).**

#### **1.4 Ferramentas para a manipulação genética de *K. phaffii***

As técnicas de manipulação genética empregadas em *K. phaffii* são majoritariamente baseadas no uso de vetores integrativos, muito utilizados na expressão heteróloga de proteínas e que apresentam como vantagem a estabilidade das construções. O uso de vetores episomais, embora menos comum, apresenta maiores eficiências de transformação e permitem o manejo em *K. phaffii* de forma mais versátil (CAMATTARI *et al.*, 2016).

Os vetores disponíveis atualmente permitem o uso de marcas de seleção como a resistência aos antibióticos zeocina, higromicina e G418 ou a auxotrofia para diversos aminoácidos e bases nitrogenadas (DROCOURT *et al.*, 1990; LINCEREGHINO *et al.*, 2001, 2008; NETT *et al.*, 2005; YANG *et al.*, 2014). Novas marcas de seleção têm sido desenvolvidas para reduzir a dependência dos vetores existentes e buscar sistemas mais eficientes. Exemplos são o gene *FLD1*, que confere resistência da levedura ao formaldeído, e o gene da toxina *mazF* usado em contrasseleção (SUNGA; CREGG, 2004; YANG; JIANG; YANG, 2009). O gene *amdS* de *Aspergillus nidulans* também já foi empregado em *K. phaffii* e é um exemplo de marca de seleção dominante que não requer uso de



antibiótico. O gene confere à levedura a habilidade de usar acetamida como fonte de nitrogênio e ao mesmo tempo é responsável por um efeito de contrasseleção em meio de cultura contendo fluoroacetamida, onde linhagens que possuírem a marca *amdS* morrem na presença deste composto (PIVA *et al.*, 2018). Outro exemplo de marca de seleção que permite contrasseleção é o gene *URA5*, uma marca auxotrófica que permite a seleção em meio mínimo sem uracila. A contrasseleção é realizada com o ácido 5-fluorótico (5'-FOA), que é convertido em um composto tóxico quando a enzima codificada pelo gene *URA5* está presente (NETT; GERNGROSS, 2003).

Um sistema também adaptado de *S. cerevisiae* utiliza a via de síntese de adenina como marca de seleção. Linhagens com a deleção do gene *ADE2*, que codifica a enzima fosforibosil aminoimidazol carboxilase (EC 4.1.1.21), já são utilizadas comercialmente em *K. phaffii* devido a sua auxotrofia para adenina (PichiaPink Expression Kit, Invitrogen). Este gene catalisa a sexta etapa da biossíntese das purinas e causa o acúmulo de 5-fosforibosil 5-aminoimidazol (AIR), um composto avermelhado que leva a colônias com esta coloração (GEDVILAITE; SASNAUSKAS, 1994). Dessa forma, colônias com este gene deletado serão avermelhadas, e colônias que receberem o gene nativo como marca de seleção auxotrófica, além de recuperarem a prototrofia, voltarão a ser brancas. Este sistema pode ser utilizado na análise de estabilidade de plasmídeos e do número de cópias integradas na levedura (DU; BATTLES; NETT, 2012). Em *S. cerevisiae*, a deleção de genes de enzimas que atuam antes da fosforibosil aminoimidazol carboxilase na via de síntese de adenina, como o gene *ADE3*, evita a acumulação do composto vermelho e faz com que linhagens *ade2 ade3* sejam brancas (UGOLINI; CV, 1996). A deleção completa do gene *ADE3* em *S. cerevisiae* também causa auxotrofia para o aminoácido histidina, provavelmente devido ao acúmulo de intermediários com ação regulatória (RÉBORA; LALOO; DAIGNAN-FORNIER, 2005). Estirpes de *K. phaffii* com múltiplas auxotrofias estão disponíveis, como a GS200 (*arg4 his4*), Jc227 (*ade1 arg4*), Jc 300 (*ade1 arg4 his4*) e Jc308 (*ade1 arg4 his4 ura3*).

É notável que muitas das ferramentas de manipulação genética desta levedura foram adaptadas de *S. cerevisiae*. Inicialmente, os plasmídeos utilizados em *K. phaffii* carregavam as origens de replicação de plasmídeos desenvolvidos para *S. cerevisiae* e as estratégias de transformação e integração

genômica eram similares. No entanto, a levedura *S. cerevisiae* possui um sistema de reparo por recombinação homóloga ativo e já bem caracterizado na literatura, diferentemente de *K. phaffii* que utiliza majoritariamente o sistema de recombinação não homóloga (*non-homologous end-joining*), causando integrações não específicas e reduzindo a eficiência da transformação (SCHWARTZ *et al.*, 2017). Sendo assim, a utilização de estratégias de integração por recombinação homóloga requer extenso trabalho de seleção de transformantes e confirmação da integração no *loci* correto. Mesmo com estes percalços, novas técnicas têm permitido uma eficiente manipulação genética de levedura, como a montagem de vetores *in vivo* utilizando sequências homólogas maiores que em *S. cerevisiae* (CAMATTARI *et al.*, 2016; YU *et al.*, 2012). Um importante avanço veio com a recente deleção do gene *KU70*, que codifica para uma proteína envolvida no mecanismo de reparo de danos na dupla fita do DNA, permitindo reduzir a taxa de recombinação não específica (NÄÄTSAARI *et al.*, 2012; SCHWARTZ *et al.*, 2017; WENINGER *et al.*, 2018). Porém, resultados posteriores demonstraram que a deleção permanente de *KU70* traz consequências deletérias às linhagens, como instabilidade genômica e baixa taxa de crescimento (NÄÄTSAARI *et al.*, 2012). Alternativas como o sistema CRISPR-Cas9 estão sendo desenvolvidas para tentar aumentar a taxa de sucesso de integração direcionada de sequências heterólogas (TKACHENKO *et al.*, 2023).

Novas alternativas para utilização desta levedura incluem diversas estratégias para o aumento do nível de transcrição de genes, a construção de bibliotecas de promotores e de fatores de transcrição, até a inserção de vias inteiras em uma só etapa de transformação (HARTNER *et al.*, 2008; VOGL *et al.*, 2014). Neste contexto, destaca-se o uso da sequência 2A proveniente do Picornavirus (família *Picornaviridae*), que permite a utilização de cassetes policistrônicos em células eucarióticas. Esta estratégia consiste na inserção de uma sequência de 66 nucleotídeos entre as regiões codantes de 2 ou mais genes controlados por um único par promotor-terminador. Com isso, durante a tradução do mRNA policistrônico, a sequência viral de 22 peptídeos produzida passa por um processo de deslizamento do ribossomo (*ribosome skipping*), que resulta na quebra das proteínas e liberação dos produtos de cada gene específico (HOEFGEN *et al.*, 2018). Sendo assim, é possível fazer a inserção de múltiplos

genes em uma única etapa de transformação, eliminando a necessidade de utilização de múltiplos promotores, terminadores e marcas de seleção. Esta estratégia se mostra bastante interessante em situações em que se quer inserir ou modificar vias metabólicas que requerem diversas etapas de processamento. Além disso, atualmente a escassa quantidade de promotores e terminadores caracterizados para uso em *K. phaffii* restringe o número de modificações genéticas possíveis.

### **1.5 Controle de expressão gênica e promotores utilizados em *K. phaffii***

*K. phaffii* já foi utilizada para expressão de mais de 5.000 proteínas segundo dados disponíveis em <http://www.pichia.com> (XU *et al.*, 2018). A maioria dos sistemas de expressão utiliza o promotor  $P_{GAP}$  (gliceraldeído 3-fosfato desidrogenase) para ativação constitutiva da transcrição e o  $P_{AOX1}$  (álcool oxidase I) induzível por metanol (VOGL; GLIEDER, 2013). Estas sequências são conhecidas pelo seu restrito controle da transcrição e por sua forte indução a partir da sinalização por carboidratos e metanol, respectivamente. Outras sequências induzíveis foram reportadas como o  $P_{FLD1}$ , promotor do gene codificador da enzima formaldeído desidrogenase dependente de glutatona, dihidroxiacetona sintase ( $P_{DAS}$ ), os promotores ativados em baixa concentração de glicose ( $P_{G1}$ ) e reprimidos por glicerol ( $P_{G6}$ ) (PRIELHOFER *et al.*, 2013; SHEN *et al.*, 1998; TSCHOPP *et al.*, 1987). Além dos promotores induzíveis, outros promotores constitutivos já foram testados, sendo eles o promotor do fator de alongamento da tradução ( $P_{TEF1}$ ) (STADLMAYR *et al.*, 2010) e o promotor da GTPase ( $P_{YPT1}$ ) (SEARS *et al.*, 1998). No entanto, estes promotores não são amplamente utilizados devido à problemas relacionados à força de expressão, condições para indução e o rigor na regulação da transcrição. Por exemplo, o  $P_{FLD1}$  necessita não somente de metanol como única fonte de carbono para sua indução, mas também de sulfato de amônio –  $(NH_4)_2SO_4$ – como fonte de nitrogênio. Já no caso de  $P_{G1}$  e  $P_{G6}$ , assim como em outros promotores, a força do promotor não suporta os altos níveis de expressão requeridos para a produção em larga escala de proteínas heteróloga (AHMAD *et al.*, 2014; SHEN *et al.*, 1998).

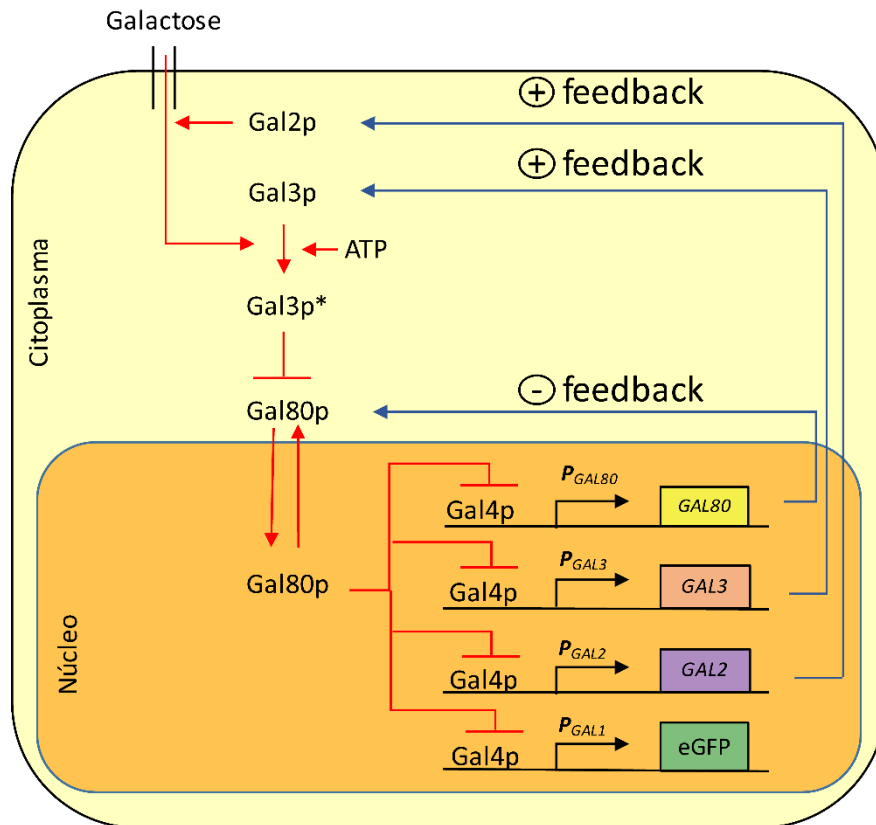
A utilização do promotor correto para as condições de expressão da proteína recombinante é fundamental, pois um promotor fraco irá produzir poucos transcritos e por consequência o rendimento proteico será reduzido. Por outro lado, um promotor excessivamente forte pode produzir uma sobrecarga metabólica na célula, diminuindo a taxa de crescimento celular, aumentando o acúmulo de proteínas mal dobradas e a produção de proteínas (GOMES *et al.*, 2018). De forma a dirimir este problema, considera-se que a utilização de um promotor constitutivo ou de um promotor induzível, permite o controle da quantidade de proteínas em determinadas condições. No entanto, nos casos em que se deseja inserir múltiplos genes de uma via metabólica na levedura, a utilização de múltiplos promotores e terminadores se torna tecnicamente desafiador. A proporcionalidade de cada proteína de uma via metabólica é cuidadosamente controlada por uma célula e replicar este mecanismo transcricional complexo com precisão ainda não se faz viável. A fim de contornar este problema, tentativas de utilização de metodologias de embaralhamento aleatório de cassetes monocistrônicos se mostraram pouco eficientes em muitos casos (GEIER *et al.*, 2015; KIM *et al.*, 2013; SHAO; ZHAO; ZHAO, 2009). Estudos mais aprofundados de caracterização de novas sequências promotoras e da utilização de cassetes de expressão policistrônicos se fazem necessários para permitir a utilização de grandes sequências de DNA recombinante em *K. phaffii* (GEIER *et al.*, 2015b; HOEFGEN *et al.*, 2018). Sendo assim, sistemas de regulação da expressão gênica de *S. cerevisiae* podem ser testados em *K. phaffii*, que permitiria a ampliação do número de ferramentas disponíveis para esta levedura.

### **1.6 O sistema de expressão gênica regulado por galactose em *S. cerevisiae***

Uma das mais importantes estratégias utilizadas para produção heteróloga de proteínas em *S. cerevisiae* é baseada no seu sistema de expressão regulado por galactose. A família de genes *GAL* em *S. cerevisiae* consiste em três genes regulatórios (*GAL4*, *GAL 80* e *GAL3*) e cinco genes da via metabólica (*GAL1*, *GAL2*, *GAL7*, *GAL10* e *MEL1*), que em conjunto permitem esta levedura o uso de galactose como fonte única de carbono (LOHR *et al.*, 2016). O gene regulatório *GAL4* codifica um fator de transcrição chamado Gal4p,

que se liga à sequência de ativação da transcrição (em inglês, UAS) dos genes GAL como um homodímero e inicia o recrutamento do complexo mediador da RNA polimerase II, iniciando a transcrição (KANG; MARTINS; SADOWSKI, 1993; SELLECK; MAJORS, 1987; TRAVEN; JELICIC; SOPTA, 2006). Os genes da família GAL podem apresentar múltiplos sítios de ligação para Gal4p, permitindo um controle fino da expressão destes genes. Já o gene GAL80 codifica para uma proteína repressora, Gal80p, que se auto associa para a formação de um homodímero, ganhando a capacidade de ligar-se a Gal4p e impedindo sua ligação ao DNA, consequentemente prevenindo o recrutamento do complexo, levando à impossibilidade de transcrição dos genes GAL (PANNALA *et al.*, 2010). Gal3p é um transdutor de sinal que interage com Gal80p em resposta a galactose. Sua atividade é dependente de ATP e galactose, para que consiga se ligar a Gal80p (PENG; HOPPER, 2000). Já a proteína Gal2p é uma permease de membrana que é responsável pelo transporte de galactose para o citoplasma (KRUCKEBERG, 1996; SELICK; CAMPBELL; REECE, 2008). O sistema está demonstrado na **Figura 2**.

Ao contrário de *S. cerevisiae*, *K. phaffii* não metaboliza este carboidrato (Jiang *et al.*, 2015), pois não possui a maquinaria enzimática necessária para tal. Um sistema de expressão induzido por galactose representaria uma alternativa mais segura que o sistema induzível por metanol. Além disso, também é uma opção mais econômica, uma vez que as quantidades necessárias para a indução seriam bastante reduzidas, já que a levedura não seria capaz de degradar o indutor. No entanto, para que este sistema seja utilizado em *K. phaffii*, devem ser incorporados ao genoma desta levedura os quatro genes do regulon GAL de *S. cerevisiae*. Porém, em *S. cerevisiae*, cada um destes genes possui promotores específicos com forças de expressão diferenciadas entre si. Desta forma, integrar múltiplos genes em *K. phaffii* representa um grande desafio, uma vez que são necessários vários *loci* para a inserção dos genes e o controle fino e balanceado da expressão de cada gene. Assim, cria-se um grande desafio de engenharia metabólica e a utilização do cassete policistrônico se mostra a estratégia mais adequada para esta situação, pois reduzirá a complexidade total da construção.



**Figura 2: Representação gráfica do circuito genético ativado por galactose.** A galactose é capaz de ativar a transcrição de genes controlados por promotores da família GAL. A presença de galactose leva à fosforilação de Gal3p que sequestra Gal80p e impede sua migração ao núcleo. A ausência de Gal80p no núcleo permite a ativação da transcrição dos genes GAL por Gal4p. Para provar que o circuito genético é funcional em *K. phaffii*, a proteína eGFP foi clonada sob controle do promotor P<sub>GAL1</sub> para que seja transcrita na presença de galactose.

### 1.7 Aptâmeros como sequências controladoras de expressão gênica e proteica

Aptâmeros são sequências sintéticas de DNA que possuem a capacidade de se dobrar em uma estrutura tridimensional e ligar-se não covalentemente com alta especificidade e afinidade a uma molécula alvo (DEROSA, 2016). Estas sequências podem ser sintetizadas aleatoriamente *in vitro* e posteriormente selecionadas de acordo com sua afinidade pela molécula alvo ou podem ser desenhadas com o auxílio de softwares de predição de estrutura tridimensional

(TUERK, 1990). A técnica mais utilizada ganhou o acrônimo SELEX (do inglês *Systematic Evolution of Ligands by EXponential enrichment*) e consiste em três etapas: i) criação das bibliotecas de oligonucleotídeos que podem conter até  $10^{16}$  sequências candidatas. Estas sequências possuem uma região aleatória flanqueada por duas regiões conhecidas. As regiões aleatórias serão testadas por sua afinidade ao alvo. Estes oligonucleotídeos podem ser marcados por fluoróforos, radioisótopos ou outra molécula que permita sua recuperação; ii) a fase de recuperação das sequências que interagiram com a molécula alvo e; iii) amplificação por PCR das sequências recuperadas permitindo a criação de uma grande massa necessária para as etapas subsequentes de teste e caracterização (DEROSA, 2016). Estas três etapas podem ser repetidas extensivamente, com modificação das condições de incubação entre alvo e oligonucleotídeos, a fim de se obter um grau maior de rigor na seleção dos aptâmeros.

Uma vez feita a seleção dos aptâmeros, estas pequenas sequências podem ser rapidamente aplicadas para ligação com diversos tipos de ligantes, como por exemplo: toxinas, antibióticos, marcadores moleculares, drogas, metais pesados, detecção de estirpes bacterianas dentre outros (MCKEAGUE, 2015). A vantagem desta metodologia em relação a outras mais amplamente utilizadas, como as tecnologias baseadas em interações com anticorpos, está no seu baixo custo e mais fácil aplicabilidade. Os aptâmeros não necessitam de cultivo de células ou insumos biológicos de alto custo para sua produção e podem ser produzidos *in vitro* em larga escala, com mínimas variações entre lotes. Além disso, a vida útil dos aptâmeros é maior que a de anticorpos e são mais estáveis mesmo após exposição a condições impróprias de armazenamento, pois possuem capacidade de renaturação mesmo após altas temperaturas ou desidratação (SMITH, 2007; ROHLOFF, 2014).

Os aptâmetros também podem ser utilizados para a construção de *riboswitches*, que são sequências inseridas na região 5'-UTR de mRNAs que exercem função regulatória sobre a transcrição ou tradução dos alvos. Este mecanismo ocorre através da interação reversível do ligante, geralmente uma pequena molécula, que causa uma mudança da estrutura tridimensional do RNA, impedindo que as proteínas do maquinário celular reconheçam e processem corretamente a sequência (Figura 3).

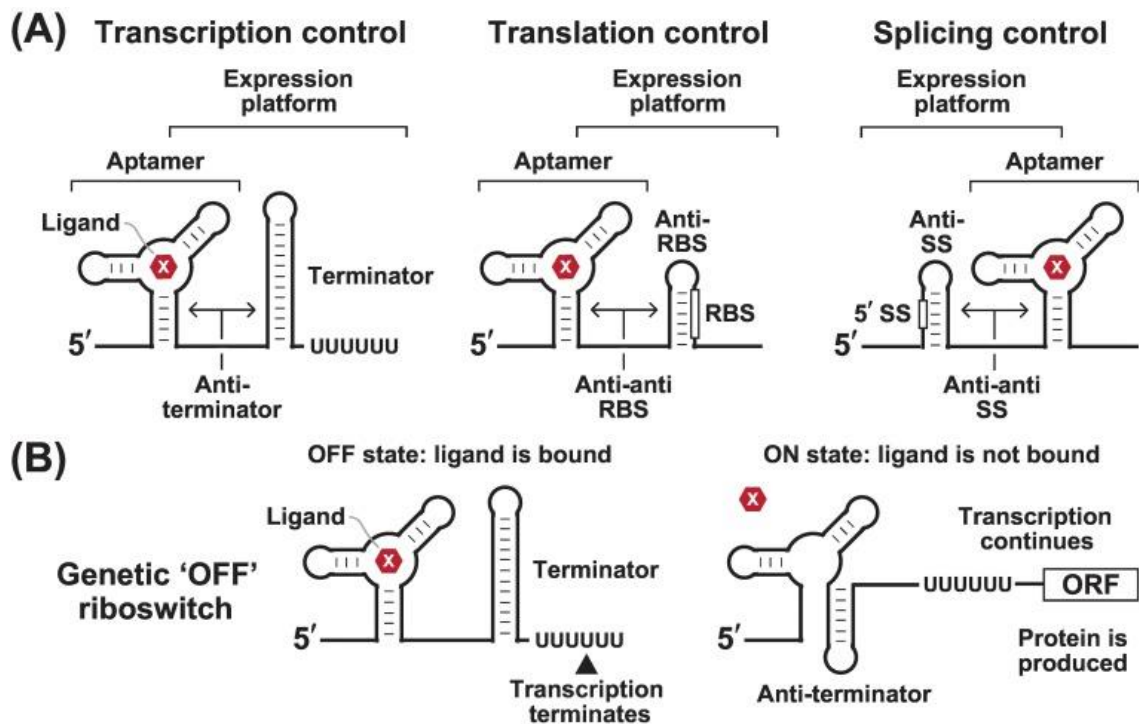


Figura 3: Mecanismo de ação de um *riboswitch*. A) Os aptâmeros são inseridos no mRNA para criar uma sequência que serve de sítio de ligação para uma molécula. Este sistema pode ser utilizado para ativar a transcrição (esquerda), tradução (centro) ou *splicing* (direita). B) O *riboswitch* também pode ter função inibitória parcial ou total, caso a sequência seja inserida em uma região controladora do mRNA, como sítios de ligação de ribossomos. Retirado de: Kavita & Breaker, 2023. Doi: 10.1016/j.tibs.2022.08.009.

Sendo assim, os aptâmeros estão ganhando cada vez mais atenção da comunidade científica por sua flexibilidade de aplicação em diversos ramos. Atualmente há uma escassa disponibilidade de vetores e linhagens comerciais para *K. phaffii*, apesar de sua relevante importância biotecnológica e, portanto, é altamente desejável que se faça o melhoramento genético e criação de novas linhagens desta levedura, visando a otimização da produção de proteínas recombinantes e outras moléculas de grande interesse econômico.

O atual sistema de expressão heteróloga de proteína é baseado majoritariamente no promotor  $P_{AOX1}$ , que como dito anteriormente, apresenta algumas desvantagens. Neste trabalho, tratamos do desenvolvimento de duas novas linhagens de *K. phaffii* onde uma possui um sistema regulável de



expressão gênica baseado no regulon GAL de *S. cerevisiae* e a outra possui um sistema de controle da expressão proteica baseado em um *riboswitch* construído por aptâmeros sensíveis à tetraciclina.

## 2. Objetivos

a) Desenvolver uma linhagem de *K. phaffii* com um sistema de controle de expressão gênica baseado no regulon GAL de *S. cerevisiae*.

Objetivos específicos:

- i. Construir vetor contendo gene repórter eGFP (*enhanced green fluorescent protein*) sob controle do promotor  $P_{GAL1}$ ;
- ii. Amplificar as regiões codificantes do regulon GAL de *S. cerevisiae* (*GAL2*, *GAL3*, *GAL4*, *GAL80*);
- iii. Montar os genes do regulon GAL em um cassete de expressão policistrônico sob controle do promotor  $P_{TEF1}$  de *S. cerevisiae* e contendo a marca de seleção dominante G418<sup>R</sup>;
- iv. Integrar o cassete do regulon GAL no *locus* do gene *AOX1*;
- v. Testar os fenótipos dos clones obtidos;
- vi. Detectar atividade de eGFP em meio contendo galactose e glicerol;
- vii. Avaliar a concentração ótima de galactose para indução do promotor  $P_{GAL1}$ ;

b) Desenvolver uma linhagem de *K. phaffii* com um sistema de controle de expressão proteica baseado no *riboswitch* sensível à tetraciclina.

Objetivos específicos:

- i. Clonar o gene *KU70* da linhagem GS115 de *K. phaffii* em um vetor epissomal com suas sequências promotoras e terminadoras;
- ii. Clonar o gene *lacZ* em um vetor epissomal para utilização como gene repórter;
- iii. Inserir os aptâmeros nas regiões promotoras dos genes *KU70* e *lacZ* para testar controle da expressão das proteínas;
- iv. Criar um mecanismo de detecção por western blotting para a proteína Ku70p;

- v. Avaliar condições de inibição da expressão proteica com exposição a tetraciclina com as proteínas Ku70p e a enzima  $\beta$ -galactosidase;

## 2. Estratégia

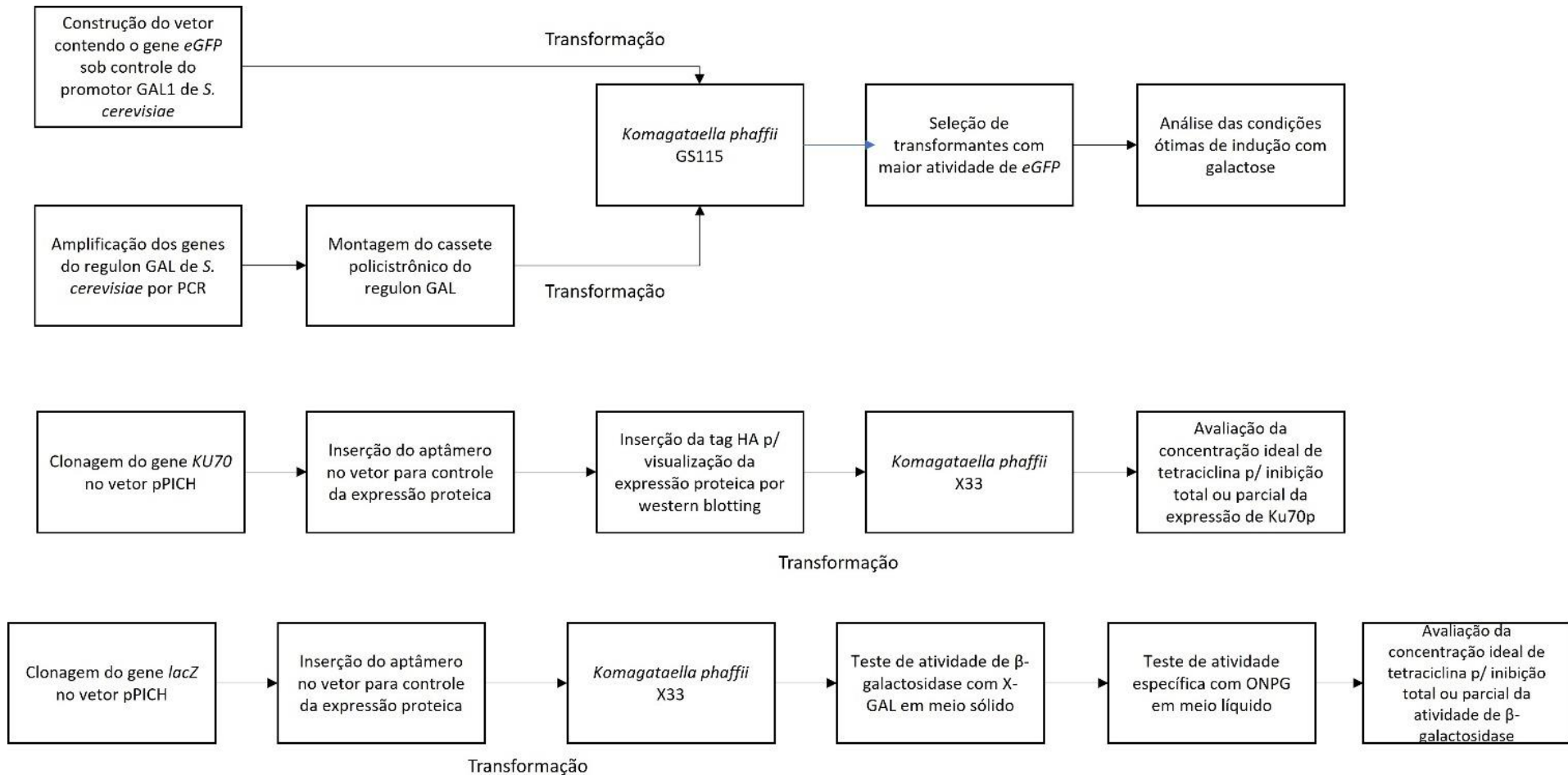


Figura 4: Fluxograma de trabalho para montagem das linhagens de controle da expressão gênica e proteica em *K. phaffii*.

### 3. Materiais e Métodos

#### 4.1. Células

##### Bactérias:

- *Escherichia coli* XL10-gold (Agilent Technologies, EUA): Tet<sup>r</sup>  $\Delta(mcrA)183$ ,  $\Delta(mcrCB-hsdSMR-mrr)173$ , *endA1*, *supE44*, *thi-1*, *recA1*, *gyrA96*, *relA1*, *lac*, *Hte* [*F*<sup>+</sup>*proAB lacI*<sup>q</sup>*Z* $\Delta$ *M15 Tn10* (Tet<sup>r</sup>) *Amy Cam*<sup>r</sup>]
- DH10-alpha (Thermo Fischer Scientific, EUA): F<sup>-</sup>, *mcrA*,  $\Delta(mrr-hsdRMS-mcrBC)$ ,  $\Phi80d$ *lacZ* $\Delta$ *M15*,  $\Delta$ *lacX74*, *endA1*, *recA1*, *deoR*,  $\Delta(ara, leu)7697$ , *araD139*, *galU*, *galK*, *nupG*, *rp sL*,  $\lambda$ <sup>-</sup>

##### Leveduras:

- *Komagataella phaffii* GS115 (ATCC20864)
- *Komagataella phaffii* X-33  $\Delta ku70::amdS$

#### 4.2. Meios de cultivo

##### LB

Extrato de levedura	0,5% (p/v)
Triptona	1% (p/v)
NaCl	1% (p/v)

O pH foi ajustado para 7,2. Para o meio sólido, foi adicionado ágar 1,5% (p/v). O meio foi esterilizado por calor úmido em autoclave. Quando necessário, foram utilizados os antibióticos zeocina a uma concentração final de 25 µg/mL, ampicilina 100 µg/mL ou kanamicina 100 µg/mL.

##### SOB

Extrato de levedura	0,5% (p/v)
Triptona	2% (p/v)
NaCl	0,5% (p/v)
KCl	0,2% (p/v)
MgCl <sub>2</sub>	10 mM
MgSO <sub>4</sub>	10 mM

O pH foi ajustado para 7,5. Soluções de  $MgCl_2$  e  $MgSO_4$  na concentração de 1M foram filtradas com membranas de poro 0,22  $\mu m$  para esterilização e adicionadas ao meio esterilizado por calor úmido em autoclave para a concentração final de 10mM.

## YPD

Extrato de levedura	1% (p/v)
Peptona	2% (p/v)
Glicose	2% (p/v)

Para o meio sólido, foi adicionado ágar 2% (p/v). Quando necessário, foram utilizados os antibióticos Zeocina a uma concentração final de 100  $\mu g/mL$  e G418 200  $\mu g/mL$ .

## MEIO MÍNIMO (MM)

YNB com Sulfato de Amônio	1,34% (p/v)
Biotina	$4 \times 10^{-5}\%$ (p/v)

Para o preparo, foram feitas soluções estoque de cada reagente e a mistura foi feita em ambiente asséptico. O meio mínimo foi suplementado com sorbitol ou galactose na concentração indicada. As soluções estoque utilizadas foram esterilizadas por filtração com filtro de nitrocelulose 0,22  $\mu m$  (Millipore).

## 4.3 Soluções

### 4.3.1 Soluções para eletroforese em gel de poliacrilamida

Acrilamida:bisacrilamida: 29% acrilamida; 1% bisacrilamida.

1,5M Tris-HCl pH 8,8

1M Tris-HCl pH 6,8

10% sódio dodecilsulfato

10% persulfato de amônio (armazenada a  $-20\text{ }^\circ C$  até o momento de uso.)

Tampão de amostra desnaturante 5X: Tris-HCl 60mM (pH 6,8)

SDS 2% (p/v)

$\beta$ -Mercaptoetanol 14,4 mM

Glicerol 25% (v/v)

Azul de bromofenol 0,1% (p/v)

Tampão de corrida: Tris-Glicina 10X (250 mM Tris-HCl ; 1,9 M)

Solução de lise: 50 mM Tris-HCl pH 8,0

Tween 20 1,25% (v/v)

0,04% (v/v) Complete Protease Inhibitor Cocktail  
(Roche)

Solução corante: Comassie brilliant blue G-250 0,2% (p/v)

Metanol 30% (v/v)

Ácido acético glacial 7% (v/v)

Solução descorante: Metanol 30% (v/v)

Ácido acético glacial 7% (v/v)

#### **4.3.2 Western Blotting**

Tampão de transferência: 25 mM Tris-HCl

190 mM Glicina

SDS 0,1 % (p/v)

Metanol 20 % (v/v)

Tampão de incubação (TBST): 20 mM Tris-HCl pH 7,5

150 mM NaCl

Tween 20 0,1 % (v/v)

Solução de bloqueio de membrana: Leite em pó integral reconstituído em  
tampão TBS-T 5% (p/v)

#### **4.3.3 Preparo de DNA plasmidial por lise alcalina**

Solução I: Tris-HCl 25mM (pH 8,0)

EDTA 10 mM (pH 8,0)

Solução II: 200mM NaOH

SDS 1% (p/v)

Solução preparada no momento de uso.

Solução III: Acetato de Sódio 3M

Ácido acético 2M

O pH foi ajustado para 5,0.

#### 4.3.4 Eletroforese em gel de agarose

Brometo de etídio: EtBr 10 mg/mL

Tampão Tris-acetato-EDTA (TAE) 50X: Tris 2M

Ácido acético glacial 1M

EDTA 0,05 M (pH 8,0)

Tampão de amostra: TAE 50X 50% (v/v)

Glicerol 30% (v/v)

Azul de bromofenol 0,25% (p/v)

#### 4.3.5 Antibióticos

- Zeocina 100 mg/mL (Thermo Fischer)
- Kanamicina (50mg/mL): Dissolvida em água destilada e esterilizada por filtração.
- Tetraciclina (1mM): Dissolvida em etanol 99,5% (v/v).

As soluções de antibióticos foram mantidas em alíquotas a -20 °C até o momento de utilização.

#### 4.4 Preparo de células bacterianas competentes por choque térmico

As células de *E. coli* competentes foram preparadas para transformação por choque térmico. XL10-gold ou DH10-alpha foram semeadas em meio LB ágar a partir de um estoque armazenado a - 80 °C e incubadas por 16 h a 37 °C.

Uma colônia isolada da bactéria foi inoculada em 10 mL de meio SOB em um frasco Erlenmeyer de 125 mL e incubada a 200 rpm e 37 °C por 16 h. A leitura da DO<sub>600</sub> foi feita e o volume de células necessário para um inóculo de 0,05 DO<sub>600</sub> em 100 mL de meio SOB foi feito e incubado por tempo suficiente para que a DO<sub>600</sub> chegasse a 0,3. A cultura foi então resfriada por 15 min em banho de água-gelo e as células foram centrifugadas a 3000 x g por 5 min a 4 °C. O precipitado foi ressuspenso em 32 mL de tampão de transformação I

(RbCl 12 g/L,  $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$  9,9 g/L, acetato de potássio 0,03 M,  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$  1,5 g/L e glicerol 150 g/L) e incubado em gelo por 15 min. As células foram então centrifugadas novamente e ressuspensas em 4 mL de tampão de transformação II (MOPS 0,02 M, RbCl 1,2 g/L,  $\text{CaCl}_2$  11 g/L e glicerol 150 g/L). Ao final do protocolo, as células foram separadas em alíquotas de 100  $\mu\text{L}$  e congeladas em banho de gelo seco e etanol. O armazenamento foi feito a  $-80$   $^\circ\text{C}$ .

O procedimento de transformação utilizou uma alíquota de células competentes descongeladas em gelo por 30 min, então submetidas a  $42$   $^\circ\text{C}$  por 90 segundos e novamente a gelo por 2 min. As células receberam 1 mL de meio LB, foram recuperadas por 1 h a  $37$   $^\circ\text{C}$  e semeadas em meio seletivo

#### **4.5 Preparo de células bacterianas competentes por eletroporação**

Uma colônia isolada da bactéria foi incubada por 16 h a  $37$   $^\circ\text{C}$  e 200 rpm em 5 mL de meio LB com 50  $\mu\text{g}/\text{mL}$  de tetraciclina. Um mililitro deste pré-inóculo foi adicionado a 100 mL de meio LB sem antibiótico num frasco Erlenmeyer de 500 mL. A cultura foi incubada a  $37$   $^\circ\text{C}$  e 200 rpm até que a  $\text{OD}_{600}$  chegasse a 0,5. O frasco foi então resfriado em gelo por 20 min e as células foram centrifugadas a  $1200 \times g$  por 5 min a  $4$   $^\circ\text{C}$ . Após a remoção de todo o sobrenadante, as células foram ressuspensas em 10 mL de glicerol 10% gelado. Em seguida, foi feita uma incubação em gelo por 20 min e as células foram novamente coletadas por centrifugação nas mesmas condições. O precipitado foi ressuspendido em 100  $\mu\text{L}$  de glicerol 10% gelado e as células foram separadas em alíquotas de 40  $\mu\text{L}$ , congeladas em gelo seco e armazenadas a  $-80$   $^\circ\text{C}$ .



O procedimento de transformação por eletroporação utilizou os seguintes parâmetros numa cubeta de 0,1 cm: 1,8 kV, 10  $\mu$ F e 600  $\Omega$ . Após o choque, as células receberam 1 mL de meio LB, foram cultivadas por 1 h a 37 °C e foram semeadas em meio seletivo.

#### **4.6 Extração de plasmídeos de *E. coli* por lise alcalina**

A extração dos plasmídeos de *E. coli* foi realizada pelo método de lise alcalina descrito por Sambrook e colaboradores, em 1989, com modificações. Uma colônia isolada da placa de transformantes foi inoculada em 5 mL de meio LB com o antibiótico adequado e este inóculo foi incubado a 37 °C por 16 h sob agitação. As células de 3 mL de cultura foram coletadas por centrifugação a 10.000 x *g* por 2 minutos e o precipitado foi resuspendido em 200  $\mu$ L de solução I. Foram adicionados 360  $\mu$ L de solução II para provocar a lise das células e a amostra foi incubada a temperatura ambiente por 5 minutos. Posteriormente foram adicionados 300  $\mu$ L de solução III gelada e a mistura foi incubada por 5 minutos em gelo. A amostra foi coletada centrifugação a 10.000 x *g* por 5 minutos e o sobrenadante foi transferido para outro tubo. Foi adicionado 750  $\mu$ L de isopropanol, homogeneizado por inversão e submetido à centrifugação a 10.000 x *g* por 5 minutos. O sobrenadante foi descartado e o precipitado resuspendido em 200  $\mu$ L de solução I. Foram adicionados 110  $\mu$ L de acetato de amônio 7,5 M e a amostra foi homogeneizada com agitação vigorosa. Em seguida, a amostra foi submetida a centrifugação a 10000 x *g* por 10 minutos. O sobrenadante foi transferido para um novo tubo e precipitado através da adição de 750  $\mu$ L de 99,5% (v/v) gelado. Seguiu-se com uma centrifugação a 10000 x *g* por 5 minutos e foi feito o descarte do sobrenadante. Uma lavagem com 500  $\mu$ L de etanol 70%

(v/v) gelado foi realizada seguida de centrifugação a 10000 x *g* por 5 minutos. O sobrenadante foi descartado e o precipitado foi seco com incubação a 45 °C. Após secagem, o precipitado foi ressuspensão em 50 µL de água destilada ultrapura contendo RNase A (20 µg/mL) e incubado a 37 °C por 15 minutos. O DNA foi armazenado a -20 °C até quando necessário.

Também foram utilizados os *kits* de extração de plasmídeos Wizard SV Plus (Promega) e Qiagen Plasmid Maxi (Qiagen).

#### **4.7 Análise de sequências de DNA por homologia**

O alinhamento de sequências de DNA foi feito utilizando a ferramenta BLASTn ([blast.ncbi.nlm.nih.gov/Blast.cgi](http://blast.ncbi.nlm.nih.gov/Blast.cgi)). As análises necessárias para design dos oligonucleotídeos utilizados neste trabalho foram feitas utilizando os softwares SnapGene e Tm Calculator (Thermo Fischer).

#### **4.8 Amplificação por PCR**

A amplificação de fragmentos por reação em cadeia da polimerase (PCR) foi realizada com diferentes enzimas, dependendo da aplicação. Foram utilizadas a Phusion Dna Polymerase (Thermo Fischer Scientific ou New England Biolabs) e a Taq DNA Polymerase (Thermo Fischer Scientific, EUA) de acordo com as instruções dos fabricantes. Os oligonucleotídeos utilizados neste trabalho estão descritos na **Tabela 1**.

**Tabela 1: Oligonucleotídeos utilizados neste trabalho para amplificação por PCR de fragmentos de DNA. Sequências em negrito representam regiões de recombinação homóloga. Sequências em itálico representam o motivo 2A.**

ID	Nome	Sequência	Características
[1]	PGAL1-GFP-F	<b>CCGCCCTAGGCAATTAATATTTACTTATTTCCG</b> GATTAGAAGCCGCCGAGC	Amplifica o cassete contendo promotor P <sub>GAL1</sub> , gene GFP e o terminador CYC1 com regiões para clonagem por recombinação homóloga no vetor pPICH
[2]	PGAL1-GFP-R	<b>GAATTCGGATCCGTTTAAACTCATTCAACCG</b> CAAATTAAGCCTTCGAGC	
[3]	Regal-TEF1-F	ATGCATGCGCGCCGCGATCTAACATCCAAA <b>GACGAAAGGTTGAATGAAACCTTTTTGCCAT</b> <b>CCGACATCCACAGGTC</b> AATGTTTCTACTCCTT TTTTACTCTTCCAG	Amplifica o promotor TEF1 do genoma de <i>S. cerevisiae</i> S288c. A extremidade 5' possui região homóloga ao locus 5' UTR do gene <i>AOX1</i> . A extremidade 3' possui região homóloga ao gene <i>GAL2</i>
[4]	RegalTEF1-R	<b>TTGCTGTGAAACAACAGGCATATTGTTCTCCT</b> <b>CAACTGCCAT</b> CCTTAGATTAGATTGCTATGCTT TCTTTCTAATG	
[5]	Regal-GAL2-F	<b>ATTACAACTTTTTTTACTTCTTGCTCATTAGAA</b> <b>AGAAAGCATAGCAATCTAATCTAAGATGGCA</b> GTTGAGGAGAACAATATGC	Amplifica o gene <i>GAL2</i> do genoma de <i>S. cerevisiae</i> S288c. A extremidade 5' possui região homóloga a extremidade 3' do P <sub>TEF1</sub> . A extremidade 3' possui região homóloga a extremidade 5' do gene <i>GAL3</i> .
[6]	Regal-GAL2-R	TATTGGAACGTTTGTATTCATGGGCCCTGGAT TAGATTCAACATCACCAGCCAATTTCAACAAAT CAAAGTTCAAAGTTTATTCTAGCATGGCCTTG TACCAC	
[7]	Regal-2AGAL3-F	<b>TTACAACATGACGACAAACCGTGGTACAAGG</b> <b>CCATGCTAGAATAAACTTTGAACTTTGATTTG</b> <i>TTGAAATTGGCTGGTGATGTTGAATCTAATCC</i>	Amplifica o gene <i>GAL3</i> do genoma de <i>S. cerevisiae</i> S288c. A extremidade 5' possui região homóloga a

		AGGGCCCATGAATACAAACGTTCCAATATTCA G TTC	extremidade 3' do gene <i>GAL2</i> adicionado do fragmento 2A. A extremidade 3' possui região homóloga a extremidade 5' do gene <i>GAL4</i> .
[8]	RegalGAL3-R	GCACGCTTGTTGATAGAGAAGACAGTAGCTTCA TGGGCCCTGGATTAGATTCAACATCACCAGCC AATTTCAACAAATCAAAGTTCAAAGTTTATTGT TCGTACAAACAAGTACCCAAG	
[9]	Regal-2A GAL4-F	<b>TTGTACGAACAATAAACTTTGAACTTTGATTT</b> GTTGAAATTGGCTGGTGATGTTGAATCTAATC CAGGGCCCATGAAGCTACTGTCTTCTATCGAA C	Amplifica o gene <i>GAL4</i> do genoma de <i>S. cerevisiae</i> S288c. A extremidade 5' possui região homóloga a extremidade 3' do gene <i>GAL3</i> adicionado do fragmento 2A. A extremidade 3' possui região homóloga a extremidade 5' do gene <i>GAL80</i> .
[10]	RegalGAL4-R	AGATCTCTTGTTGTAGTCCATGGGCCCTGGAT TAGATTCAACATCACCAGCCAATTTCAACAAAT CAAAGTTCAAAGTTTACTCTTTTTTTGGGTTT GTGGG	
[11]	Regal-2A GAL80-F	<b>CCAAAAAAGAGTAAACTTTGAACTTTGATTT</b> GTTGAAATTGGCTGGTGATGTTGAATCTAATC CAGGGCCCATGGACTACAACAAGAGATCTTC GG	Amplifica o gene <i>GAL80</i> do genoma de <i>S. cerevisiae</i> S288c. A extremidade 5' possui região homóloga a extremidade 3' do gene <i>GAL4</i> adicionado do fragmento 2A. A extremidade 3' possui região homóloga a extremidade 5' do gene <i>G418<sup>R</sup></i> .
[12]	RegalGAL80- R	TCCCGTTGAATATGGCCCCATGGGCCCTGGA TTAGATTCAACATCACCAGCCAATTTCAACAAA TCAAAGTTCAAAGTTTATAAACTATAATGCGAG ATATTGCTAACGTTTAATG	
[13]	Regal-2A G418-F	<b>TCTCGCATTATAGTTTATAAACTTTGAACTTTG</b> ATTTGTTGAAATTGGCTGGTGATGTTGAATCT AATCCAGGGCCCATGGGGCCATATTCAACGG GAAACGTC	Amplifica o gene <i>G418<sup>R</sup></i> junto do terminador <i>CYC1</i> a partir do vetor pPICK2. A extremidade 5' possui região homóloga a extremidade 3' do gene <i>GAL80</i>

[14]	Regal-G418-CYC1t-R	ATGCATGCGAATTCATAGGTTACAAATAAAAA AGTATCAAAAATGAAGCCTGCATCTCTCAGGC AAATGGCATTCTGACATCCTCTTGAAGCTTGC AAATTAAGCCTTCGAGC	adicionado do fragmento 2A. A extremidade 3' possui região homóloga ao <i>locus</i> 3'UTR do gene <i>AOX1</i> .	
[15]	ampregal-TEF1-F	ATGCATGCGCGGC	Esta série de primers foi utilizada para o alongamento das extremidades dos fragmentos amplificados pelos oligonucleotídeos da série "Regal" (IDs [1] a [14]). Os alvos e sentido de amplificação estão identificados pelo nome do oligonucleotídeo.	
[16]	ampregal-TEF1-R	TTGCTGTGAAACAACAGGC		
[17]	ampregal-GAL2-F	ATTACAACTTTTTTTACTTCTTGCT		
[18]	ampregal-GAL2-R	TATTGGAACGTTTGTATTCA		
[19]	ampregal-GAL3-F	TTACAACATGACGACAAACC		
[20]	ampregal-GAL3-R	GCACGCTTGTTTCGATAG		
[21]	ampregal-GAL4-F	TTGTACGAACAATAAACTTTGAAC		
[22]	ampregal-GAL4-R	AGATCTCTTGTGTAGTCCATG		
[23]	ampregal-GAL80-F	CCAAAAAAGAGTAACTTTGAACTTTGA		
[24]	ampregal-GAL80-R	TCCCGTTGAATATGG		
[25]	ampregal-G418-F	TCTCGCATTATAGTTTA		
[26]	ampregal-G418-R	ATGCATGCGAATTCATAGGTT		
[27]	newGAL2-R	TTATTCTAGCATGGCCTTGTACCAC		Anela no terminal 3' do fragmento amplificado pela reação de PCR com o oligonucleotídeo [6].
[28]	newGAL3-R	TTATTGTTTCGTACAAACAAGTA		Anela no terminal 3' do fragmento amplificado pela reação de PCR com o oligonucleotídeo [8].

[29]	newGAL4-R	TTACTCTTTTTTTGGGTTTGGTGGGGTAT	Anela no terminal 3' do fragmento amplificado pela reação de PCR com o oligonucleotídeo [10].	
[30]	newGAL80-R	TTATAAACTATAATGCGAGATATTG	Anela no terminal 3' do fragmento amplificado pela reação de PCR com o oligonucleotídeo [12].	
[31]	mid-3'TEF1-F	AAAGAAAGCATAGCAATCTAATCTAAG	A série "mid" de oligonucleotídeos foi utilizada para confirmação da montagem (ordem e sentido) dos genes no cassete policistrônico. Os alvos e o sentido de amplificação estão descritos no nome do oligonucleotídeo. Este conjunto pode ser utilizado em diversas combinações para gerar amplicons parciais do cassete.	
[32]	mid-5'GAL2-R	GAAATACTTTGGTCTTGGTGG		
[33]	mid-3'GAL2-F	CCACCAAGACCAAAGTATTTTC		
[34]	mid-5'GAL3-R	ACCACCAGTAGGAATATCGC		
[35]	mid-3'GAL3-F	GCGATATTCCTACTGGTGGT		
[36]	mid-5'GAL4-R	CATGTCATCCTCTTCAGACC		
[37]	mid-3'GAL4-F	GTAGCTATATGGACAATCATAATGT		
[38]	mid-5'GAL80-R	TCTCAAAATATATGGTGATTTAC		
[39]	mid-3'GAL80-F	GTAATCACCATATATTTTGAGA		
[40]	mid-5'G418-R	GCATAAATTCCGTCAGCC		
[41]	pPIC9-F	CTTGCATGCACCATTTCCTTGC		Amplifica o vetor pPIC9 excluindo-se o terminador <i>AOX1</i> , o fragmento do gene <i>HIS4</i> , o sítio de clonagem múltipla e o sinal de secreção de proteínas fator- $\alpha$ .
[42]	pPIC9-R	TTAGCATGCTTCGAATAATTAGTTGTTTT		
[43]	5'UTR-KU70-F	ATAGCGGCCGCCCTGAAAAAGTACTTTGAC		Amplifica a ORF de <i>KU70</i> junto de 1.000 pb up/downstream a partir do genoma de <i>K. phaffii</i> GS115. Foram adicionadas sequências de endonucleases para clonagem no vetor pPICH.
[44]	3'UTR-KU70-R	ATAGAATTCAAGAAGATTTTGAATTTACG		
[45]	KU70-3HA-F	ATACATATGTACCCATACGATGTTCCAGATTA CGCTTACCCATACGATGTTCCAGATTACGCTT ACCCATACGATGTTCCAGATTACGCTAGTGTT GTCAGCAAGCAATA	Amplifica a ORF <i>KU70</i> adicionando um sítio de <i>NdeI</i> no códon inicial e a sequência 3xHA para substituição no gene nativo.	

[46]	Mid-KU70-R	ATATCTAGAAATTCGTTGGCGGATCTCCTCAA T	Anela no sítio de <i>Xba</i> I na posição 661 da ORF de <i>KU70</i> .
[47]	lacZ-F	TCGAGAACATATGGGGGATCCATGACCGGAT CCGGAG	Amplifica o gene <i>lacZ</i> a partir do vetor pZ0. O fragmento contém extremidades para recombinação homóloga com o vetor pPICH70-3tc, onde substitui-se o gene <i>KU70</i> por <i>lacZ</i> .
[48]	lacZ-R	ACGTTGCGAGGAAAAAGTAAGGAACACGAATT ATTTTTGACAC	

#### **4.9 Construção dos plasmídeos pV2A-T para o cassete policistrônico GAL**

Os genes do regulon GAL e o promotor TEF1 foram obtidos a partir do DNA genômico de *S. cerevisiae* S288c. A marca de seleção G418 e o terminador *CYC1* foram amplificados a partir do plasmídeo pPICK2. As primeiras amplificações dos fragmentos citados acima foram realizadas com os oligonucleotídeos da série Regal ([1] a [14]). Para a enzima Phusion, foram utilizadas as seguintes condições de reação: desnaturação inicial a 98 °C por 1 minuto seguido de 40 ciclos de desnaturação a 98 °C por 30 segundos, anelamento por 30 segundos com temperatura variável de acordo com os pares de oligonucleotídeos utilizados (gradiente de 60 a 72 °C variando 0,3 °C por ciclo) e extensão a 72 °C com tempo variável calculado de 30 segundos para cada quilobase amplificada. A extensão final foi feita a 72 °C por 5 minutos e as reações foram armazenadas a -20 °C até sua utilização.

A ligação de vetores a produtos de PCR ou a fragmentos provenientes de digestão utilizou a T4 DNA ligase e a solução tampão da fabricante USB Technologies. As ligações foram feitas a 16 °C por 16 horas. Quando houve necessidade de defosforilação do vetor a ser utilizado, foi utilizada a Shrimp Alkaline Phosphatase (Promega) com o tampão e protocolo fornecido pelo fabricante. A digestão de fragmentos de DNA com enzimas de restrição utilizou a solução tampão e a temperatura indicados pelo fabricante por um mínimo de duas horas.

Também foi utilizado o método de clonagem por recombinação *in vivo* como descrito por (JACOBUS; GROSS, 2015). Os fragmentos foram amplificados com sequências homólogas ao vetor para permitir a montagem *in*



vivo dos plasmídeos, onde quantidades equimolares dos fragmentos foram utilizados para a transformação da bactéria juntamente com 250 ng do vetor onde as sequências foram clonadas, respeitando-se a relação 2:1 de inserto:vetor.

No caso de clonagens que não puderam ser feitas em *E. coli*, o vetor e os fragmentos contendo sequências homólogas para montagem por recombinação foram transformados diretamente em *K. phaffii* através do protocolo de transformação de levedura por eletroporação.

#### **4.10 Transformação de leveduras por eletroporação**

Uma colônia fresca de levedura foi inoculada em 10 mL de meio YPD em um frasco Erlenmeyer de 250 mL e incubada a 30 °C e 200 rpm por 24 h. Essa cultura foi utilizada para inocular 100 mL de meio YPD em frasco Erlenmeyer de um litro, o qual foi incubado a 30 °C e 200 rpm pelo tempo necessário para que a OD<sub>600</sub> chegasse a 1,0. As células foram então incubadas em banho de água e gelo por 10 minutos e em seguida, centrifugadas a 3000 x g por 5 minutos a 4 °C. Foram feitas duas lavagens subsequentes com água destilada estéril gelada, sendo elas de 100 mL e 50 mL, respectivamente. Uma terceira lavagem com 4 mL de sorbitol 1 M gelado foi feita. Ao final, as células foram ressuspensas em 100 µL de sorbitol 1 M gelado e 80 µL de células foram usados para a eletroporação com 5-10 µg de DNA. A eletroporação utilizou uma cubeta de 2 mm e parâmetros de 1,5 kV, 25 µF e 200 Ω. Um mL de meio YPD gelado foi imediatamente adicionado à cubeta. No caso de plasmídeos com marca de seleção de resistência a antibióticos, as células foram recuperadas por 2 h a 30 °C em meio YPD sem agitação e depois semeadas em meio seletivo.

#### **4.11 Microscopia de fluorescência e triagem de transformantes positivos**

Uma colônia da cepa X-33 de *K. phaffii* cotransformada com os plasmídeos pV2A\_regal e pPICH-Pgal1-gfp foi cultivada em meio mínimo (MM) suplementado com sorbitol 1% (p/v) durante 48 horas a 30 °C sob agitação a 200 rpm. As células foram coletadas por centrifugação a 2.000 x g por 5 minutos e lavadas com água destilada estéril duas vezes. As células foram imediatamente reinoculadas em meio YNB fresco contendo galactose 2% (p/v). As células foram cultivadas em condições similares por 24 horas. Uma alíquota de 10 µL do meio de cultivo foi colhida e depositada em uma lâmina de vidro para análise no microscópio invertido de fluorescência Zeiss Axio Observer Z1 equipado com lente objetiva de 63x de aumento NA 1,4. A detecção da proteína eGFP foi feita com excitação a 488 nm, emissão a 507nm e ganho de sinal automático. As imagens foram capturadas com a câmera AxioCam MRm (Carl Zeiss) e processadas pelo software Zen (Carl Zeiss). As figuras demonstradas na sessão de Resultados são representativas de três experimentos independentes. O controle negativo foi realizado com células X-33 com a ausência dos plasmídeos.

#### **4.12 Extração de proteínas de *K. phaffii***

Uma colônia isolada obtida de placa de petri contendo meio YPD suplementado com o antibiótico Zeocina (100 µg/mL) e ágar 2% (v/v) foi inoculada em 10 mL de meio YPD líquido suplementado com o antibiótico Zeocina em frasco de 250mL a partir de uma colônia isolada em meio YPD também (YPDZeo) (100 µg/mL) a 30 °C sob agitação de 200 rpm até atingir fase

estacionária de crescimento (aproximadamente 24 horas). Um novo inóculo foi feito em 50mL de meio líquido YPDZeo a uma  $DO_{600nm}$  de 0,1 a 30 °C sob agitação de 200 rpm até atingir  $DO_{600nm}$  de 1,0. As células foram coletadas por centrifugação por 10 min a 5.000 x *g* a 4 °C. As células foram lavadas com água destilada gelada por duas vezes, repetindo-se as etapas de centrifugação entre as lavagens. Foi adicionado 450 µL de tampão de extração (50 mM Tris-HCl, Tween 20 1,25% (v/v), pH 8,8) ao precipitado e as células foram colocadas sob estresse mecânico com vigorosa agitação na presença de pérolas de vidro de tamanho entre 425 e 600 nm, em ciclos de 30 seg sob agitação intercalados por 30 segundos em banho de água e gelo, em um total de 5 ciclos. A fase líquida foi transferida por pipetagem para um novo tubo e centrifugada a 4 °C por 10 minutos a 12.000 x *g*. O sobrenadante contendo a fração solúvel de proteínas foi transferido para um novo tubo e mantido sob refrigeração até o momento do uso ou congelado a -20 °C caso não fosse utilizado imediatamente após a extração.

#### **4.13 Dosagem de proteínas solúveis em extrato total de *K. phaffii***

As proteínas foram dosadas pela metodologia modificada de Bradford (BioRad). O método consiste na reação em microplaca entre o reagente Coomassie Brilliant Blue G-250 que se liga às proteínas em condições acídicas, mudando da coloração vermelha para azul. A quantidade de corante azul formada é proporcional à quantidade de proteína e mensurada pela absorbância a 595 nm. A curva de calibração foi feita utilizando diluições de Albumina Sérica Bovina (BSA) variando a concentração de proteína total entre 0,125 a 2 mg/mL. Foram utilizados 5 µL do padrão/amostra para 250 µL de reagente. A reação

ocorreu por 5 minutos. O branco foi realizado com reagente e água destilada substituindo a amostra.

#### **4.14 Extração rápida de DNA genômico para confirmação de transformantes**

Para confirmação da integração dos cassetes no *locus* desejado, as colônias transformantes a serem testadas foram ressuspensas em 50 µL de água destilada estéril. Em seguida, foi adicionada 1 µL de solução de Liticase (2 mg/mL) e incubada a 30 °C por 1 hora. A lise dos esferoblastos formados foi feita pela adição de 20 mM NaOH a 100 °C por 5 minutos. Os debris celulares foram centrifugados a 10.000 x g por 1 minuto. O volume de 2 µL foi utilizado nas reações de PCR com a Taq Polimerase (NEB).

#### **4.15 Construção dos vetores de expressão de *KU70* e *lacZ***

O vetor episomal pPICHOLI (Mobictec Molecular Biotechnology, Alemanha) foi utilizado como arcabouço inicial para a construção dos vetores utilizados neste trabalho. O vetor inicial foi digerido com as enzimas de restrição *Bam*HI e *Not*I para que fosse feita a linearização do vetor e extração do cassete de promotor e terminador do gene *AOX1*, nativo de *K. phaffii*. O vetor resultante foi denominado como pPICH. Então, o gene repórter *lacZ* foi amplificado por PCR a partir do vetor pLGΔ178 (GUARENTE, 1983) utilizando a Phusion DNA polimerase (New England Biolabs, EUA) sob as seguintes condições: desnaturação inicial a 98 °C por 1 minuto seguido de 40 ciclos de desnaturação a 98 °C por 30 segundos, anelamento de oligonucleotídeos a 60 °C por 30

segundos e extensão a 72 °C por 2 minutos. Uma etapa adicional de extensão a 72 °C por 5 minutos foi realizada logo em seguida e a reação foi finalizada e armazenada a -20 °C até posterior análise. As concentrações dos reagentes seguiram as recomendações do fabricante. Foram utilizados os oligonucleotídeos [47] e [48] a uma concentração final de 0,25 µM para amplificação do gene *lacZ*. O fragmento de DNA amplificado nesta reação foi clonado no vetor pPICH utilizando o DNA assembly kit (New England Biolabs). Foi utilizada a proporção vetor-inserto de 1:2 com massa total de 50 ng de DNA e uma sequência de sobreposição de 25 nucleotídeos complementares. A reação foi incubada a 50 °C por 60 minutos e subsequentemente transformada na cepa quimiocompetente XL-10 Gold de *E. coli*. As células foram semeadas em LB ágar contendo zeocina (25 µg/mL) e os transformantes corretos foram selecionados via confirmação por PCR e análise de fragmentos gerados por digestão por enzimas de restrição. O vetor resultante foi chamado de **pPICH-lacZ**.

O gene *KU70* (Número de acesso GenBank CP014717.1) foi amplificado do DNA genômico extraído da cepa GS115 de *K. phaffii* (Invitrogen). Os oligonucleotídeos [43] e [44] foram utilizados para amplificação em condições de reação de PCR idênticas às usadas para o gene *lacZ*. A clonagem do gene *KU70* no vetor pPICH se deu de forma similar à utilizada para o gene *lacZ*. O vetor resultado foi chamado de **pPICH-ku70**.

Uma sequência de 3 tags HA *in tandem* com o códon inicial dos genes foi adicionada ao gene *KU70* via PCR (oligonucleotídeos [45] e [46]) para permitir que o epítipo da proteína Hemaglutinina A pudesse ser utilizado como marcador da proteína Ku70p por western blotting posteriormente.

O aptâmero de tetraciclina foi sintetizado pela empresa GenOne (Brasil) e clonado no plasmídeo pUC19. A sequência do aptâmero foi amplificada via PCR, purificada, digerida pelas enzimas de restrição *SphI* e *NdeI* (New England Biolabs) e clonados nos vetores pPICH-lacZ e pPICH-ku70. Os novos vetores doravante denominados **pPICH-lacZ-3tc** e **pPICH-ku70-3tc** foram transformados em *K. phaffii* pelo método de eletroporação. As transformantes resultantes deste processo que contém o plasmídeo pPICH-lacZ-3tc foram identificadas como TRS Blue (do inglês *tetracycline-responsive strain*). As transformantes do plasmídeo pPICH-ku70-3tc foram identificadas neste trabalho como TRS ku70. Ambos os plasmídeos contêm 3 sequências repetidas do aptâmero de tetraciclina (*riboswitch*) de tamanho total 295pb clonadas *in tandem* ao códon inicial. Os plasmídeos pPICH-lacZ e pPICH-ku70 foram utilizados como controles negativos nas reações por não ter o *riboswitch* de tetraciclina.

#### **4.16 Atividade de $\beta$ -galactosidase**

Uma cultura inicial de células de *K. phaffii* contendo os plasmídeos pPICH-lacZ ou pPICH-lacZ-3tc foram crescidas por 12 horas a 30 °C e 200 rpm de agitação em meio YPD suplementado com zeocina e com presença ou ausência de tetraciclina. As células foram coletadas por centrifugação a 3.000 x g por 5 minutos e reinoculadas em meio YPD fresco a uma concentração final de 0,1 DO<sub>600nm</sub>. As células foram crescidas até uma concentração entre 0,7 – 0,8 DO<sub>600nm</sub> e foram coletadas por centrifugação. O pellet foi ressuspensionado em 1 mL de tampão Z (60 mM Na<sub>2</sub>HPO<sub>4</sub>, 10 mM KCl and 1 mM MgSO<sub>4</sub>). A permeabilização das células ocorreu com a adição de 25  $\mu$ L de uma solução de SDS 0,1% (p/v) e 50  $\mu$ L de clorofórmio. Então, foram feitos 3 ciclos de vigorosa

agitação por 30 segundos, incubação a 30 °C por 15 minutos e 5 minutos a - 80 °C (*freeze and thaw*). Após realizada a ruptura das células, a reação de dosagem da atividade enzimática foi iniciada com a adição de 200 µL de uma solução de ONPG 4 mg/mL a 30 °C. A reação foi incubada até que obtida a coloração amarela após aproximadamente 5 minutos. A interrupção da reação foi feita a partir da adição de 500 µL de 1M NaCO<sub>3</sub>. As amostras foram imediatamente centrifugadas para separação do debris celular e a absorbância a 600nm e 420nm foi dosada imediatamente. A atividade de β-Galactosidase foi calculada através da fórmula:

$$Miller\ units = 1,000 \times \left( \frac{OD_{420nm}}{OD_{600nm} * V * T} \right)$$

onde T é o tempo de reação em minutos e V é o volume de cultura utilizado para obtenção da quantidade de células necessária.

#### **4.17 Análise estatística**

Toda a análise estatística descritiva dos resultados apresentados neste trabalho foi realizada utilizando o software GraphPad Prism. O teste t não pareado entre duas amostras foi calculado a partir dos dados obtidos de três experimentos independentes com triplicatas experimentais. As figuras e gráficos apresentados são representativos dos resultados médios obtidos nos experimentos, demonstrando o desvio padrão quando possível.

#### **4.18 Eletroforese de proteínas em gel de poliacrilamida (PAGE) e Western Blotting**

Os extratos de proteína realizados nas amostras foram analisados em gel de poliacrilamida SDS-Page (5% stacking, 12% resolving). Previamente à aplicação no gel, as amostras foram incubadas a 100 °C por 5 minutos em tampão de carregamento contendo 100 mM Tris-HCl (pH 6,8), 8M ureia, 20% (v/v) glicerol, 4% (p/v) SDS, 0.2% (p/v) azul de bromofenol and 100 mM  $\beta$ -mercaptoetanol. A eletroforese ocorreu em temperatura ambiente a uma corrente constante de 0,05 A por gel presente na cuba. O conteúdo do gel foi revelado com a incubação em solução coloidal de Coomassie Brilliant Blue G-250 (Merck, Alemanha). Todos os géis foram preparados em duplicata onde um foi utilizado para revelação e outro para o western blotting. A transferências das proteínas do gel de poliacrilamida para membranas de nitrocelulose (Merck Millipore) foram feitas em cuba semiúmida com todo o material embebido em tampão de transferência (25 Mm Tris-HCl, 190mM glicina, metanol 20% (v/v)) por 50 minutos a temperatura ambiente e corrente constante de 0,15 A. A qualidade da transferência foi verificada com a coloração reversível da membrana de nitrocelulose com solução de Ponceau S (Ponceau S 0,1% [p/v], ácido acético glacial 5% [v/v]). O bloqueio da membrana foi realizado com incubação por 1 hora em 5% (p/v) leite em pó reconstituído em tampão TBST. Em seguida, a membrana foi incubada em tampão TBST contendo 1  $\mu$ g/mL de anticorpo monoclonal anti-HA de coelho (SAB5600193, Sigma Aldrich) e 3% (p/v) BSA por 1 hora em temperatura ambiente. Em seguida, a membrana foi exposta por 1 hora a temperatura ambiente ao anticorpo secundário conjugado com peroxidase anti-imunoglobulina de coelho proveniente de cabra (ab6721, Abcam). A revelação



foi feita após incubação da membrana conjugada aos anticorpos com o substrato quimioluminescente (Abcam) por 2 minutos em câmara escura. A imagem foi obtida em modo semiautomático após exposição de 1 minutos através do Imager 680 (Amersham). Entre cada etapa de incubação (bloqueio, anticorpo primário, anticorpo secundário e revelação) foram feitas três etapas de lavagem da membrana que consistiam na incubação da mesma em tampão TBST a temperatura ambiente por 5 minutos. Medidas da intensidade de banda por densitometria foram feitas utilizando o software ImageJ 1.54 através do cálculo de área sob a curva. O background foi normalizado. Nenhuma edição foi realizada nas figuras apresentadas. Brilho e contraste foram ajustados para melhor visualização dos dados. Estimativa de tamanho de Ku70p é de aproximadamente 60 kDa.

## **5 Resultados e Discussão**

Para melhor compreensão dos resultados obtidos, esta sessão foi dividida em duas partes de acordo com os Objetivos deste trabalho apresentados acima.

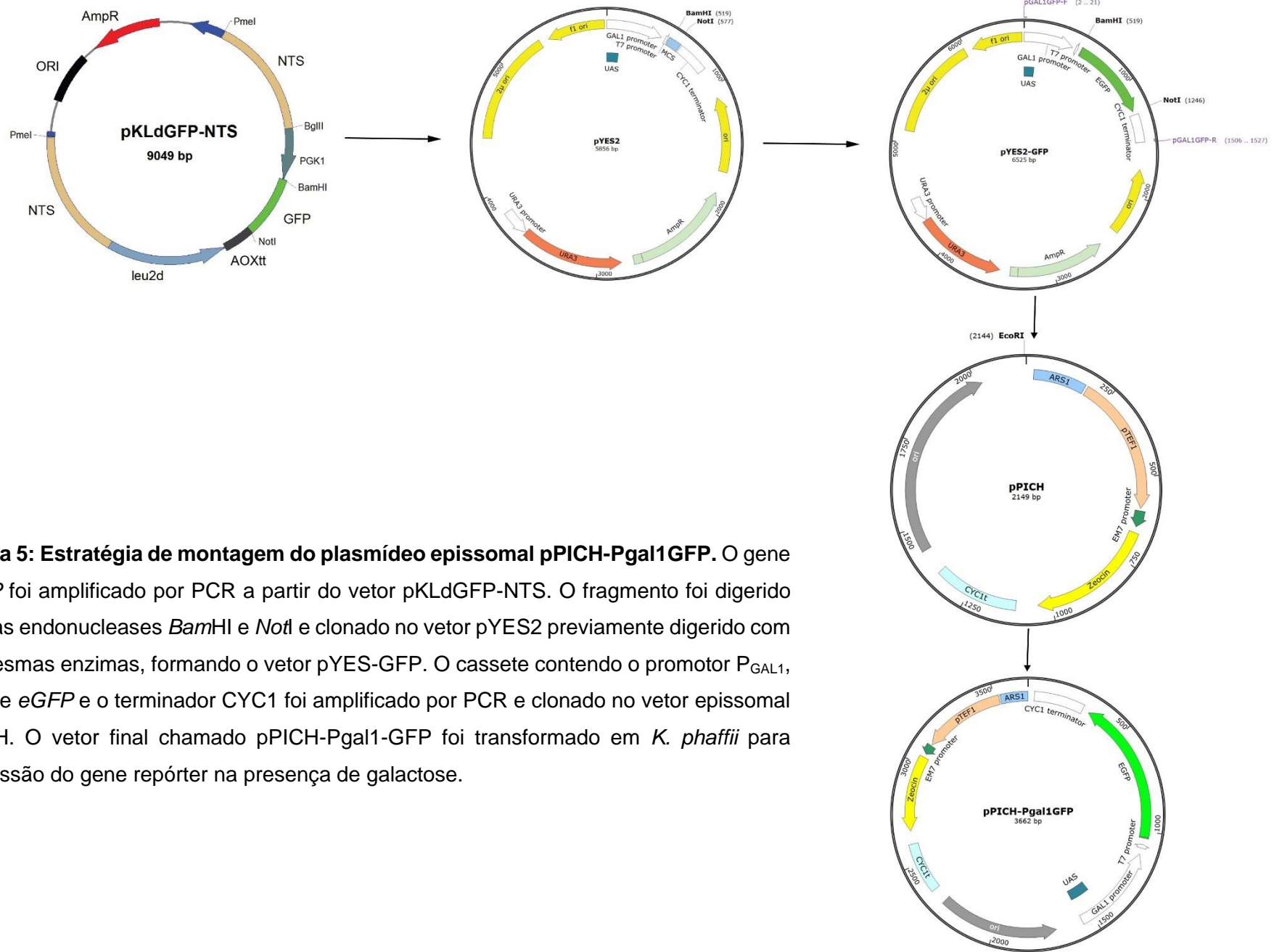
### **5.1 Desenvolvimento da linhagem de *K. phaffii* contendo o sistema de controle de expressão gênica baseado no regulon GAL de *S. cerevisiae*.**

#### **5.1.1 Construção do sistema repórter ativado por galactose**

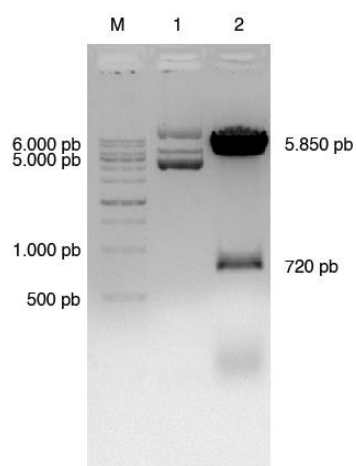
O vetor **pPICH-Pgal1-GFP** foi desenvolvido com o objetivo de expressar a proteína *eGFP* como gene repórter na presença de galactose. Para isso, foram necessárias várias etapas de clonagens para permitir a montagem correta do

cassete policistrônico GAL (Figura 5). O gene *eGFP* foi obtido através de clivagem com as endonucleases *Bam*HI e *Not*I a partir do plasmídeo **pKLDGFP-NTS** (BETANCUR, 2017). A digestão do plasmídeo foi analisada em gel de agarose 1% (p/v) e o fragmento com peso molecular correspondente ao *eGFP* (720 pb) foi excisado do gel (Figura 6). O fragmento de DNA foi eluído da agarose e purificado utilizando o kit Wizard Gel and PCR Clean-Up System (Promega) como recomendado pelo fabricante. Subsequentemente o gene *eGFP* foi ligado ao vetor **pYES2**, previamente digerido com as enzimas *Bam*HI e *Not*I, através da reação com a enzima T4 DNA ligase (NEB). Com isso, o vetor resultante, adiante a ser chamado de **pYES2-GFP** contém a ORF *eGFP* sob controle do promotor  $P_{GAL1}$ . O produto da ligação foi então transformado por choque térmico na linhagem DH10 alpha de *E. coli* para amplificação de massa do vetor. As células transformantes foram selecionadas em placas de LB suplementadas com ampicilina (200 µg/mL). Foram escolhidas 6 colônias aleatoriamente para extração de DNA plasmidial por lise alcalina e confirmação da correta clonagem. O plasmídeo extraído das colônias transformantes foi testado através da digestão com endonucleases e amplificação por PCR do fragmento contendo o promotor  $P_{GAL1}$ , o gene *eGFP* e o terminador *CYC1*, com os oligonucleotídeos [1] e [2] (Figura 7A). O cassete  $P_{GAL1}$ -GFP amplificado foi purificado para clonagem no vetor epissomal pPICH, previamente linearizado com a endonuclease *Eco*RI e defosforilado com a enzima fosfatase alcalina de camarão (SAP, do inglês *shrimp alkaline phosphatase*). Esta etapa foi necessária pois o vetor pPICH contém a sequência de replicação autônoma (ARS) necessária para o funcionamento em *K. phaffii*. O vetor linearizado foi cotransformado na linhagem DH10α de *E. coli* juntamente com o cassete

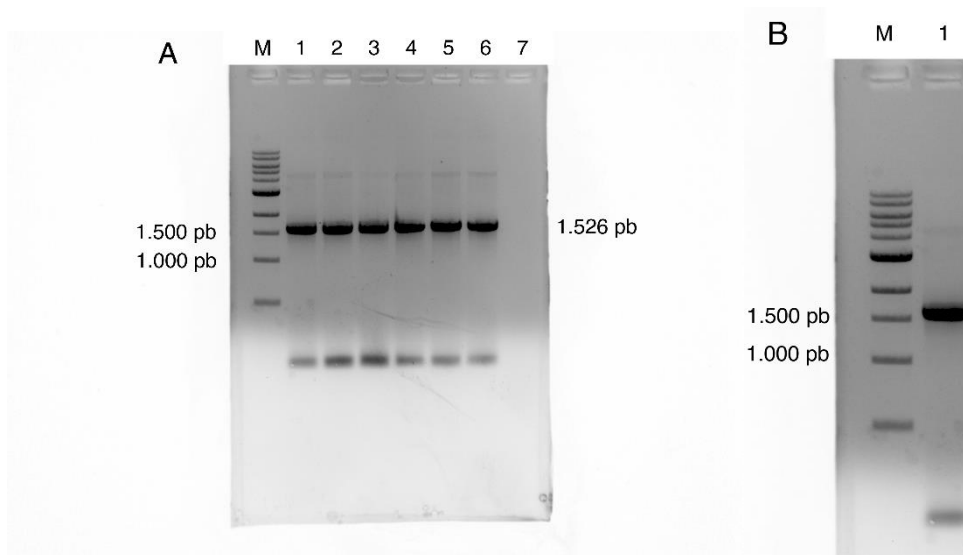
PGAL1-*eGFP*, em razão molar vetor:inserir de 1 para 2, com 250 ng de massa total de DNA, pela técnica de choque térmico. Uma vez que o inserir e o vetor possuem regiões homólogas com tamanho de 30 pb em cada extremidade, as células bacterianas são capazes de fazer a recombinação entre o vetor e inserir, finalizando a clonagem. As colônias transformantes foram selecionadas em meio LB suplementados de zeocina (25 µg/mL). O vetor resultante foi denominado pPICH-PGAL1-GFP. Após seleção de transformantes e extração do plasmídeo, a confirmação da correta inserção do fragmento de interesse foi avaliada por PCR com amplificação do cassete PGAL1-GFP-CYC1TT (oligonucleotídeos [1] e [2]) e análise de produtos de digestão por endonucleases (Figura 7B). Com este vetor é possível testar o funcionamento do regulon GAL, uma vez que o GFP será ativado somente na presença de galactose e de todas as proteínas reguladoras do sistema.



**Figura 5: Estratégia de montagem do plasmídeo episomal pPICH-Pgal1GFP.** O gene *eGFP* foi amplificado por PCR a partir do vetor pKLdGFP-NTS. O fragmento foi digerido com as endonucleases *Bam*HI e *Not*I e clonado no vetor pYES2 previamente digerido com as mesmas enzimas, formando o vetor pYES-GFP. O cassete contendo o promotor P<sub>GAL1</sub>, o gene *eGFP* e o terminador *CYC1* foi amplificado por PCR e clonado no vetor episomal pPICH. O vetor final chamado pPICH-Pgal1-GFP foi transformado em *K. phaffii* para expressão do gene repórter na presença de galactose.



**Figura 6: Digestão do plasmídeo pYES2-GFP para retirada do gene *eGFP* e confirmação da correta clonagem.** Os fragmentos gerados na digestão são compatíveis com a correta clonagem. Este plasmídeo foi utilizado para amplificação do cassete PGAL1-GFP-CYC1TT utilizado para o teste de ativação de *eGFP* por galactose. M: Marcador de peso molecular 1 kb DNA ladder (Sinapse Biotecnologia); 1: pYES2 não digerido; 2: pYES2 digerido com *Bam*HI e *Not*I (tamanho esperado 720 pb).



**Figura 7: Digestão do vetor pPICH-Pgal1-GFP para confirmação da inserção do fragmento de *eGFP* regulado por galactose.** A) Fragmentos de PCR do cassete PGAL1-GFP-CYC1TT amplificado do vetor pYES2-GFP. Tamanho esperado: 1.526 pb. M: Marcador de peso molecular 1 kb DNA ladder (Sinapse Biotecnologia); 2-6: fragmentos de PCR resultantes da amplificação do conteúdo plasmidial das colônias transformadas com o vetor pYES2-GFP; 7: controle negativo. B) Amplificação do

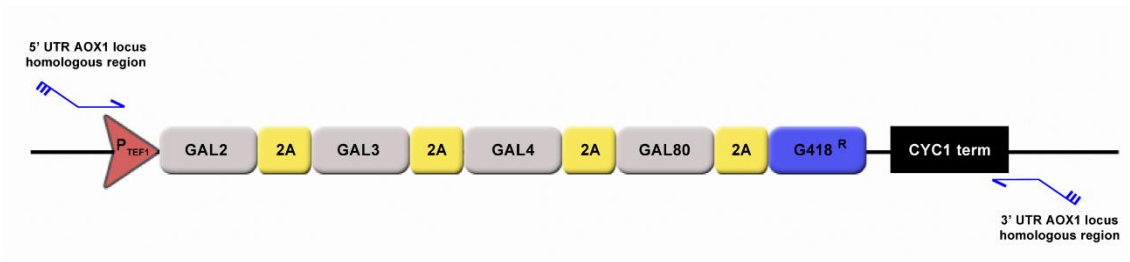
cassete clonado no vetor pPICH-Pgal1GFP. Tamanho esperado: 1.526 pb. M: Marcador de peso molecular 1 kb DNA ladder (Sinapse Biotecnologia); 1 – fragmento de PCR do cassete.

### 5.1.2 Construção e inserção do regulon GAL em *K. phaffii*

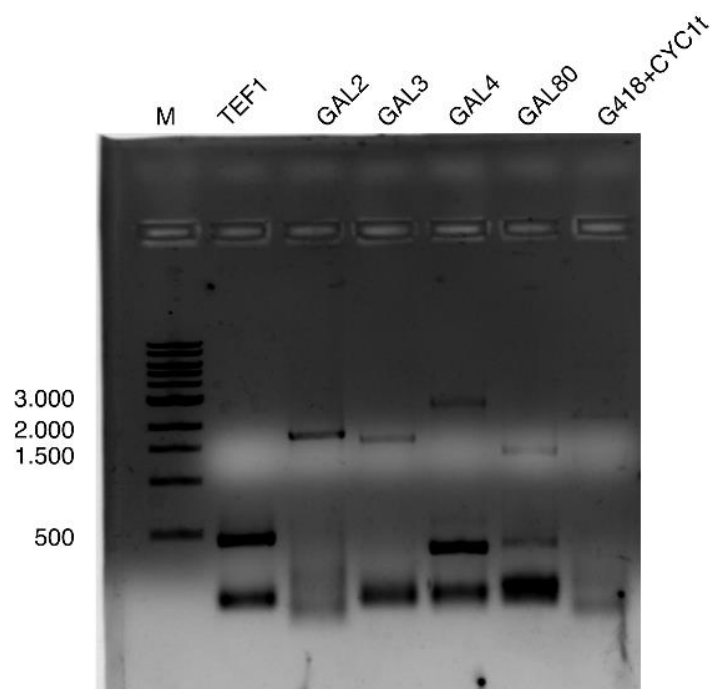
Para que o correto funcionamento do sistema fosse alcançado, utilizou-se a estratégia de construção de um cassete policistrônico contendo os genes da família GAL responsáveis pela ativação da transcrição na presença de galactose. Na **Figura 8** temos a representação gráfica da estrutura geral do regulon GAL. O cassete tem sua transcrição regulada pelo promotor *TEF1* (389 pb) de *S. cerevisiae*, que já havia sido utilizado por nosso grupo de pesquisa e demonstrado eficiência similar ao *TEF1* nativo de *K. phaffii*. Os genes *GAL2* (1.725 pb), *GAL3* (1.563 pb), *GAL4* (2.646 pb) e *GAL80* (1.308 pb), que foram amplificados via PCR a partir do genoma de *S. cerevisiae* S288c com a retirada dos códons de parada de cada gene, para permitir a completa transcrição do cassete policistrônico. Esta etapa é necessária pois o ribossomo poderia interromper a tradução do mRNA prematuramente ou gerar um erro na fase leitura, gerando uma proteína não funcional. Cada gene foi intercalado pelo fragmento 2A, que gera uma quebra da proteína durante a tradução do mRNA. Isto ocorre através do mecanismo de deslizamento do ribossomo, liberando as proteínas individualmente contendo 6 aminoácidos no C-terminal e 2 aminoácidos no N-terminal das proteínas liberadas. Além disso, foi adicionado ao cassete a marca de seleção dominante  $G418^R$ , que confere resistência ao antibiótico geneticina, a fim de permitir a correta seleção das células transformantes. O gene  $G418^R$  foi posicionado ao final do regulon para garantir que todas as outras sequências codantes estariam sendo corretamente

transcritas. Isto se deve ao fato de que todo o cassete possui apenas uma sequência iniciadora de transcrição, um códon inicial e um códon de parada. Do contrário, o posicionamento do gene de resistência ao antibiótico em outra posição poderia levar ao aparecimento de falsos positivos, onde as células poderiam abortar a tradução do mRNA logo após a tradução da marca de resistência ao antibiótico. Portanto, os genes subsequentes deixariam de ser traduzidos e as outras proteínas necessárias ao sistema de regulação não seriam expressas. Todas as ORFs possuem cerca de 20 – 25 pb de homologia entre si para permitir a correta montagem direcionada do cassete. O terminador CYC1 foi escolhido para controlar a finalização da transcrição do cassete. Nos terminais 5' e 3' do cassete policistrônico foram inseridas sequências homólogas ao *locus* do gene *AOX1*, a fim de posteriormente direcionar a integração genômica deste fragmento de DNA e gerar a deleção do gene *AOX1*. A deleção do gene *AOX1* confere o fenótipo Mut<sup>S</sup> (baixa taxa de crescimento utilizando metanol como fonte exclusiva de carbono). Desta forma, seria possível fazer a dupla seleção dos clones.

Uma vez feita a amplificação e purificação de todos os fragmentos de DNA do cassete (Figura 9), a reação de transformação por eletroporação de *K. phaffii* foi montada contendo quantidades equimolares dos fragmentos com um total de 10 µg de DNA, objetivando a montagem *in vivo* e integração genômica do cassete.



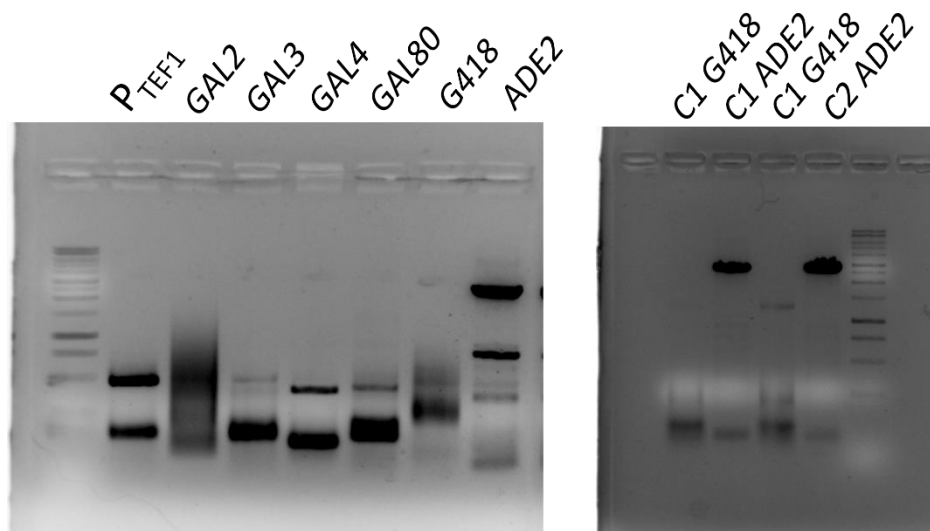
**Figura 8: Representação do cassete policistrônico desenhado para expressão dos genes da família GAL.** O cassete tem sua transcrição controlada pelo promotor constitutivo forte *TEF1* de *S. cerevisiae*. Cada gene da família GAL está intercalado por uma sequência de quebra 2A para permitir a separação das proteínas durante a tradução do mRNA. O gene *G418<sup>R</sup>* confere resistência ao antibiótico geneticina para seleção de transformantes. Os terminais 5' e 3' possuem regiões homólogas ao *locus* do gene *AOX1* para permitir a integração do cassete e simultânea deleção do gene. Desta forma estes genes seriam expressos constitutivamente e permitiriam a regulação da expressão de *eGFP* a partir da presença de galactose no meio de cultivo.



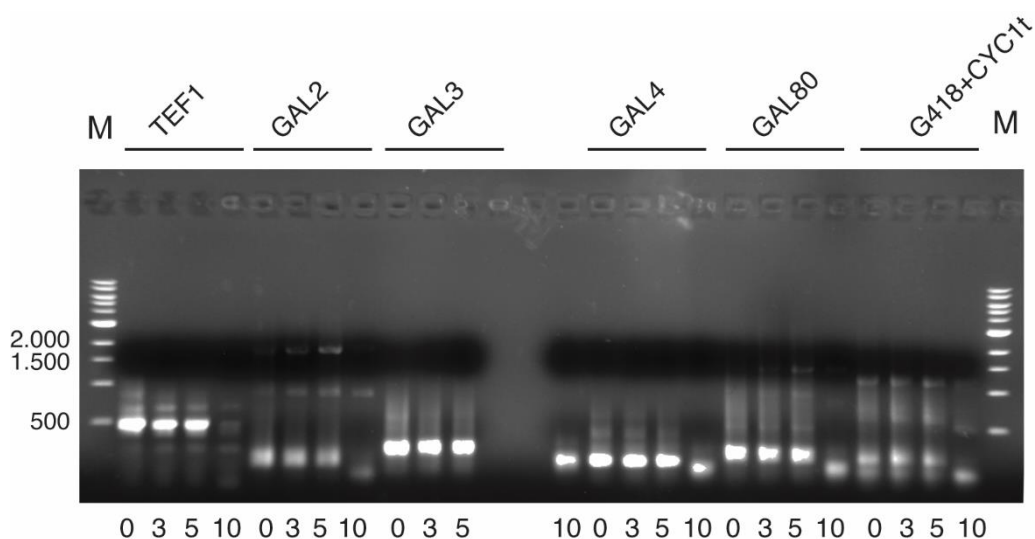
**Figura 9: Amplificação dos elementos do regulon GAL.** Os genes *GAL2*, *GAL3*, *GAL4* e *GAL80* foram amplificados por PCR a partir do DNA genômico de *S. cerevisiae* S288c para montagem do cassete policistrônico. O gene *G418<sup>R</sup>* junto do terminador *CYC1* foi amplificado do vetor pPICK2. Tamanhos esperados: *GAL2* (1.722pb); *GAL3* (1.560 pb); *GAL4* (2.643 pb); *GAL80* (1.305 pb); *G418<sup>R</sup>* + *CYC1*\_TT (1.134 pb). M: Marcador de peso molecular 1kb DNA ladder (Sinapse Biotecnologia).



Após a transformação, cerca de 200 colônias foram selecionadas e transferidas para uma nova placa de YPD suplementado com geneticina. O DNA genômico destas colônias foi extraído e um subseqüentemente foi testada a presença dos genes do regulon GAL por PCR utilizando-se os oligonucleotídeos da série “mid” ([31] a [40]). Não houve amplificação em nenhum dos clones testados. Porém, ao testar a amplificação individual dos genes (oligonucleotídeos [3] a [14]), houve amplificação de apenas parte dos genes GAL e da marca de seleção G418<sup>R</sup> (Figura 10). Esta estratégia se mostrou ineficiente para a correta montagem e integração do cassete GAL devido ao fato de *K. phaffii* possuir um sistema de recombinação homóloga (HR) pouco eficiente. O resultado demonstra que os fragmentos de DNA foram integrados em regiões distintas do genoma por recombinação não homóloga (NHEJ). Uma nova tentativa foi feita utilizando maiores sequências de homologia entre os fragmentos (45-50 pb), após alongamento com os oligonucleotídeos [15] a [26], mas o mesmo resultado negativo foi obtido. Além disso, aumentou-se significativamente o grau de dificuldade para obtenção de produtos de PCR de boa qualidade, uma vez que vários produtos inespecíficos eram amplificados mesmo após extensivas adaptações de protocolo (Figura 11). Dada as dificuldades de clonagem *in vivo* em *K. phaffii*, optou-se por adaptar a estratégia de montagem utilizando a bactéria *E. coli* para clonagem dos fragmentos no plasmídeo **pPIC9**.

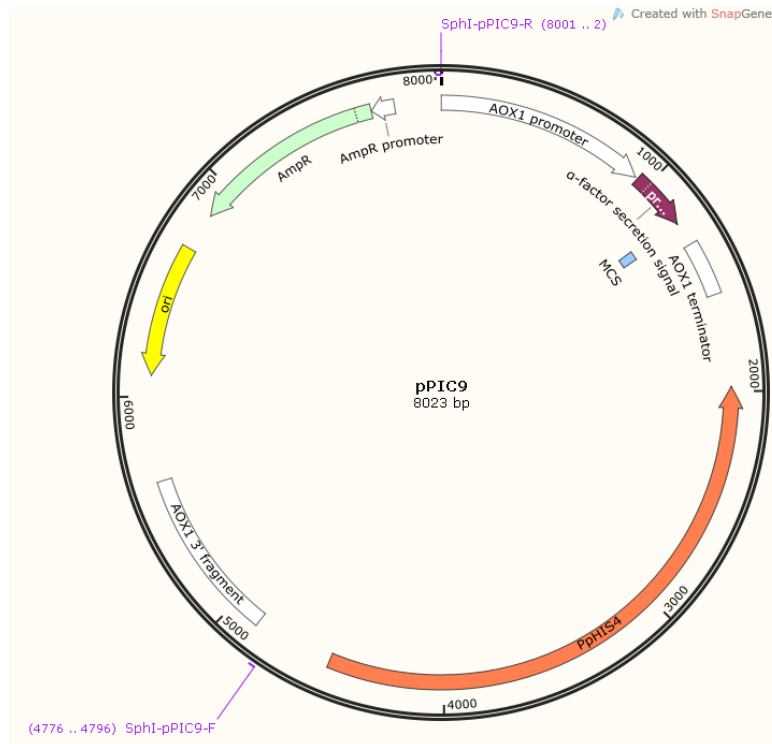


**Figura 10: Fragmentos de PCR obtidos a partir do DNA genômico dos clones transformantes de *K. phaffii* com o regulon GAL.** A) A confirmação da transformação com o cassete policistrônico foi realizada por PCR. As colônias transformantes semeadas em YPD + geneticina (200 µg/mL) tiveram seu DNA genômico extraído e cada fragmento do cassete foi testado. B) O DNA genômico de dois clones foram digeridos com a endonuclease NotI e testados por PCR para os genes G418<sup>R</sup> e ADE2, para eliminar a possibilidade de resultado negativo por interferência ou baixa eficiência de PCR devido ao super enovelamento do DNA genômico. O gene ADE2 foi utilizado nos dois casos como controle positivo reacional.



**Figura 11: Amplificação por PCR dos genes do regulon GAL com o conjunto de oligonucleotídeos ampRegal.** Os novos primers foram utilizados para realizar o alongamento dos fragmentos, facilitando a montagem e integração do cassete policistrônico. Foram testadas diversas condições reacionais, incluindo o uso de DMSO,

para obtenção dos novos fragmentos. M: Marcador de peso molecular 1 kb DNA ladder (Sinapse Biotecnologia); Inferior: concentração (%) de DMSO utilizada na reação; Superior: fragmento alvo da reação de PCR.



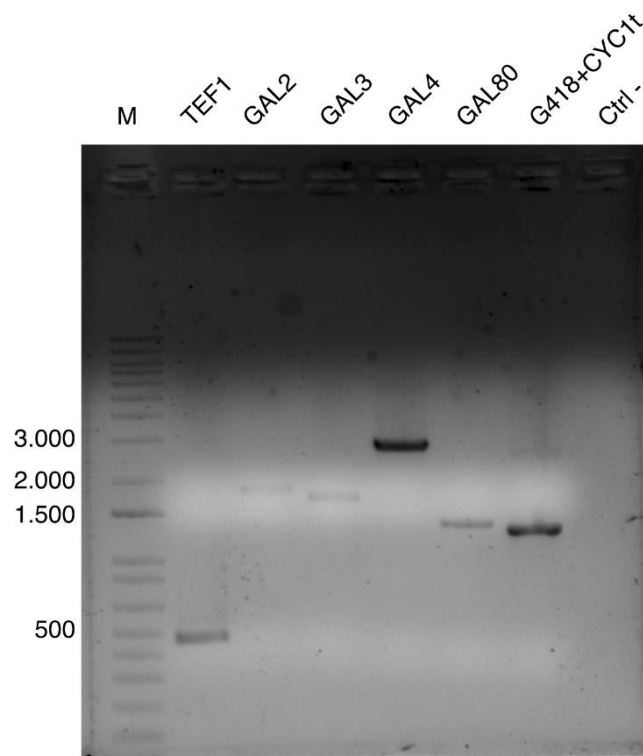
**Figura 12: Mapa do vetor pPIC9 utilizado para montagem do regulon GAL em *E. coli*.** Este vetor foi utilizado devido a suas características de estabilidade e direcionamento integrativo em *K. phaffii*, podendo ser clonado em *E. coli* para fácil manipulação do material genético.

O vetor integrativo pPIC9 (Figura 12) é um *shuttle vector* utilizado em *K. phaffii* para produção de proteínas recombinantes. Este vetor possui um sítio de clonagem entre o promotor e terminador AOX1, assim como o gene *HIS4* que pode ser utilizado para reversão de auxotrofia para histidina e o fragmento do terminal 3' do gene *AOX1* para integração genômica.

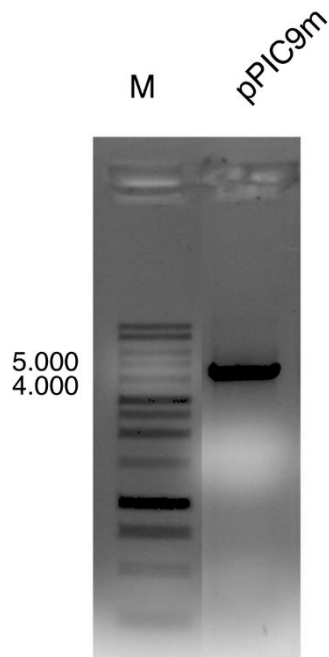
A Figura 13 mostra os novos fragmentos amplificados com regiões de homologia somente nos terminais 5' de cada fragmento. Para se fazer a

clonagem dos fragmentos do regulon GAL, o vetor pPIC9 foi amplificado por PCR para que se retirasse o gene *HIS4*, o terminador *AOX1* e o sinal de secreção do fator alpha (oligonucleotídeos [41] e [42]) (Figura 14). Assim, foi obtido o vetor mínimo pPIC9m. O vetor linearizado e as novas sequências foram amplificadas, purificadas e transformadas por eletroporação na cepa XL10 Gold de *E. coli* em quantidades equimolares, respeitando-se a proporção de 2:1 entre inserto e vetor. As células transformantes foram semeadas em meio LB contendo ampicilina. Após o crescimento, apenas uma colônia apresentou crescimento satisfatório e foi novamente plaqueada em meio contendo ampicilina. O DNA plasmidial foi extraído pelo método de lise alcalina e posteriormente testado por PCR para os genes do regulon GAL (Figura 15). Vê-se que somente o promotor  $P_{TEF1}$ , o gene *GAL2* e a sequência G418+CYC1t foram amplificados. Este resultado demonstra que a célula apenas foi capaz de montar o vetor parcialmente. Posteriormente, foi verificado que o promotor e o gene *GAL2* haviam sido corretamente montados. Sendo assim, as duas sequências unidas foram amplificadas utilizando-se os oligonucleotídeos [15] e [27] (Figura 16). Com isso, diminui-se o número de eventos de recombinação e aumenta-se a probabilidade da correta montagem. A transformação foi repetida em condições similares, agora contendo o fragmento único TEF1-GAL2 obtido na etapa anterior. Porém, o mesmo resultado foi obtido. Desta maneira, pode-se hipotetizar que a bactéria esteja recombinao internamente parte do vetor montado devido ao número repetitivo de sequências 2A. Uma alternativa seria a modificação dos códons das sequências 2A, mantendo-se o padrão de aminoácidos. No entanto, a execução destas modificações poderia levar a erros na sequência e seria extremamente laborioso. Sendo assim, optou-se por uma

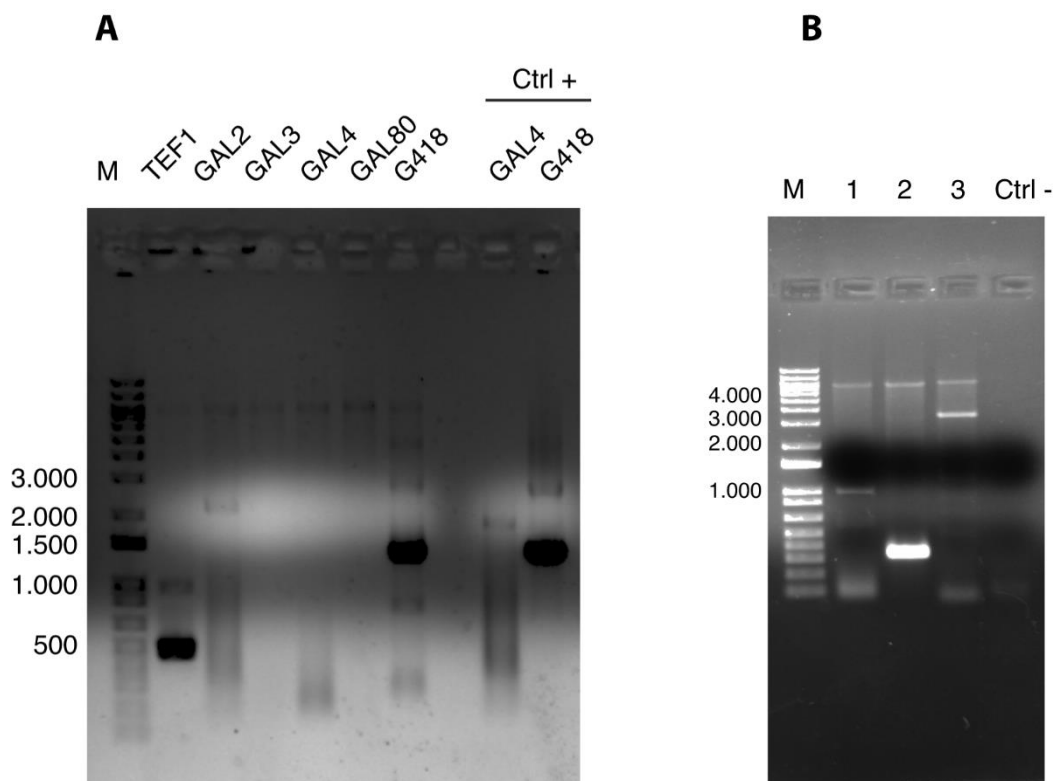
nova estratégia para montagem do cassete policistrônico utilizando o plasmídeo **pV2A-T**.



**Figura 13: Fragmentos de PCR amplificados com os oligonucleotídeos ampRegal após purificação.** A obtenção de material purificado foi realizada para aumento da eficiência de montagem do cassete policistrônico. M: marcador de peso molecular 1 kb plus DNA ladder (Thermo Scientific).

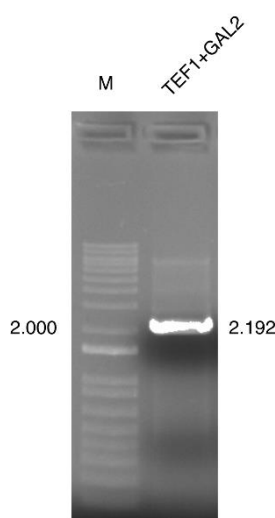


**Figura 14: Amplificação do vetor pPIC9m por PCR.** O material foi excisado do gel de agarose e purificado para utilização no processo de clonagem e montagem do cassete policistrônico. Tamanho esperado: 4.186 pb. M: marcador de peso molecular 1 kb dna ladder (Promega).



**Figura 15: Montagem do regulon GAL no vetor pPIC9m em *E. coli*.** Os fragmentos dos genes do cassete policistrônico foram transformados em *E. coli* juntamente do vetor pPIC9m para montagem *in vivo*. Foram feitas dois conjuntos de reações de PCR com

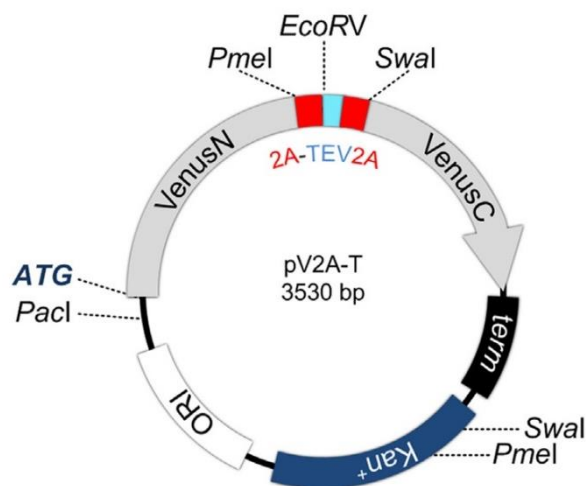
os oligonucleotídeos(ampRegal) utilizados para a clonagem e oligonucleotídeos(mid) que se ligam em duas ORFs consecutivas, para garantir a correta montagem *in tandem*. A) Amplicons obtidos da transformação de *E. coli* XL10 Gold com o regulon GAL e o vetor pPIC9m. Ctrl+: controle positivo utilizando DNA genômico de *S. cerevisiae* S288c como fita molde; B) Amplicons obtidos a partir da combinação de oligonucleotídeos mid, onde 1: TEF1-F + GAL2-R, 2: TEF1-F + G418-R e 3: GAL2-F + G418-R. M: marcador de peso molecular 1 kb plus dna ladder (Thermo Scientific) em ambas as figuras.



**Figura 16: Fragmento de PCR amplificado do vetor pPIC9Gal contendo o promotor  $P_{TEF1}$  e o gene *GAL2*.** Este fragmento foi corretamente montado *in vivo* e foi amplificado para a próxima rodada de transformação, diminuindo assim o número de fragmentos a serem clonados simultaneamente.

O plasmídeo pV2A-T (**Figura 17**) é um *shuttle vector*, ou seja, funcional em *E. coli* e *K. phaffii*, que permite a recombinação e clonagem de múltiplos fragmentos de DNA em sequência orientada. Seu primeiro uso foi descrito por HOEFGEN *et al.*, 2018 em *Aspergillus nidulans* onde foi inserido com sucesso toda a via biossintética do alcaloide psicotrópico psilocibina. Cada ORF é separada pela sequência 2A de picornavirus e uma sequência alvo da endopeptidase TEV. A estratégia demonstrada na **Figura 18** consiste na clonagem individual de cada fragmento de DNA a ser incorporado no cassete

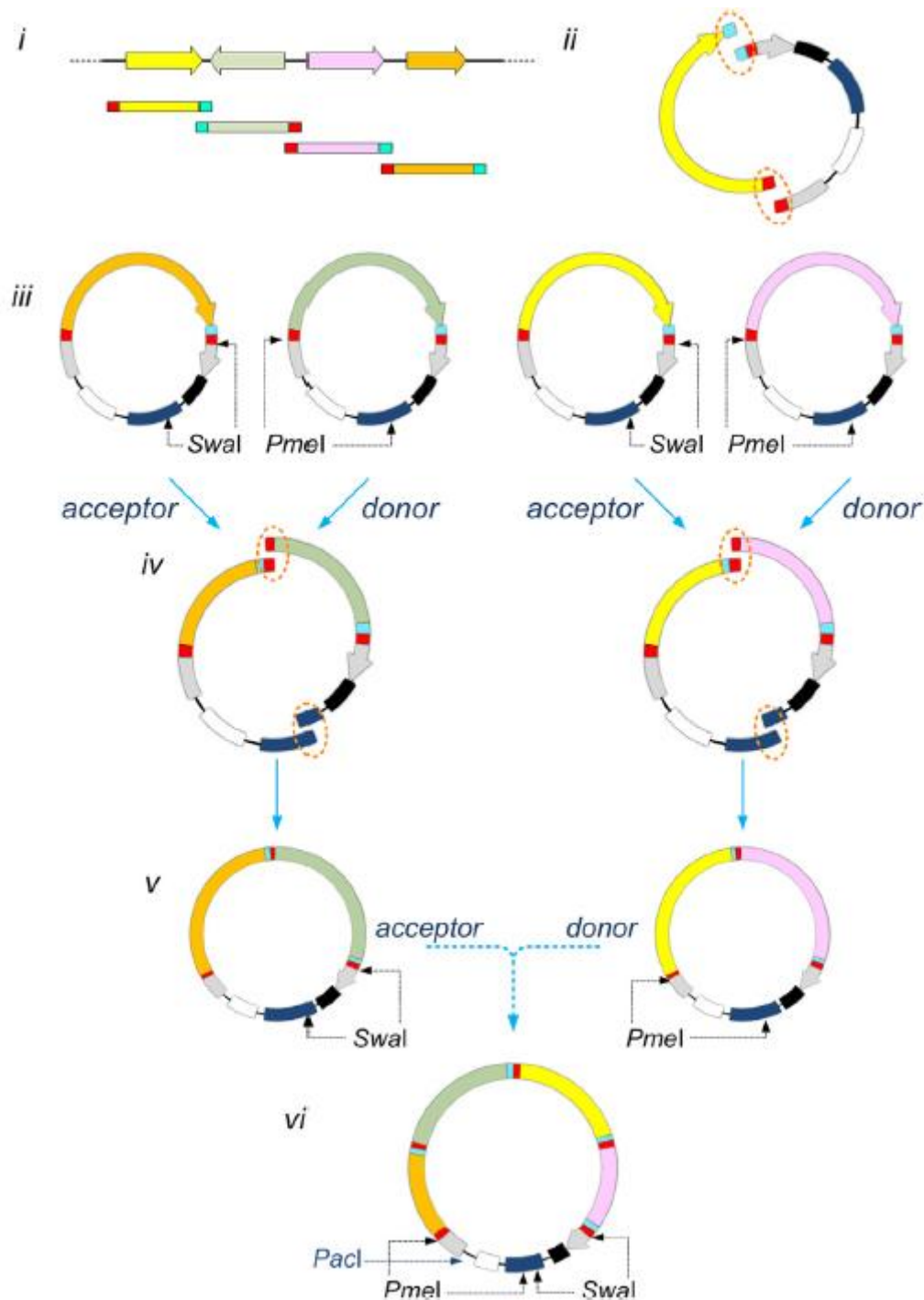
policistrônico no vetor inicial pV2A-T. Esta etapa pode ser realizada através de reações de digestão por endonucleases seguido de ligação ou via PCR e recombinação homóloga *in vivo*. Em seguida, faz-se a propagação e purificação do vetor para obtenção de massa necessária para as etapas subsequentes. A fim de unir dois fragmentos alvos em um único vetor, faz-se a digestão de um dos vetores (doador) com a endonuclease *Swal* e do outro vetor (receptor) com *PmeI*. Assim, criam-se duas extremidades homólogas entre os plasmídeos que são utilizadas para recombinação *in vivo*. O vetor resultante agora contém dois fragmentos *in tandem* separados pela sequência 2A. Este vetor pode ser reprocessado para adição de novos fragmentos de DNA utilizando-se a mesma lógica. Desta forma, o vetor resultante final contém todos os fragmentos de DNA inseridos na ordem desejada, formando um cassete policistrônico.



**Figura 17: Descrição visual do vetor pV2A-T.** O vetor contém a origem de replicação ORI do plasmídeo pUC19, gene modificado *Kan<sup>r</sup>* que confere resistência a kanamicina. O sítio de clonagem possui duas sequências 2A separadas pela região de reconhecimento da endopeptidase TEV. O gene repórter *Venus* está dividido em duas partes pelo sítio de clonagem. Retirado de HOEFGEN *et al.*, 2018. <https://doi.org/10.1016/j.ymben.2018.05.014>

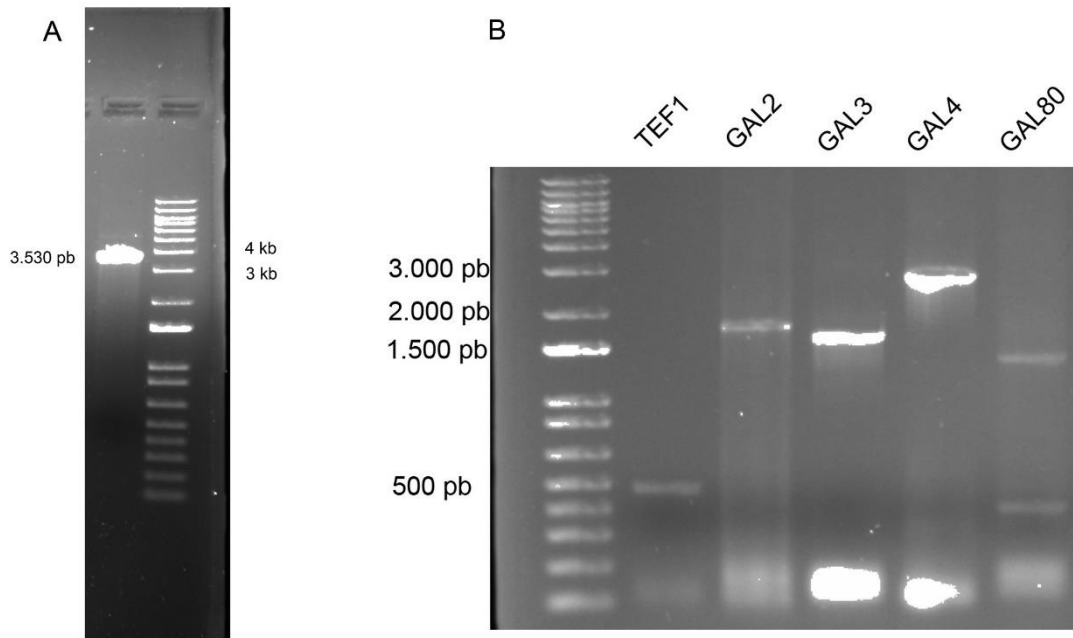


Desta forma, adotando-se esta estrutura de trabalho para a construção do regulon GAL com fim de aplicação em *K. phaffii*, foi feita a clonagem individual dos genes *GAL2*, *GAL3*, *GAL4* e *GAL80* no plasmídeo pV2A-T. Na **Figura 19** é demonstrada a amplificação via PCR dos fragmentos de interesse, utilizando oligonucleotídeos que modificam a sequência original com a retirada dos códon de parada e inserção das sequências homólogas ao vetor contendo os sítios de restrição necessários. Os fragmentos foram ligados ao vetor utilizando a técnica de clonagem por extensão circular com polimerase, onde os fragmentos são clonados no vetor através de uma reação de PCR em que ocorre o anelamento dos fragmentos de DNA e a consequente circularização do vetor contendo o inserto (QUAN e TIAN, 2011). Todas as etapas de clonagem foram verificadas por padrão de digestão dos plasmídeos com endonucleases assim como a presença e ordenação dos genes via PCR. Também foi realizada a recuperação do plasmídeo transformado na levedura através da extração de DNA total e este material foi transformado em *E. coli* para garantir que o produto final estava corretamente montado e funcional. O sequenciamento do vetor não foi realizado em tempo hábil, porém a correta transcrição do gene de resistência ao antibiótico garante que os fragmentos estão na fase de leitura correta. O mapa do vetor final e sua confirmação de construção encontram-se na **Figura 20**. O promotor *TEF1* foi inserido no vetor após linearização com a endonuclease *PacI* através de reação de ligação com a enzima T4 DNA ligase. Desta forma, espera-se que o regulon GAL seja transcrito de maneira constitutiva forte.

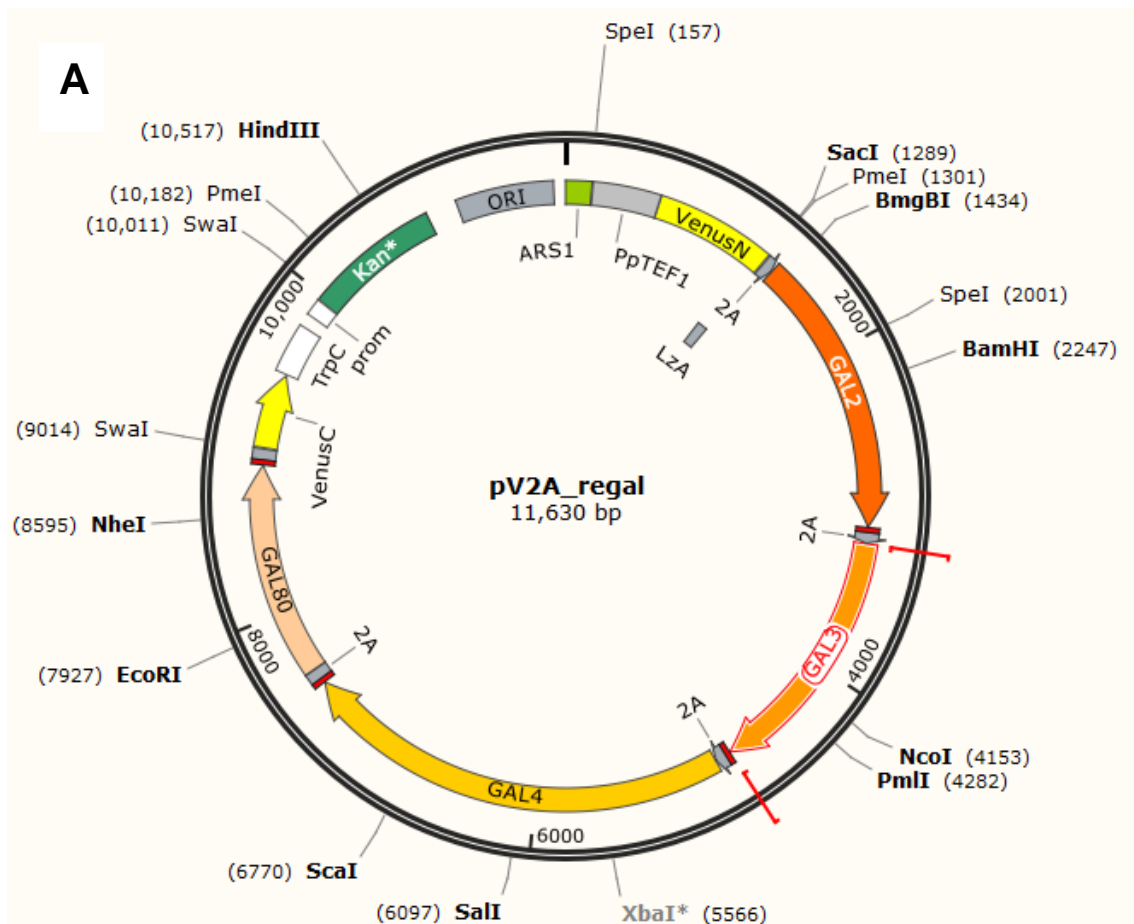


**Figura 18: Esquema de montagem do cassete policistrônico com auxílio do vetor pV2A-T.** i) Os genes de interesse são individualmente amplificados com oligonucleotídeos que contêm sequências homólogas ao vetor para recombinação orientada. ii) Vetor linearizado com *EcoRV*. iii) Os plasmídeos podem ser digeridos com as endonucleases *Swal* (plasmídeo acceptor) ou *PmeI* (plasmídeo doador). A escolha da enzima depende da ordem desejada dos fragmentos no cassete policistrônico. iv) *Swal* e *PmeI* criam terminais com sequências homólogas que permitem a remontagem do plasmídeo por recombinação. v) Após a fusão de dois genes, o vetor resultante recria

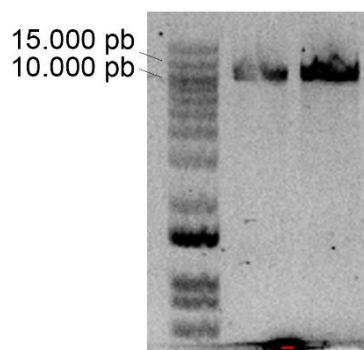
os sítios das endonucleases utilizadas, permitindo a adição de novos fragmentos. vi) O vetor final resultante possui todos os fragmentos de DNA em ordenados e separados pela sequência 2A. Os fragmentos devem ser modificados para retirar os códons de parada para permitir que o cassete policistrônico funcione corretamente. O sítio de *PacI* permite a adição de novos elementos ao plasmídeo antes do cassete, como promotores e sequências de integração genômica. O correto funcionamento do plasmídeo também pode ser verificado pela presença da proteína funcional YFP na célula transformada. Retirado de HOEFGEN *et al.*, 2018. <https://doi.org/10.1016/j.ymben.2018.05.014>



**Figura 19: Eletroforese em gel de agarose dos fragmentos de DNA utilizados para construção do regulon GAL no plasmídeo pV2A-T.** A) Vetor pV2A-T linearizado com a endonuclease *EcoRV*. B) Fragmentos de DNA amplificados para clonagem individual no plasmídeo pV2A-T. Tamanho esperado dos amplicons: Promotor TEF1: 407 pb. GAL2: 1.722 pb. GAL3: 1.560 pb. GAL4: 2.643 pb. GAL80: 1.305 pb.



**B**

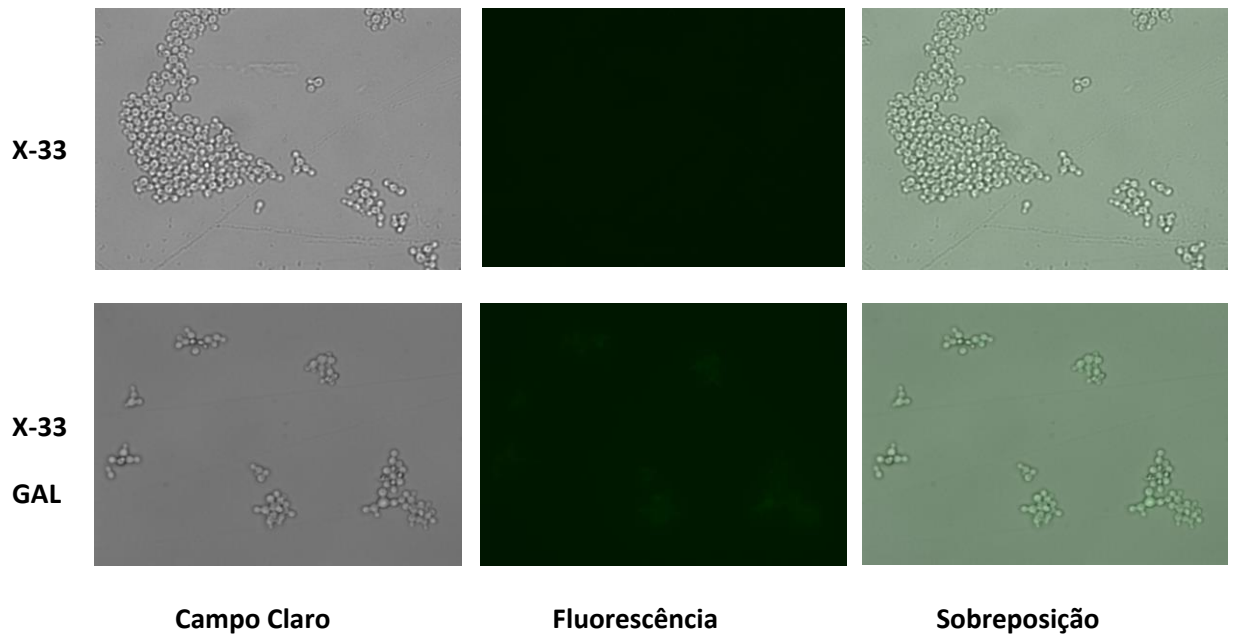


**Figura 20: Confirmação da construção do vetor pV2A\_regal contendo o cassete policitrônico para expressão do regulon GAL em *K. phaffii*.** A) Mapa do vetor construído. B) Digestão do vetor pV2A\_regal com *PacI* extraídos de duas colônias transformantes. Tamanho esperado de 11.000 pb.

Com o intuito de testar o funcionamento do sistema de expressão gênica regulado por galactose em *K. phaffii*, os plasmídeos pV2A\_regal e pPICH-

PGal1GFP foram cotransformados na cepa X33 por eletroporação. Desta forma, espera-se que na presença de galactose ocorra a produção da proteína *eGFP* que é capaz de emitir fluorescência a 507nm.

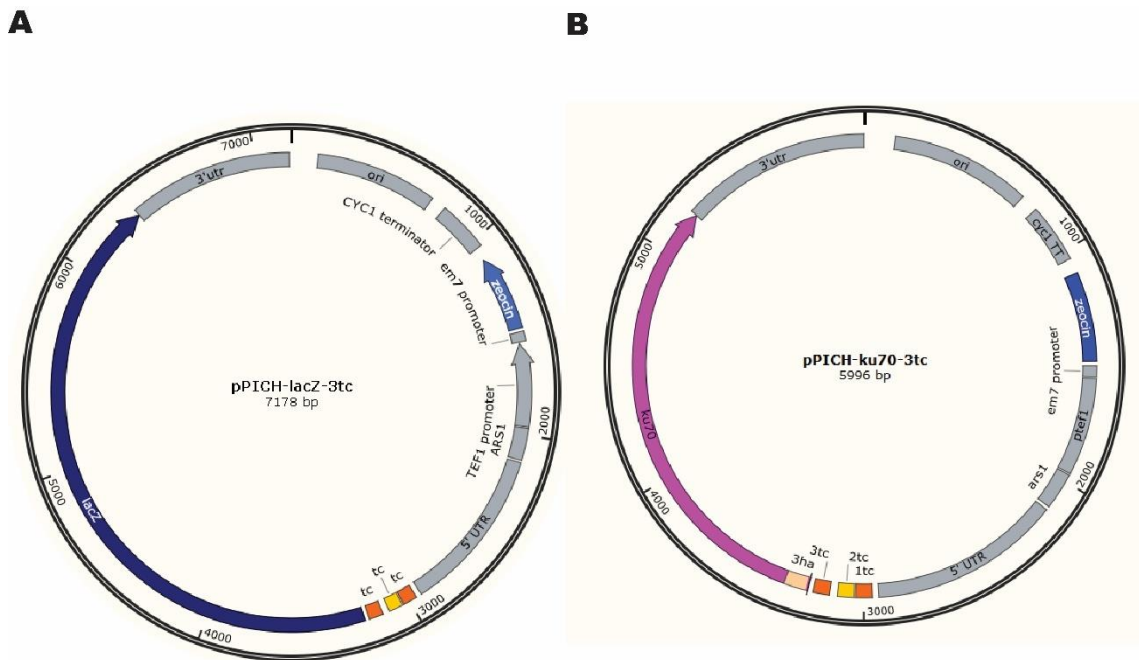
Na **Figura 21** é possível observar que não houve sinal de fluorescência significativo quando excluída a autofluorescência das células ou o sinal de fundo. Existem várias hipóteses levantadas pelo grupo de trabalho para explicar o resultado obtido, como: falta de expressão balanceada dos genes do regulon GAL, mau funcionamento do sistema de expressão de eGFP controlado pelo promotor  $P_{GAL1}$ , falha na sinalização do sistema pela galactose e problemas na separação das proteínas pelo sistema 2A após a tradução do mRNA policistrônico ou até mesmo ao baixo número de cópias do plasmídeo. Devido à grande complexidade do sistema construído pelo número de genes envolvidos e pela dificuldade de se obter transformantes positivos, decidiu-se desenvolver uma cepa de *K. phaffii* laboratorial com características de recombinação homóloga mais pronunciadas para facilitar o desenvolvimento posterior do sistema do regulon GAL realizando a integração genômica dos fragmentos de DNA componentes.



**Figura 21: Prova de conceito do funcionamento do regulon GAL em *K. phaffii*.** As células foram crescidas em meio mínimo YNB com sorbitol 1% (p/v) e galactose 2% (p/v) durante 24 horas. As imagens foram obtidas com 40x de aumento. A excitação ocorreu a 488nm e a emissão foi obtida a 507nm. Não foi obtido sinal suficientemente significativo da proteína fluorescente *eGFP* durante os testes, sendo necessário reavaliar o sistema de expressão policistrônico regulado por galctose.

## 5.2 Desenvolvimento da linhagem de *K. phaffii* responsiva à tetraciclina

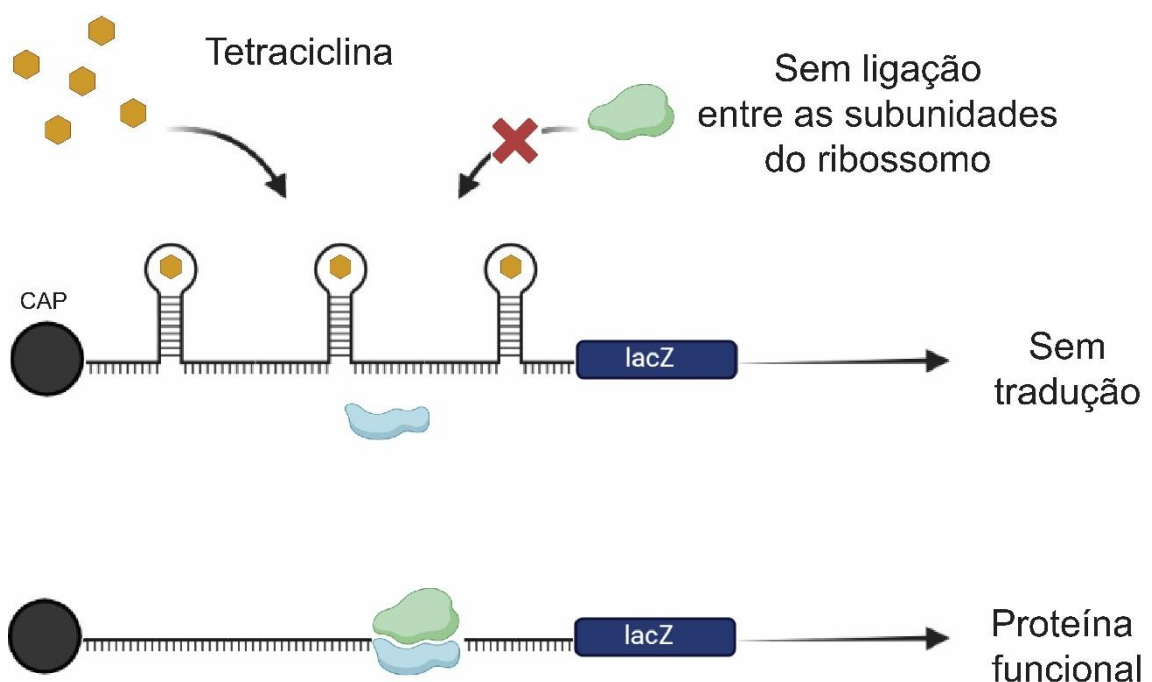
A fim de desenvolver um sistema de expressão baseado no aptâmero regulado por tetraciclina, foi utilizado o vetor epissomal pPICH como arcabouço para clonagem dos genes *KU70* e *lacZ* sob controle do *riboswitch*. Desta forma, poderíamos garantir a estabilidade e múltiplas cópias, eliminando a possibilidade de um resultado falsamente positivo devido à baixa expressão do gene.



**Figura 22: Mapa dos plasmídeos contendo o aptâmero de tetraciclina (tc) controlando a produção de proteínas.** A) pPICH-lacZ-3tc vetor episomal contendo o gene *lacZ* sob controle de 3 sequências repetitivas do aptâmero tc. B) pPICH-ku60-3tc vetor episomal contendo o gene *KU70* sob controle do aptâmero tc. Todas as sequências foram amplificadas por PCR e clonadas no vetor pPICH por recombinação homóloga ou por PCR.

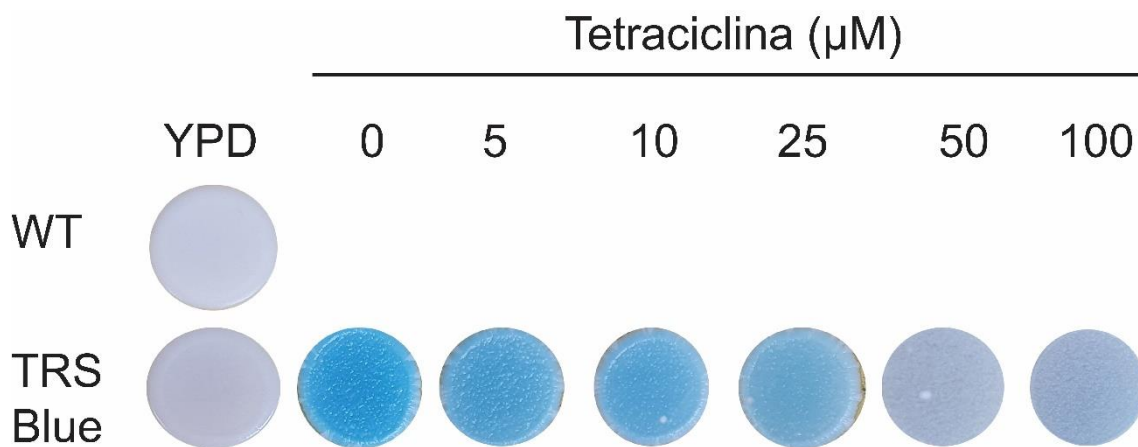
A prova de conceito do funcionamento do sistema foi iniciada com a expressão da enzima  $\beta$ -galactosidase codificada pelo gene *lacZ* controlado pelo *riboswitch* no plasmídeo pPICH-lacZ-3tc (Figura 21A). O sistema foi testado na presença e na ausência de tetraciclina. De acordo com a estratégia utilizada, a presença do *riboswitch* não influencia na expressão do mRNA do gene, mas apenas em sua tradução. Ou seja, desta forma, células de *K. phaffii* contendo o sistema de controle de expressão proteica que fossem expostas à presença do composto X-Gal e apresentassem atividade enzimática satisfatória de  $\beta$ -galactosidase devem ganhar a coloração azul que se intensifica de acordo com o tempo de exposição. Na presença de tetraciclina, espera-se que o *riboswitch*

assuma uma conformação tridimensional que impede o acoplamento correto das duas subunidades ribossomais e as células mantenham a coloração branca devido à reduzida ou ausente produção da enzima (Figura 23). Como demonstrado na Figura 24, as colônias semeadas em YPD contendo X-Gal expostas a concentrações de tetraciclina iguais ou superiores a 50  $\mu\text{M}$  apresentaram redução na intensidade da cor azul, compatível com o resultado esperado em caso de funcionamento correto do *riboswitch*. Este resultado também é corroborado pelo trabalho publicado por HANSON *et al.*, 2003, onde o aptâmero mostrou a capacidade de redução da produção de luciferase a partir das mesmas concentrações de tetraciclina.



**Figura 23: Mecanismo de funcionamento do *riboswitch* durante a tradução do mRNA.** Na ausência de tetraciclina, o ribossomo se liga ao mRNA e o traduz a uma proteína. Na presença de tetraciclina, o aptâmero assume uma estrutura tridimensional que impede a montagem correta do ribossomo e reduz a produção da proteína.

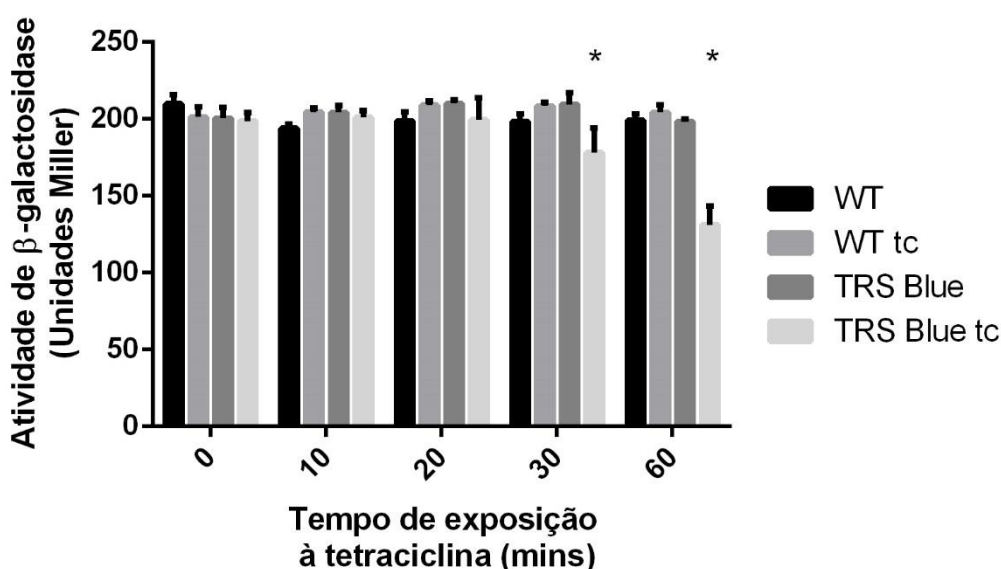




**Figura 24: O aptâmero de tetraciclina reduz a tradução de  $\beta$ -galactosidase em *K. phaffii*.** Células transformadas com o vetor pPICH-lacZ-3tc foram semeadas em meio YPD contendo X-GAL por 3 dias a 30 °C seguido de 4 dias a 4 °C para intensificar a coloração azul. Concentrações crescentes de tetraciclina mostram uma resposta inibitória da tradução dose-dependente, prevenindo a metabolização de X-GAL e diminuindo a intensidade da cor azul em concentrações maiores do antibiótico. Esta figura é representativa de três experimentos independentes que continham triplicatas de cada colônia em cada placa.

Uma vez que o experimento em placa possui um tempo prolongado de incubação para que ocorra o aparecimento da coloração azul, foi necessário avaliar a resposta celular em um período menor de tempo. Para isso, a atividade de  $\beta$ -galactosidase foi mensurada entre 10 e 60 minutos de exposição a 100  $\mu\text{M}$  de tetraciclina através do ensaio utilizando ONPG. O resultado mostrou uma redução de aproximadamente 34% na atividade enzimática após 60 minutos de exposição (Figura 25). Sendo assim, foi comprovado o funcionamento do sistema de controle de expressão proteica baseado no *riboswitch* regulado por tetraciclina. Considerando estes resultados, caso as colônias expostas à tetraciclina que se tornaram brancas sejam retiradas da presença do estímulo, a

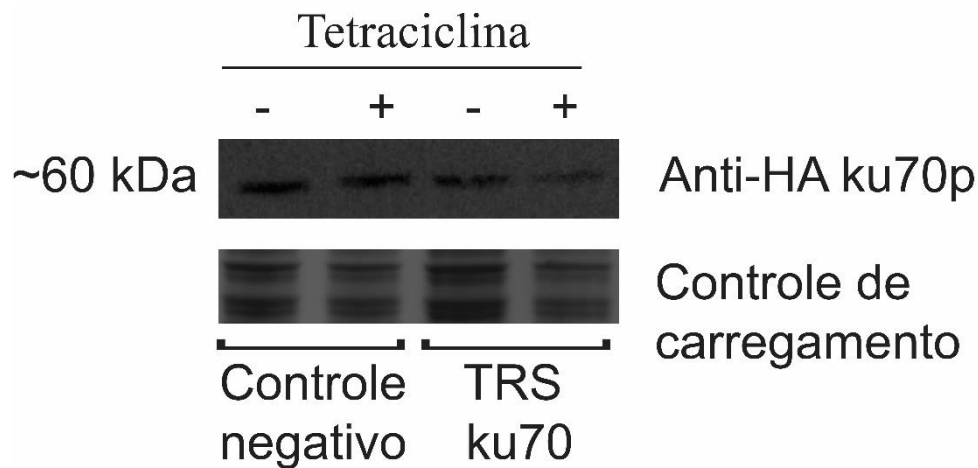
tradução da enzima retornará e as células adquirirão a coloração azul rapidamente, uma vez que o mRNA já produzido seria imediatamente traduzido. Um ponto de atenção para futuras aplicações desta tecnologia está na necessidade de adaptar o tempo de exposição e a concentração de tetraciclina para cada proteína recombinante a ser produzida, pois é sabido que a enzima  $\beta$ -galactosidase possui uma meia-vida curta na célula e, portanto, o resultado observado aqui necessita de adaptações para cada proteína.



**Figura 25: Atividade de  $\beta$ -galactosidase é reduzida após exposição da célula a tetraciclina.** A cepa TRS Blue mostrou uma redução da atividade enzimática específica de  $\beta$ -galactosidase após 30 minutos de exposição à tetraciclina em meio YPD líquido. \*O teste T não pareado mostra que há uma diferença significativa entre os dados de 30 e 60 mins da cepa TRS Blue na presença de tetraciclina em relação aos outros grupos testados em um intervalo de confiança de 95%. n = 3.

Após demonstrada a prova de conceito do sistema, este trabalho focou seus esforços em demonstrar o funcionamento do *riboswitch* em outra proteína, a fim de avaliar se o resultado observado seria similar. Para isso, utilizamos a proteína

Ku70p, envolvida no reparo da quebra da dupla fita de DNA. Para isso, foi necessária fazer a deleção do gene *KU70* na cepa X-33 de *K. phaffii*. Em seu lugar foi adicionado o gene *amdS*, demonstrado por PIVA *et al.*, 2018 ser uma efetiva marca de contra seleção quando combinado com o composto Acetamida. A nova cepa  $\Delta ku70::amdS$  de X-33 foi transformada com o plasmídeo pPICH-ku70-3tc contendo o gene modificado com a sequência HA (Figura 22B). Esta etapa de marcação do N-terminal da proteína foi necessária pois a deleção do gene *KU70* não gera nenhum fenótipo visual e desta forma, podemos verificar a redução da tradução da proteína por Western Blotting. Após a exposição das células a 250  $\mu$ M de tetraciclina por 8 horas em meio de cultivo líquido YPD, foi observada a redução da concentração de Ku70p em aproximadamente 40% após análise de densitometria da imagem (Figura 26). Portanto, foi demonstrado que a presença de tetraciclina é capaz de ativar o *riboswitch* e consequentemente reduziu a tradução da proteína. Ademais, é importante destacar a reversibilidade do sistema, característica que torna a utilização do *roboswitch* vantajosa. Ao se retirar as células da presença da tetraciclina, a célula pode retomar a tradução dos mRNA disponíveis imediatamente, fazendo com que a célula tenha uma resposta mais célere.



**Figura 26: Redução da tradução de Ku70p após ativação do *riboswitch* de tetraciclina.** A cepa TRS *ku70* foi exposta a 250  $\mu$ M de tetraciclina por 8h e foi observada uma redução de 40% na quantidade de proteína expressa. Tamanho estimado da proteína Ku70p = ~60 kDa.

O uso do *riboswitch* de tetraciclina se mostrou uma estratégia efetiva e de baixo custo que pode ser aplicado à virtualmente qualquer proteína. Posto que o sistema permite o controle preciso da expressão proteica no nível traducional, genes que estejam sendo controlado por promotores fracos ou fortes podem ser controlados da mesma maneira, pois a força de restrição da tradução pode ser ajustada com o aumento ou redução do número de sítios de ligação a tetraciclina do *riboswitch* (HANSON 2003; SUESS 2003; BERENS 2015; RUSCITO 2016). Isto significa que menos manipulações genéticas são necessárias para estudos de produção de proteínas recombinantes, assim como a utilização de diversas combinações de promotores e terminadores no caso da inserção de múltiplos genes. Em especial, no caso do gene *KU70*, sua deleção pode gerar efeitos deletérios na célula (SCHWARZHANS 2016).

## 6 Conclusões e perspectivas

A montagem do regulon GAL em *K. phaffii* se mostrou tecnicamente desafiadora e se faz necessária a reavaliação do processo de montagem e integração do cassete no genoma. Como prova de conceito, pode-se fazer a expressão constitutiva do gene *GAL4* para se testar a capacidade de ativação do promotor  $P_{GAL1}$  controlando o gene repórter *lacZ*. Desta forma, pode-se mensurar o desempenho deste sistema em condições ótimas, seguido da comparação do desempenho do sistema com todas as proteínas controladoras.

Além disso, este trabalho conseguiu demonstrar que o uso de aptâmeros para o controle da produção de proteína em *K. phaffii* é viável e representa uma contribuição significativa para o desenvolvimento de novos estudos de aplicação deste microrganismo como plataforma de produção de produtos biológicos de alto valor agregado. A cepa TRS *KU70* produzida neste trabalho, que contém o gene *KU70* sob controle do *riboswitch* de tetraciclina, poderá ser testada para o aumento da recombinação homóloga de fragmentos exógenos de DNA quando na presença de tetraciclina, facilitando o desenvolvimento de novos estudos de manipulação genética desta levedura e potencializando o uso de *K. phaffii* na indústria.

## 7 Referências bibliográficas

ACHARYA, T., KENNEDY, R., DAAR, AS., SINGER, PA. **Biotechnology to improve health in developing countries – a review**. Mem Inst Oswaldo Cruz 99(4), p. 341-350. 2004. doi: 10.1590/s0074-02762004000400001.

AHMAD, M., HIRZ, M., PICHLER, M., SCHWAB, H. **Protein expression in *Pichia pastoris*: Recent achievements and perspectives for heterologous protein production**. *Applied Microbiology and Biotechnology*, v. 98, nº 12, p. 5301–5317, 2014. ISBN: 10.1007%2Fs00253-014-5732-5, ISSN: 14320614,

DOI: 10.1007/s00253-014-5732-5.

ATA, O., ERGÜN, B., FICKER, P., HEISTINGER, L., MATTANOVICH, D., REBNEGGER, C., GASSER, B. **What makes *Komagataella phaffii* non-conventional?** *FEMS Yeast Research*. 2021. Vol 21. DOI: 10.1093/femsyr/foab059

ARAUJO, J.A., FERREIRA, T.C., RUBINI, M.R.M DURAN, A.G.G., MARCO, J.L., MORAES, L.M.P, TORRES, F.A.G. **Coexpression of cellulases in *Pichia pastoris* as a self - processing protein fusion.** *AMB Express*, 2015. ISBN: 1356801501, ISSN: 2191-0855, DOI: 10.1186/s13568-015-0170-z.

ARAYA-GARAY, J. M., FEIJOO-SIOTA, L., VEIGA-CRESPO, P., VILLA, T.G. **Construction of new *Pichia pastoris* X-33 strains for production of lycopene and  $\beta$ -carotene.** *Applied Microbiology and Biotechnology*, v. 93, n° 6, p. 2483–2492, 2012. ISBN: 0175-7598, ISSN: 01757598, DOI: 10.1007/s00253-011-3764-7.

BAGHBAN, R., FARAJNIA, S., RAJABIBAZL, M., GHASEMI, Y., MAFI, A.A., HOSEINPOOR, R., RAHBARNIA, L., ARIA, M. **Yeast Expression Systems: Overview and Recent Advances.** *Molecular Biotechnology*, v. 0, n° 0, p. 0, 2019. ISBN: 1203301900164, ISSN: 1073-6085, DOI: 10.1007/s12033-019-00164-8.

BETANCUR, M.O., REIS, V.C.B., NICOLA, A.M., DEMARCO, J.L., MORAES, L.M.P., TORRES, F.A.G. **Multicopy integration in *Komagataella phaffii* mediated by a defective auxotrophic marker.** *Microbial cell factories*, v. 16, n° 99, 2017. DOI: 10.1186/s12934-017-0715-8 .

CAMATTARI, A., GOH, A., YIP, Y.L., TAN, A.H.M., NG, S.W., TRAN, A., LIU, G., LIACHKO, I., DUNHAM, M.J., RANCATI, G. **Characterization of a panARS-based episomal vector in the methylotrophic yeast *Pichia pastoris* for recombinant protein production and synthetic biology applications.** *Microbial Cell Factories*, v. 15, n° 1, p. 1–11, 2016. ISBN: 1475-2859, ISSN: 14752859, DOI: 10.1186/s12934-016-0540-5.

ÇELİK, E.; ÇALIK, P. **Production of recombinant proteins by yeast cells.**

*Biotechnology Advances*, v. 30, n° 5, p. 1108–1118, 2012. ISBN: 1873-1899 (Electronic)r0734-9750 (Linking), ISSN: 07349750, DOI: 10.1016/j.biotechadv.2011.09.011.

CEREGHINO, J. L.; CREGG, J. M. **Heterologous protein expression in the methylotrophic yeast *Pichia pastoris***. *FEMS Microbiology Reviews*, v. 24, n° 1, p. 45–66, 2000. ISBN: 0168-6445 (Print)n0168-6445 (Linking), ISSN: 1574-6976, DOI: 10.1111/j.1574-6976.2000.tb00532.x.

CHURCH, G. M., ELOWITZ, M.B., SMOLKE, C.D., VOIGT, C.A., WEISS, R. **Realizing the potential of synthetic biology**. *Nature Reviews Molecular Cell Biology*, v. 15, n° 4, p. 289–294, 2014. ISBN: 1471-0080 (Electronic)r1471-0072 (Linking), ISSN: 14710080, DOI: 10.1038/nrm3767.

CLARKE, L., KITNEY, R. **Developing synthetic biology for industrial biotechnology applications**. *Biochemical Society Transactions* 48, p. 113-122. Doi: 10.1042/BST20190349 .

DEROSA, M.C., RUSCITO, A. **Small-Molecule Binding Aptamers: Selection strategies, Characterization and Applications**. *Frontiers in Chemistry*, v. 10. 2016. DOI: 10.3389/fchem.2016.00014.

DESCHUTTER, K., LIN, Y-C., TIELS, P., HECKE, A.V., GLINKA, S., WEBER-LEHMANN, J., ROUZÉ, P., DE PEER, Y.V., CALLEWAERT, N. **Genome sequence of the recombinant protein production host *Pichia pastoris***. *Nature Biotechnology*, v. 27, n° 6, p. 561–566, 2009. ISBN: 1546-1696, ISSN: 10870156, DOI: 10.1038/nbt.1544.

D'ESPAUX, L., MENDEZ-PEREZ, D., LI, R., KEASLING, J.D. **Synthetic biology for microbial production of lipid-based biofuels**. *Current Opinion in Chemical Biology*, v. 29, p. 58–65, 2015. ISBN: 1879-0402 (Electronic)r1367-5931 (Linking), ISSN: 18790402, DOI: 10.1016/j.cbpa.2015.09.009.

DHAMANKAR, H.; PRATHER, K. L. J. **Microbial chemical factories: Recent advances in pathway engineering for synthesis of value-added chemicals**. *Current Opinion in Structural Biology*, v. 21, n° 4, p. 488–494, 2011. ISBN: 0959-440X, ISSN: 0959440X, DOI: 10.1016/j.sbi.2011.05.001.

DICIONÁRIO OXFORD ESCOLAR PARA ESTUDANTES BRASILEIROS DE INGLÊS. 2018. Washington, Ed. Oxford University Press.

DROCOURT, D., CALMELS, T., REYNES, J.P., BARON, M., TIRABY, G. **Cassettes of the *Streptoalloteichus hindustanus* ble gene for transformation of lower and higher eukaryotes to phleomycin resistance.** *Nucleic Acids Research*, v. 18, nº 13, p. 4009, 1990. DOI: 10.1093/nar/18.13.4009

DU, M.; BATTLES, M. B.; NETT, J. H. **A color-based stable multi-copy integrant selection system for *Pichia pastoris* using the attenuated ADE1 and ADE2 genes as auxotrophic markers.** *Bioengineered Bugs*, v. 3, nº 1, p. 32–37, 2012. ISBN: 1949-1026 (Electronic)r1949-1018 (Linking), ISSN: 19491018, DOI: 10.4161/bbug.3.1.17936.

GALANIE, S., THODEY, K., TRENCHARD, I.J., INTERRANTE, M.F., SMOLKE, C.D. **Biosynthesis of opioids in yeast.** *Science*, v. 349, nº 6252, p. 1095–1100, 2015. DOI: 10.1126/science.aac9373.Complete.

GEDVILAITE, A.; SASNAUSKAS, K. **Control of the expression of the ADE2 gene of the yeast *Saccharomyces cerevisiae*.** *Current Genetics*, v. 25, nº 6, p. 475–479, 1994. ISBN: 0172-8083 (Print)r0172-8083 (Linking), ISSN: 01728083, DOI: 10.1007/BF00351665.

GEIER, M., FAULAND, P., VOGL, T., GLIEDER, A. **Compact multi-enzyme pathways in *P. pastoris*.** *Chemical Communications*, v. 51, p. 1643–1646, 2015a. ISSN: 1359-7345, DOI: 10.1039/C4CC08502G.

GOMES, A.M.V., CARMO, T.S., CARVALHO, L.S., BAHIA, F.M., PARACHIN, N.S. **Comparison of Yeasts as Hosts for Recombinant Protein Production.** *Microorganisms*, v. 6, nº 2, p. 38, 2018. ISBN: 2076-2607, ISSN: 2076-2607, DOI: 10.3390/microorganisms6020038.

HANSON, S., BERTHELOT, K., FINK, B., MCCARTHY, J.E.G., SUESS, B. **Tetracycline-aptamer-mediated translations regulation in yeast.** *Molecular Microbiology*, 2003, v. 49, no 6, pgs. 1627-1637. DOI: 10.1046/j.1365-2958.2003.03656.x.



HARTNER, F. S. *et al.* **Promoter library designed for fine-tuned gene expression in *Pichia pastoris*.** *Nucleic Acids Research*, v. 36, n° 12, p. 1–15, 2008. ISBN: 0305-1048, ISSN: 03051048, DOI: 10.1093/nar/gkn369.

HE, J. DENG, J., ZHENG, Y., GU, J. **A synergistic effect on the production of S-adenosyl-L-methionine in *Pichia pastoris* by knocking in of S-adenosyl-L-methionine synthase and knocking out of cystathionine- $\beta$  synthase.** *Journal of Biotechnology*, v. 126, n° 4, p. 519–527, 2006. ISBN: 0168-1656, ISSN: 01681656, DOI: 10.1016/j.jbiotec.2006.05.009.

HOEFGEN, S., LIN, J., FRICKE, J., STROE, M.C., MATTERN, D.J., KUFS, J.E., HORTSCHANSKY, P., BRAKHAGE, A.A., HOFFMEISTER, D., VALIANTE, V. **Facile assembly and fluorescence-based screening method for heterologous expression of polycistronic transcripts in eukaryotes.** *Metabolic Engineering*, p.2018. ISSN: 1096-7176, DOI: 10.1016/j.ymben.2018.05.014.

JACOBUS, A. P.; GROSS, J. **Optimal Cloning of PCR Fragments by Homologous Recombination in *Escherichia coli*.** p. 1–17, 2015. DOI: 10.1371/journal.pone.0119221.

JENSEN, M. K.; KEASLING, J. D. **Recent applications of synthetic biology tools for yeast metabolic engineering.** *FEMS Yeast Research*, v. 15, n° 1, p. 1–10, 2015. ISBN: 15671356, ISSN: 15671364, DOI: 10.1111/1567-1364.12185.

KANG, T.; MARTINS, T.; SADOWSKI, I. **Wild type GAL4 binds cooperatively to the GAL1-10 UASG in vitro.** *Journal of Biological Chemistry*, v. 268, n° 13, p. 9629–9635, 1993. ISBN: 6048224524, ISSN: 00219258, DOI: 10.1006/jipa.1996.0101.

KAVITA, K., BREAKER, RR. **Discovering riboswitches: the past and the future.** *Trends in Biochemical Sciences* 48 (2), p. 119-141. DOI: 10.1016/j.tibs.2022.08.009.

KIM, B., DU, J., ERIKSEN, D.T., ZHAO, H. **Combinatorial Design of a Highly Efficient Xylose-Utilizing Pathway in *Saccharomyces cerevisiae* for the production of cellulosic biofuels.** *Applied and Environmental Microbiology*, v.

79, n° 3, p. 931–941, 2013. ISSN: 0099-2240, DOI: 10.1128/AEM.02736-12.

KRUCKEBERG, A. L. **The hexose transporter family of *Saccharomyces cerevisiae***. *Archives of Microbiology*, v. 166, n° 5, p. 283–292, 1996. ISBN: 0302-8933 (Print)r0302-8933 (Linking), ISSN: 03028933, DOI: 10.1007/s002030050385.

KÜBERL, A., SCHNEIDER, J., THALLINGER, G.G., ANDERL, I., WIBBERG, D., HAJEK, T., JAENICKE, S., BRINKROLF, K., GOESMANN, A., SZCZEPANOWSKI, R., PÜHLER, A., SCHWAB, H., GLIEDER, A., PICHLER, H. **High-quality genome sequence of *Pichia pastoris* CBS7435**. *Journal of Biotechnology*, v. 154, n° 4, p. 312–320, 2011. ISBN: 0168-1656, ISSN: 01681656, DOI: 10.1016/j.jbiotec.2011.04.014.

KURTZMAN, C. P. **Biotechnological strains of *Komagataella (Pichia) pastoris* are *Komagataella phaffii* as determined from multigene sequence analysis**. *Journal of Industrial Microbiology and Biotechnology*, v. 36, n° 11, p. 1435–1438, 2009. ISBN: 13675435 (ISSN), ISSN: 13675435, DOI: 10.1007/s10295-009-0638-4.

LI, P., ANUMANTHAN, A., GAO, X-G., ILANGO VAN, K., SUZARA, V.V., DÜZGÜNES, N., RENUGOPALAKRISHAM, V. **Expression of recombinant proteins in *Pichia pastoris***. *Applied Biochemistry and Biotechnology*, v. 142, n° 2, p. 105–124, 2007. ISBN: 0273-2289 (Print)r0273-2289 (Linking), ISSN: 02732289, DOI: 10.1007/s12010-007-0003-x.

LIN-CEREGHINO, G. P., CEREGHINO, J.L., SUNGA, A.J., JOHNSON, M.A., LIM, M., GLESSON, M.A.G., CREGG, J.M. **New selectable marker/auxotrophic host strain combinations for molecular genetic manipulation of *Pichia pastoris***. *Gene*, v. 263, p. 159–169, 2001. DOI: 10.1016/S0378-1119(00)00576-X.

LIN-CEREGHINO, J., HASHIMOTO, M.D., MOY, A., CASTELO, J., ORAZEM, C.C., KUO, P., XIONG, S., GANDHI, V., HATAE, C.T., CHAN, A., LIN-CEREGHINO, G.P. **Direct selection of *Pichia pastoris* expression strains using new G418 resistance vectors**. *Yeast*, v. 25, n° 4, p. 293–299, 2008. ISSN: 0749503X, DOI: 10.1002/yea.1587.

LOHR, D., VENKOV, P., ZLATANOVA, J. **Transcriptional regulation in the yeast gal gene family: a complex genetic network.** *The FASEB Journal*, v. 9, n° 9, p. 777–787, 2016. DOI: 10.1096/fasebj.9.9.7601342

MACAULEY-PATRICK, S., FAZENDA, M.L., MCNEILL, B., HARVEY, L.M. **Heterologous protein production using the *Pichia pastoris* expression system.** *Yeast*, v. 22, n° 4, p. 249–270, 2005. ISBN: 0749-503X, ISSN: 0749503X, DOI: 10.1002/yea.1208.

MARX, H.; MATTANOVICH, D.; SAUER, M. **Overexpression of the riboflavin biosynthetic pathway in *Pichia pastoris*.** *Microbial Cell Factories*, v. 7, p. 1–11, 2008. ISBN: 1475-2859, ISSN: 14752859, DOI: 10.1186/1475-2859-7-23.

MATTANOVICH, D., GRAF, A., STADLMANN, J., DRAGOSITS, M., REDL, A., MAURER, M., KLEINHEINZ, M., SAUER, M., ALTMANN, F., GASSER, B. **Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*.** *Microbial Cell Factories*, v. 8, p. 1–13, 2009. ISBN: 1475-2859, ISSN: 14752859, DOI: 10.1186/1475-2859-8-29.

MCKEAGUE, M., MCCONNELL, E.M., CRUZ-TOLEDO, J., BERNARD, E.D., PACH, A., MASTRONARDI, E., ZHANG, X., BEKING, M., FRANCIS, T., GIAMBERARDINO, A., CABECINHA, A., RUSCITO, A., ARANDA-RODRIGUEZ, R., DUMONTIER, M., DEROSA, M.C. **Analysis of *in vitro* aptamer selection parameters.** *Journal of Molecular Evolution*, n° 81, p. 150-161. DOI:10.1007/S00239-015-9708-6

NÄÄTSAARI, L., MISTLBERGER, B., RUTH, C., HAJEK, T., HARTNER, F.S., GLIEDER, A. **Deletion of the *Pichia pastoris* ku70 homologue facilitates platform strain generation for gene expression and synthetic biology.** *PLoS ONE*, v. 7, n° 6, 2012. ISBN: 1932-6203, ISSN: 19326203, DOI: 10.1371/journal.pone.0039720.

NETT, J. H., HODEL, N., RAUSCH, S., WILDT, S. **Cloning and disruption of the *Pichia pastoris* ARG1, ARG2, ARG3, HIS1, HIS2, HIS5, HIS6 genes and their use as auxotrophic markers.** *Yeast*, v. 22, n° 4, p. 295–304, 2005. ISBN: 0749-503X, ISSN: 0749503X, DOI: 10.1002/yea.1202.

NETT, J. H.; GERNGROSS, T. U. **Cloning and disruption of the PpURA5 gene and construction of a set of integration vectors for the stable genetic modification of *Pichia pastoris***. *Yeast*, v. 20, n° 15, p. 1279–1290, 2003. ISBN: 0749-503X, ISSN: 0749503X, DOI: 10.1002/yea.1049.

NIELSEN, A. A. K.; SEGALL-SHAPIRO, T. H.; VOIGT, C. A. **Advances in genetic circuit design: Novel biochemistries, deep part mining, and precision gene expression**. *Current Opinion in Chemical Biology*, v. 17, n° 6, p. 878–892, 2013. ISBN: 1367-5931, ISSN: 13675931, DOI: 10.1016/j.cbpa.2013.10.003.

PANNALA, V. R., BHAT, P.J., BHARTIYA, S., VENKATESH, K.V. **Systems biology of GAL regulon in *Saccharomyces cerevisiae***. *Wiley Interdisciplinary Reviews*, n°2(1), p. 98-106, 2010. DOI: 10.1002/wsbm.38.

PENG, G.; HOPPER, J. E. **Evidence for Gal3p's Cytoplasmic Location and Gal80p's Dual Cytoplasmic-Nuclear Location Implicates New Mechanisms for Controlling Gal4p Activity in *Saccharomyces cerevisiae***. *Molecular and Cellular Biology*, v. 20, n° 14, p. 5140–5148, 2000. ISBN: 0270-7306 (Print)r0270-7306 (Linking), ISSN: 0270-7306, DOI: 10.1128/MCB.20.14.5140-5148.2000.

PICATAGGIO, S. **Potential impact of synthetic biology on the development of microbial systems for the production of renewable fuels and chemicals**. *Current Opinion in Biotechnology*, v. 20, n° 3, p. 325–329, 2009. ISBN: 0958-1669, ISSN: 09581669, DOI: 10.1016/j.copbio.2009.04.003.

PINHEIRO, V. B., TAYLOR, A.I., COZENS, C., ABRAMOV, M., RENDERS, M., ZHANG, SU., CHAPUT, J.C., WENGEL, J., PEAK-CHEW, S-Y., MCLAUGHLIN, S.H., HERDEWIJN, P., HOLLIGER, P. **Synthetic Genetic Polymers Capable of Heredity and Evolution**. *Science*, v. 336, n° 6079, p. 341–344, 2012. ISBN: 0036-8075, ISSN: 0036-8075, DOI: 10.1126/science.1217622.

PIVA, L. C.; DE MARCO, J. L.; DE MORAES, L. M. P.; REIS, V. C. B.; TORRES, F. A. G. **Acetamidase as a dominant recyclable marker for *Komagataella phaffii* strain engineering**. *Applied Microbiology and Biotechnology*, v. 102, n. 6, p. 2753–61, 2018.

POTVIN, G.; AHMAD, A.; ZHANG, Z. **Bioprocess engineering aspects of heterologous protein production in *Pichia pastoris*: A review.** *Biochemical Engineering Journal*, v. 64, p. 91–105, 2012. ISBN: 1369-703X, ISSN: 1369703X, DOI: 10.1016/j.bej.2010.07.017.

PRIELHOFER, R., MAURER, M., KLEIN, J., WENGER, J., KIZIAK, C., GASSER, B., MATTANOVICH, D. **Induction without methanol: Novel regulated promoters enable high-level expression in *Pichia pastoris*.** *Microbial Cell Factories*, v. 12, n° 1, p. 1–10, 2013. ISBN: 1475-2859, ISSN: 14752859, DOI: 10.1186/1475-2859-12-5.

QUAN, J; TIAN, J. **Circular polymerase extension cloning for high-throughput cloning of complex and combinatorial DNA libraries.** *Nature Protocols*, v. 6, n°2, p. 242 – 251, 2011. DOI:10.1038/nprot.2010.181

RÉBORA, K.; LALOO, B.; DAIGNAN-FORNIER, B. **Revisiting purine-histidine cross-pathway regulation in *Saccharomyces cerevisiae*: A central role for a small molecule.** *Genetics*, v. 170, n° 1, p. 61–70, 2005. ISBN: 0016-6731 (Print)n0016-6731 (Linking), ISSN: 00166731, DOI: 10.1534/genetics.104.039396.

ROHLOFF, J.C., GELINAS, A.D., JARVIS, T.C., OCHSNER, U.A., SCHNEIDER, D.J., GOLD, L., JANJIC, N. **Nucleic Acid Ligands with Protein-like Side Chain: Modified Aptamers and their Use as Diagnostic and Therapeutic Agents.** *Molecular Therapy – Nucleic Acids*, n° 10:e201, 2014. DOI:10.1038/mtna.2014.49.

SCHWARTZ, C., FROGUE, K., RAMESH, A., MISA, J., WHEELDON, I. **CRISPRi repression of nonhomologous end-joining for enhanced genome engineering via homologous recombination in *Yarrowia lipolytica*.** *Biotechnology and Bioengineering*, v. 114, n° 12, p. 2896–2906, 2017. ISSN: 10970290, DOI: 10.1002/bit.26404.

SCHWARZHANS JP, WIBBERG D, WINKLER A, LUTTERMANN T, KALINOWSKI J, FRIEHS K. **Non-canonical integration events in *Pichia pastoris* encountered during standard transformation analysed with genome sequencing.** *Sci Rep* 6:38952. <https://doi.org/10.1038/srep38952>.

SEARS, I. B., O'CONNOR, J., ROSSANESE, O.W., GLICK, B.S. **A versatile set of vectors for constitutive and regulated gene expression in *Pichia pastoris*.** *Yeast*, v. 14, n° 8, p. 783–790, 1998. ISBN: 0749-503X, ISSN: 0749503X, DOI: 10.1002/(SICI)1097-0061(19980615)14:8<783::AID-YEA272>3.0.CO;2-Y.

SELLECK, S. B.; MAJORS, J. E. **In vivo DNA-binding properties of a yeast transcription activator protein.** *Mol Cell Biol*, v. 7, n° 9, p. 3260–3267, 1987. ISBN: 0270-7306, ISSN: 0270-7306, DOI: 10.1128/MCB.7.9.3260.

SELLICK, C. A.; CAMPBELL, R. N.; REECE, R. J. **Galactose Metabolism in Yeast — Structure and Regulation of the Leloir Pathway Enzymes and the Genes Encoding Them.** v. 269, n° 08, p. 111–150, 2008. DOI: 10.1016/S1937-6448(08)01003-4.

SHAO, Z.; ZHAO, H.; ZHAO, H. **DNA assembler , an in vivo genetic method for rapid construction of biochemical pathways.** v. 37, n° 2, p. 1–10, 2009. DOI: 10.1093/nar/gkn991.

SHEN, S., SULTER, G., JEFFRIES, T.W., CREGG, J.M. **A strong nitrogen source-regulated promoter for controlled expression of foreign genes in the yeast *Pichia pastoris*.** *Gene*, v. 216, n° 1, p. 93–102, 1998. ISBN: 0378-1119, ISSN: 03781119, DOI: 10.1016/S0378-1119(98)00315-1.

SMITH, J.E., MEDLEY, C.D., TANG, Z., SHANGGUAN, D., LOFTON, C., ANDTAN, W. **Aptamer-Conjugated Nanoparticles for the Collection and Detection of Multiple Cancer Cells.** *Analytical Chemistry*, n° 79, p. 3075 – 3082, 2007. DOI:10.1021/AC062151B.

SOLÀ, A., MAAHEIMO, H., YLÖNEN, K., FERRER, P., SZYPERSKI, T. **Amino acid biosynthesis and metabolic flux profiling of *Pichia pastoris*.** *European Journal of Biochemistry*, v. 271, n° 12, p. 2462–2470, 2004. ISBN: 0014-2956 (Print)r0014-2956 (Linking), ISSN: 00142956, DOI: 10.1111/j.1432-1033.2004.04176.x.

STADLMAYR, G., MECKLENBRÄUKER, A., ROTHMÜLLER, M., MAURER, M., SAUER, M., MATTANOVICH, D., GASSER, B. **Identification and characterisation of novel *Pichia pastoris* promoters for heterologous**

**protein production.** *Journal of Biotechnology*, v. 150, n° 4, p. 519–529, 2010. ISBN: 0168-1656, ISSN: 01681656, DOI: 10.1016/j.jbiotec.2010.09.957.

SUNGA, A. J.; CREGG, J. M. **The *Pichia pastoris* formaldehyde dehydrogenase gene (FLD1) as a marker for selection of multicopy expression strains of *P. pastoris*.** *Gene*, v. 330, n° 1–2, p. 39–47, 2004. ISBN: 0378-1119, ISSN: 03781119, DOI: 10.1016/j.gene.2003.12.015.

TUERK, C., AND GOLD, L. **Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase.** *Science* 249, 505–510, 1990. DOI: 10.1126/science.2200121

TKACHENKO, A.A., BORSHCHEVSKAYA L.N., SINEOKY, S.P., GORDEEVA, T.L. **CRISPR/Cas9-Mediated Genome Editing of the *Komataella phaffii* to Obtain a Phytase-Producer Markerless Strain.** *Biochemistry*, v. 88, p. 1338-1346. 2023.

TRAVEN, A.; JELICIC, B.; SOPTA, M. **Yeast Gal4: a transcriptional paradigm revisited.** *EMBO reports*, v. 7, n° 5, p. 2–5, 2006. DOI: 10.1038/sj.embor.7400679.

TSCHOPP, J. F., BRUST, P.F., CREGG, J.M., STILLMAN, C.A., GINGERAS, T.R. **Expression of the lacZ gene from two methanol-regulated promoters in *Pichia pastoris*.** *Nucleic Acids Research*, v. 15, n° 9, p. 3859–3876, 1987. ISBN: 0305-1048, ISSN: 03051048, DOI: 10.1093/nar/15.9.3859.

UGOLINI, S.; CV, B. **The red/white colony color assay in the yeast *Saccharomyces cerevisiae*: epistatic growth advantage of white *ade8-18, ade2* cells over red *ade2* cells.** *Current genetics*, v. 30, n° 6, p. 485–492, 1996. ISSN: 0172-8083.

VOGL, T, RUTH, C., PITZER, J., KICKENWEIZ, T., GLIEDER, A. **Synthetic core promoters for *Pichia pastoris*.** *ACS Synthetic Biology*, v. 3, n° 3, p. 188–191, 2014. ISBN: 1215421109, ISSN: 21615063, DOI: 10.1021/sb400091p.

VOGL, T.; GLIEDER, A. **Regulation of *Pichia pastoris* promoters and its consequences for protein production.** *New Biotechnology*, v. 30, n° 4, p. 385–404, 2013. ISBN: 1871-6784, ISSN: 18716784, DOI: 10.1016/j.nbt.2012.11.010.

VOGL, T.; HARTNER, F. S.; GLIEDER, A. **New opportunities by synthetic biology for biopharmaceutical production in *Pichia pastoris***. *Current Opinion in Biotechnology*, v. 24, n° 6, p. 1094–1101, 2013. ISBN: 0958-1669, ISSN: 09581669, DOI: 10.1016/j.copbio.2013.02.024.

WENINGER, A., HATZL, A-M., SCHMID, C., VOGL, T., GLIEDER, A. **Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris***. *Journal of Biotechnology*, v. 235, p. 139–149, 2016. ISBN: 0275987973, ISSN: 18734863, DOI: 10.1016/j.jbiotec.2016.03.027.

WENINGER, A., FISCHER, J.E., RASCHMANOVÁ, H., KNIELY, C., VOGL, T., GLIEDER, A. **Expanding the CRISPR/Cas9 toolkit for *Pichia pastoris* with efficient donor integration and alternative resistance markers**. *Journal of Cellular Biochemistry*, v. 119, n° 4, p. 3183–3198, 2018. ISSN: 10974644, DOI: 10.1002/jcb.26474.

XU, N., ZHU, J., ZHU, Q., XING, Y., CAI, M., JIANG, T., ZHOU, M., ZHANG, Y. **Identification and characterization of novel promoters for recombinant protein production in yeast *Pichia pastoris***. *Yeast*, v. 35, n° 5, p. 379–385, 2018. ISBN: 8621642530, ISSN: 10970061, DOI: 10.1002/yea.3301.

YAMADA, Y., MATSUDA, M., MAEDA, K., MIKATA, K. **The Phylogenetic Relationships of Methanol-assimilating Yeasts Based on the Partial Sequences of 18S and 26S Ribosomal RNAs: The Proposal of *Komagataella* Gen. Nov. (*Saccharomycetaceae*)**. *Bioscience, Biotechnology, and Biochemistry*, v. 59, n° 3, p. 439–444, 1995. ISBN: 0916-8451, ISSN: 0916-8451, DOI: 10.1271/bbb.59.439.

YANG, J., NIE, L., CHEN, B., LIU, Y., KONG, Y., WANG, H., DIAO, L. **Hygromycin-resistance vectors for gene expression in *Pichia pastoris***. *Yeast (Chichester, England)*, v. 31, n° 4, p. 115–25, 2014. ISSN: 1097-0061. DOI: 10.1002/yea.3001.

YANG, J.; JIANG, W.; YANG, S. **MazF as a counter-selectable marker for unmarked genetic modification of *Pichia pastoris***. *FEMS Yeast Research*, v. 9, n° 4, p. 600–609, 2009. ISBN: 1567-1356, ISSN: 15671356, DOI:



10.1111/j.1567-1364.2009.00503.x.

YU, X.-W., WANG, R., ZHANG, M., XU, Y., XIAO, R. **Enhanced thermostability of a *Rhizopus chinensis* lipase by in vivo recombination in *Pichia pastoris*.** *Microbial Cell Factories*, v. 11, n° 1, p. 102–112, 2012. ISBN: 1475285911102, ISSN: 1475-2859, DOI: 10.1186/1475-2859-11-102.