UNIVERSIDADE DE BRASÍLIA

INSTITUTO DE CIÊNCIAS BIOLÓGICAS

DEPARTAMENTO DE BIOLOGIA CELULAR

PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR

MICRORGANISMOS GENETICAMENTE MODIFICADOS APLICADOS À BIOCATÁLISE

João Heitor Colombelli Manfrão Netto

Brasília/DF

2021

UNIVERSIDADE DE BRASÍLIA

INSTITUTO DE CIÊNCIAS BIOLÓGICAS

DEPARTAMENTO DE BIOLOGIA CELULAR

PÓS-GRADUAÇÃO EM BIOLOGIA MOLECULAR

MICRORGANISMOS GENETICAMENTE MODIFICADOS APLICADOS À BIOCATÁLISE

Autor: João Heitor Colombelli Manfrão Netto

Orientadora: Prof.ª Dr.ª Nádia Skorupa Parachin

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

Brasília

2021

Este trabalho foi desenvolvido no Laboratório de Biologia Molecular da Universidade de Brasília (UnB) sob orientação da professora Drª. Nádia Skorupa Parachin e na *Division of Applied Microbiology* em *Lund University* sob supervisão do Dr. Magnus Carlquist.

Título: Microrganismos geneticamente modificados aplicados à biocatálise

---

João Heitor Colombelli Manfrão Netto

Tese de Doutorado

Banca Examinadora

Presidente: Profª. Drª. Nádia Skorupa Parachin

Programa de Pós-Graduação em Biologia Molecular (IB) - UnB

Membro interno: Prof. Dr Edivaldo Ximenes Ferreira Filho

Programa de Pós-Graduação em Biologia Molecular (IB) - UnB

Membro externo: Profª. Drª Betania Ferraz Quirino

Embrapa Agroenergia

Membro externo: Drª Léia Cecília de Lima Fávaro

Embrapa Agroenergia

Suplente: Profª. Drª Eliane Ferreira Noronha

Programa de Pós-Graduação em Biologia Molecular (IB) - UnB

Brasília

2021

Acknowledgments

Primeiramente, agradeço a Deus por sempre estar ao meu lado e iluminando meu caminho ao longo da minha jornada acadêmica. Mesmo não sendo uma pessoa tão religiosa, sinto que Deus sempre está ao meu lado me dando forças para nunca desistir do meu sonho de infância que foi concretizado neste trabalho.

Agradeço a minha esposa Joyce Sharloth Tavares de Freitas por estar ao meu lado do início ao fim do doutorado e pelo caminho que estamos trilhando juntos. Você me escolheu como companheiro de vida e vou sempre me esforçar o máximo para ser o melhor marido para você e cuidar da nossa família. Obrigado por sempre acreditar em mim e por sempre me ajudar quando eu estava perdido. Você me deu forças para fazer meu sonho de ser cientista se tornar a minha meta de vida e hoje me vejo como um pesquisador de verdade. Sinto que hoje eu consigo mudar o mundo com minha pesquisa e isso se deve principalmente ao nosso amor. Dedico esse trabalho a você meu amor, pois sem você não teria conseguido concluí-lo. Te amo hoje e para sempre.

Agradeço a minha orientadora Nádia Skorupa Parachin por sempre acreditar em mim e sempre apoiar minhas ideias. Você me ajudou enormemente ao longo da minha jornada acadêmica e vou ser eternamente grato a você. Você é um exemplo de vida para mim e para todos que trabalham com pesquisa no Brasil. Espero que possamos colaborar novamente no futuro. Muito obrigado professa!

I would like to thank Magnus Carlquist for the work developed during my Ph.D. exchange at Lund University. You help me a lot when I was living in Sweden and I appreciate working with you. I hope we can collaborate again shortly. Thank you a lot for everything.

Agradeço a toda minha família, principalmente a minha mãe Tânia, meu pai Franscisco, minha irmã Carol e minha irmã Cibeli por todo apoio ao longo da minha jornada acadêmica.

Agradeço a todos os meus colegas de trabalho do nosso grupo de pesquisa Engenharia de Biocatalisadores. Um agradecimento especial ao Enzo e a Kelly que ajudaram diretamente no desenvolvimento do trabalho referente à produção de ácido hialurônico por *Hansenula polymorpha*.

Faço um agradecimento individual ao meu amigo Henrique que me ajudou tanto enquanto estava morando na Suécia. Você me recebeu e me fez sentir em casa mesmo estando milhares de km de distância do Brasil. Me ajudou a iniciar meu trabalho no laboratório, me ensinou todos os macetes de viver em Lund e fez uma surpresa no dia do meu aniversário. Nunca vou esquecer tudo que você por mim Henrique e obrigado pela amizade que fizemos.

Agradeço aos colegas do Laboratório de Biologia Molecular, em especial ao pessoal do Lab 3, principalmente a professora Ildinete, o Fabian e o Jhones pela ajuda com os equipamentos utilizados para o desenvolvimento desse trabalho.

I would like to thank the staff from the Division of Applied Microbiology (Teknisk Mikrobiologi) at Lund University, especially Marie-Francoise Gorwa-Grauslund, Peter Rådström and Anette Ahlberga special thank to Nina Muratovska and Fredrik Lund helped develop the work related to *Pseudomonas putida* performing the bioconversions. Thank you a lot guys.

Agradeço aos meus companheiros do Jiu-jitsu, principalmente ao meu mestre Eric e ao Tiaguinho por todos os ensinamentos ao longo desses anos de treino. Agradeço também aos meus amigos de treino Edgar, Magal, Albert, Leandro, Igor, Luquinhas e todos os demais membros do Medeiros Team.

Por fim, agradeço as agências de fomento brasileiras FAP/DF, CNPq e CAPES. Agradeço a CAPES pelas bolsas concedidas durante o doutorado e agradeço ao Programa Institucional de Internacionalização da CAPES (CAPES/PrInt) pelo auxílio concedido para o desenvolvimento do trabalho na Universidade de Lund.

RESUMO

Biocatálise refere-se à utilização de enzimas ou células inteiras de microrganismos como catalisadores em uma reação química representando uma rota sustentável para síntese de químicos, principalmente devido à possibilidade de se utilizar resíduos industriais como substrato, por exemplo, a biomassa lignocelulósica. Com os avanços no campo da biotecnologia, várias moléculas podem ser sintetizadas através de biocatálise, incluindo compostos não produzidos naturalmente por microrganismos. Nestes casos, ferramentas de engenharia genética são utilizadas para inserção de genes exógenos ou para deleção de vias endógenas que competem pelo substrato, otimizando o processo de produção. Além disso, novas técnicas para edição de genes estão constantemente sendo desenvolvidas, ampliando os limites da biocatálise. Neste trabalho, o conceito de biocatalisador foi aplicado em dois microrganismos para o desenvolvimento de rotas alternativas para síntese de ácido hialurônico (AH) e vanililamina (VA).

AH é um biopolímero de alto valor, presente na composição de diversos produtos médicos e cosméticos. Abordagens alternativas para sua síntese estão sendo desenvolvidas para substituir o atual método utilizando processos fermentativos com espécies produtoras naturais de AH do gênero *Streptococcus*. Neste trabalho, AH foi produzido a partir linhagens geneticamente modificadas da levedura *Ogataea* (*Hansenula*) *polymorpha*. Para isso, o gene *hasA* (*hasAp* de *Pasteurella multocida* e *hasA*s de *S. zooepidemicus*), o qual codifica a enzima hyaluronan synthase (HAS), bem como o gene *hasB* (de *Xenopus laevis*), codificador da enzima UDP-glucose dehydrogenase, necessários para a síntese heteróloga de AH foram inseridos no genoma desta levedura. Além disso, foi implementado um interruptor genético para a regulação de *hasA* e *hasB* por uma serina integrase (integrase-13, Int13). Ao todo foram desenvolvidas três linhagens produtoras de AH, com a maior produção (≅ 200 mg/L) obtida pela linhagem EMB103 contendo o *interruptor genético*. Esse valor é inferior aos obtidos por outros produtores heterólogos de AH, incluindo leveduras. Contudo, o potencial de *O. polymorpha* para produção de AH através da expressão de diferentes genes *hasA* bem como a utilização de um *interruptor genético* nessa levedura foram demonstrados de forma inédita.

A segunda parte deste trabalho foi relacionada à produção de VA a partir de substratos derivados da lignina (vanilina e ácido ferúlico) utilizando a bactéria *Pseudomonas putida* como biocatalisador, demonstrando uma nova rota biotecnológica para síntese de aminas aromáticas.

A produção de VA por uma linhagem geneticamente modificada de *P. putida* (GN442ΔPP_2426) foi realizada através da superexpressão do gene *CV-ATA* que codifica para uma amina transaminase (ATA) de *Chromobacterium violaceum* bem como três ATAs endógenas de *P. putida* KT2440 identificadas através de análises *in silico*. Além disso, os genes codificadores das três ATAs endógenas e o gene *CV-ATA* foram superexpressos em *P. putida* KT2440 resultando em linhagens capazes de utilizar VA, que não é prontamente metabolizada como única fonte de carbono por esta bactéria. A maior produção de VA (0.915 mM) a partir de vanilina (10 mM) foi obtida pela linhagem TMB-JH004 superexpressando *CV-ATA*. Contudo, o processo foi limitado principalmente pela reassimilação do produto VA e pela formação de dois coprodutos, o ácido vanílico e o álcool vanilil. A expressão do gene *AlaDH*, o qual codifica uma alanina desidrogenase de *Bacillus subtilis*, levou a uma redução da reassimilação de VA representando uma estratégia eficaz para melhorar a produção de VA a partir de vanilina. Portanto, esses resultados demonstram o potencial de *P. putida* como biocatalisador para produção de aminas aromáticas a partir de derivados da lignina.

Palavras-chave: Biocatálise, *Ogataea polymorpha,* Ácido hialurônico*, Pseudomonas putida*, aminas aromáticas

ABSTRACT

Biocatalysis refers to the utilization of enzymes or whole cells of microorganisms as catalysts in a chemical reaction representing a sustainable route for chemical synthesis, mainly due to the possibility of using industrial waste as substrate, such as the lignocellulosic biomass. With advances in the biotechnology field, a wide range of molecules can be synthesized through biocatalysis, including compounds not naturally produced by microorganisms. In these cases, genetic engineering tools are utilized to insert exogenous genes or to delete endogenous pathways that compete for the substrate, optimizing the production process. Also, new techniques for gene editing are constantly being developed, expanding the limits of biocatalysis. In this work, the biocatalyst concept was applied to two microorganisms for the development of alternative routes for the synthesis of hyaluronic acid (HA) and vanillylamine (VA).

HA is a high-value biopolymer, present in the composition of several medical and cosmetics products. Alternative approaches to its synthesis are being developed to replace the current method utilizing fermentative processes with natural HA producing species of the genus *Streptococcus*. In this work, HA was produced from genetically modified strains of the yeast *Ogataea* (*Hansenula*) *polymorpha*. For this, the genes *hasA* (*hasAp* from *Pasteurella multocida* and *hasAs* from *S. zooepidemicus*), which encodes the enzyme hyaluronan synthase (HAS), as well as the *hasB* gene (from *Xenopus laervis*) encoding the UDP-glucose dehydrogenase enzyme, necessary for the heterologous synthesis of HA were inserted into the genome of this yeast. Additionally, a genetic switch was implemented for the regulation of *hasA* and *hasB* genes by a serine-integrase (integrase-13, Int13). Altogether, three strains producing HA were developed, with the highest production (≅ 200 mg/L) obtained by strain EMB103 containing the genetic switch. This value is lower than titers obtained by other heterologous HA producers, including yeasts. Nevertheless, the potential of *O. polymorpha* for HA production by the expression of different *hasA* genes and the utilization of a genetic switch in this yeast were demonstrated for the first time.

The second part of this work was related to VA production from lignin-derived substrates (vanillin and ferulic acid) utilizing the bacterium *Pseudomonas putida* as a biocatalyst, demonstrating a new eco-friendly route for the synthesis of aromatic amines. The production of VA by a genetically modified *P. putida* strain (GN442ΔPP_2426) was performed through the overexpression of the *CV-ATA* gene encoding an amine transaminase (ATA) from

Chromobacterium violaceum as well as by three endogenous *ATAs* from *P. putida* KT2440 identified by *in silico* analysis. Furthermore, the genes encoding the three endogenous ATAs and the *CV-ATA* gene were overexpressed in *P. putida* KT2440 resulting in strains capable to utilize VA, which is not promptly metabolized as the sole carbon source by this bacterium. The highest VA production (0.915 mM) from vanillin (10 mM) was obtained by the strain TMB-JH004 overexpressing *CV-ATA*. However, the process was limited mainly by the re-assimilation of VA and by the formation of two co-products, vanilic acid and vanillyl alcohol. The expression of *AlaDH* gene encoding an alanine dehydrogenase from *Bacillus subtilis* led to a reduction in VA reassimilation representing an effective strategy to improve VA production from vanillin. Therefore, these results demonstrate the potential of *P. putida* as a biocatalyst to produce aromatic amines from lignin-derived substrates.

Keywords: Biocatalysis, *Ogataea polymorpha*, Hyaluronic acid, *Pseudomonas putida*, aromatic amines

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ATA: Amine transaminase

ATP: Adenosine triphosphate

CAGR: Compound annual growth rate

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

DNA: Deoxyribonucleic acid

HA: Hyaluronic acid

HAS: Hyaluronan synthase

Int: Integrase

IPTG: Isopropyl ß-D-1-thiogalactopyranoside

MW: Molecular weight

NHEJ: Non-homologous end joining

OD: Optical density

PCR: Polymerase chain reaction

TALENs: Transcription activator-like effector nucleases

TCA: Tricarboxylic acid cycle

UDP-GlcNac: UDP-N-Acetyl-Glucosamine

UDP-GlcUA: UDP-Glucuronic Acid

UTP: Uridine triphosphate

ORGANIZATION OF THE THESIS

This thesis describes the construction and application of genetically engineered microorganisms as biocatalysts to produce two value products: hyaluronic acid and vanillylamine. Chapter I provides an overview of biocatalysis and how metabolic engineering can be applied to improve the biocatalyst's performance. In this chapter it is also discussed the difference between enzymatic and microbial biocatalysis.

Chapter II was dedicated to the construction of metabolically engineered strains of methylotrophic yeast *Ogataea polymorpha* to synthesize the biopolymer hyaluronic acid in a fermentative process. The genes necessary for its production (*hasA* and *hasB*) were inserted into the *O. polymorpha* genome in different combinations of promoters to control their expressions. It was also implemented a genetic switch to control the expression of both enzymes by a serine-integrase. This work was the aim of Paper II.

Chapter III shows the production of an aromatic amine in the context of lignin valorization applying the soil bacterium *Pseudomonas putida* as the biocatalyst for the biotransformation of two lignin-derived substrates vanillin and ferulic acid. The cells overexpressing a gene encoding an amine transaminase from *Chromobacterium violaceum* utilizing a replicative plasmid were able to produce vanillylamine from vanillin and ferulic acid. In this work, it was also identified three endogenous amine transaminases from *P. putida* KT2440 with activity against vanillylamine. The overexpression of all amine transaminases generated strains able to produce vanillylamine from vanillin. Lastly, it was evaluated the co-expression of an alanine dehydrogenase of *Bacillus subtilis* encoding gene with the amine transaminase from *C. violaceum* to improve the titers of vanillylamine obtained. These results are described in Paper III.

TABLE OF CONTENTS

1. CHAPTER 1

Introduction

1.1 Biocatalysts as a sustainable alternative for chemical synthesis

There is growing interest in utilizing renewable feedstocks to produce daily-consumed chemicals following the *Twelve Green Chemistry principles* (IVANKOVIĆ, 2017). Substituting petroleum-based starting material for raw vegetal material to obtain such chemicals is a greener approach gaining attention in recent years (DESSBESELL et al., 2020). In this context, the utilization of biocatalysts represents a feasible alternative to the current chemical processes utilized. Biocatalysis is defined as the application of isolated enzymes or whole-cell as a catalyst for a chemical reaction (SHELDON; BRADY, 2018). According to the Green chemistry principles (Table 1), this is a sustainable alternative to substitute the chemical synthesis to obtain desired products. It is based on a biomimetic approach inspired by the natural synthesis in living systems (ANASTAS; EGHBALI, 2010).

Table 1. How biocatalysis fits in the *Twelve Principles of Green Chemistry*. It was adapted from (SHELDON; WOODLEY, 2018).

	Green Chemistry Principle	Biocatalysis
1	waste prevention	significantly reduced waste
2	atom economy	more atom- and step-economical
3	less hazardous syntheses	generally low toxicity
4	design for safer products*	not relevant (product not process)
5	safer solvents and auxiliaries	usually performed in water
6	energy efficiency	mild conditions/energy-efficient
7	renewable feedstocks	enzymes are renewable
8	reduced derivatization	avoids protection/deprotection steps
9	catalysis	enzymes are catalysts
10	design for degradation*	not relevant (product not process)
11	real-time analysis	applicability to biocatalytic processes
12	inherently safer processes	mild and safe conditions

*Principles 4 and 10 are related to the product rather than the process.

Although biocatalysis has gained the scientific community's attention since the 80s, the utilization of enzymes or microorganisms for chemical synthesis is centenary. For instance, in 1858, Louis Pasteur accomplished a kinetic resolution of an aqueous racemic solution of (+)-tartaric acid applying the fungi *Penicillium glaucum*, resulting in the enrichment of (-)-tartaric acid in the solution (GARZÓN-POSSE et al., 2018). Several authors consider this experiment the milestone for biocatalysis, although fermentative processes to produce beer, wine, and vinegar have been utilized for many years (or even centuries). Subsequent studies in which the

microorganisms or enzymes were explored for chemical synthesis are considered the *first wave* of biocatalysis. The advances in the engineering of biocatalysts accompanied the biotechnology revolution, especially regarding protein engineering, recombinant DNA techniques, directed evolution, and more recently "*omics*" approaches, metabolic engineering, and synthetic biology (GUAZZARONI; SILVA-ROCHA; WARD, 2015) (BADENHORST; BORNSCHEUER, 2018). Therefore, consonant with the growing demand for sustainable alternatives for chemical synthesis, it is expected that biocatalysis becomes an essential source of chemicals and other products at the industrial level. According to IndustryARC™, the global market size of Biocatalysis & Biocatalysts is expected to grow at CAGR of 5.1% in the period of 2020-2025 reaching $120 billion at 2025 (https://www.industryarc.com; Report Code: CMR 0628).

1.2 What biocatalyst to choose? Comparison between enzymatic and microbial biocatalysis

Purified enzymes or whole-cells can be applied as biocatalysts to produce fine chemicals, bulk chemicals, biofuels, biopolymers, pharmaceuticals, among other products (DE CARVALHO, 2017). In both cases, biocatalysis have higher selectivity and specificity than chemical synthesis, particularly for kinetic resolutions of racemic mixtures. Furthermore, it is a greener approach for obtaining chemicals, as discussed above. However, there are some important differences to be considered when choosing between enzymatic and microbial biocatalysis (Table 2).

Table 2. Advantages and disadvantages of catalysis using isolated enzymes or whole-cells. Adapted from (PINTO; CONTENTE; TAMBORINI, 2020)

Comparison between whole cells and isolated enzymes as biocatalysts.		
Adv./Disadv.	**Isolated enzymes**	**Whole cells**
Advantages	No side reactions Higher permeability Tolerance to high substrate concentrations	No exogenous addition of cofactors Possibility to perform cascade reactions with the same biocatalyst No purification processes
Disadvantages	Addition of exogenous cofactors Increased time and costs associated with their purification processes Lower stability outside the cell environment	Side reactions due to metabolic pathways Difficult product recovery Low permeability due to cellular membrane

In isolated enzyme systems, no side reactions are expected, and product inhibition is less frequent. Besides, enzymes are not affected by cytotoxicity related to the substrate as the microorganisms. In contrast, the utilization of isolated enzymes includes a step for the production and purification of the biocatalyst, and cofactors usually need to be provided, increasing the process costs. Cost is a relevant aspect of multi-step reactions to synthesize the desired product from a specific substrate. For example, the production of HA by heterologous microorganisms requires de insertion of one gene (*hasA*) in bacteria and two genes in yeasts (*hasA* and *hasB*). However, the synthesis of HA utilizing enzymatic biocatalysis requires six enzymes that should be produced and purified before utilization (GOTTSCHALK et al., 2019). Additionally, the cofactors for each reaction are provided exogenously, making the process more expensive. In this case, microbial biocatalysis is advantageous since the cell metabolism, and cascade reactions that supply cofactors are easier (and cheaper) to be performed once the cell metabolism can provide some of these reactions. Besides, side reactions can help to recycle coproducts, which is essential for reactions operating in equilibrium. For instance, in the work described in Chapter III, the production of the product (VA) by the *P. putida* strain GN442ΔPP_2426 expressing an ATA encoding gene was limited by an unfavored equilibrium driving the reaction to the production of vanillin. The overexpression of an alanine dehydrogenase encoding gene led to product accumulation and a slight increase in the titer of VA produced by the cells. This enzyme helps to recycling the coproduct pyruvate producing alanine utilized for the ATA to produce VA. Therefore, microbial biocatalysis is advantageous for multi-step reactions and for supplying cofactors.

Nevertheless, the main bottlenecks for whole-cell biocatalysis are: the permeability selectivity of the cell membrane, which acts as transport barrier for some substrates; reaction inhibition by the substrate, intermediates, or even the product; side reactions competing with the main reaction (biocatalysis) for the substrate or intermediates, and product re-assimilation. These factors directly affect the cost, yield and feasibility of the process, especially for production at the industrial level (LEE; KIM, 2015). In this perspective, the biocatalyst design should be determined for matching its application, choosing the proper biocatalysts for each situation.

1.3 Application of whole-cells for biocatalysis

Whole-cell biocatalysis is usually divided into (i) Biotransformations and (ii) Fermentations. In biotransformation, cell growth and the bioconversion occur separately, while in fermentations, the target product is derived from the growth substrate and produced by the cell metabolic framework (LIN; TAO, 2017). Some authors also consider biotransformation as the process in which the growth substrate and the substrate to be converted are different while in the fermentations both substrates are the same (KUHN et al., 2010). In both approaches, the native cell metabolism can be explored for biocatalysis allowing multi-step reactions and cofactor supplying/recycling for these reactions. However, some side reactions from endogenous enzymes compete with the main reaction for the substrate leading to low concentrations (titers) of the final product. Therefore, cell metabolism should be funnelled to product synthesis by metabolic engineering approaches (Figure 1).

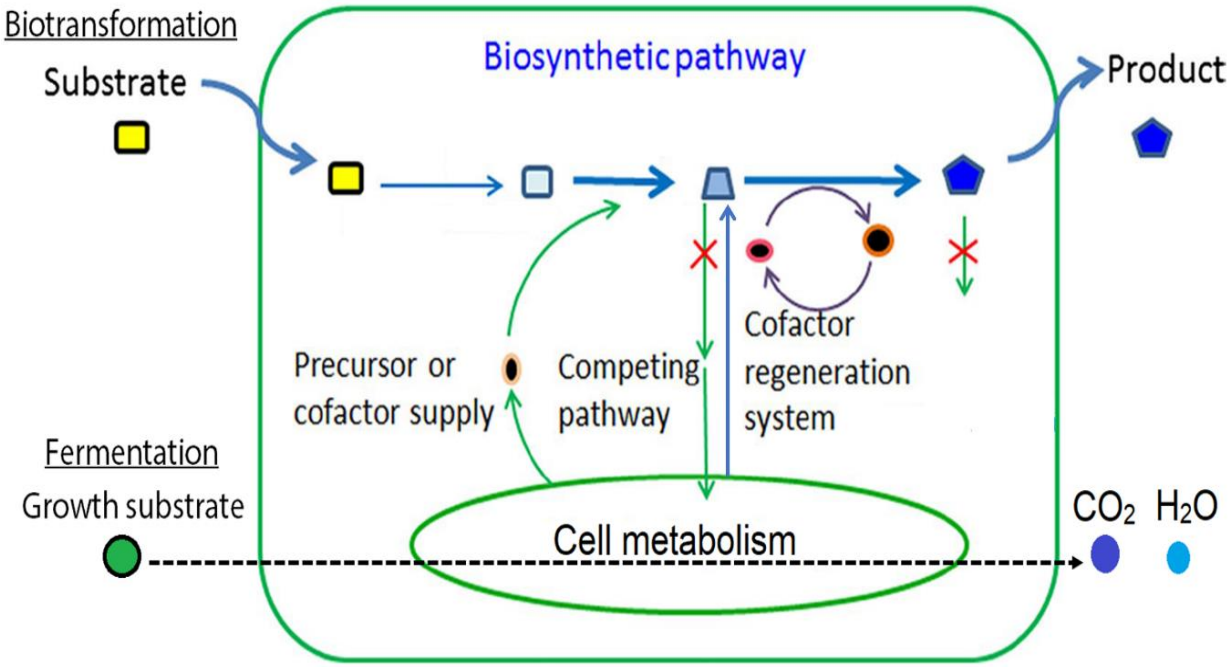


Figure 1. Overview of whole-cell biocatalysis highlighting the main bottlenecks for this type of bioconversion. The red arrows represent deletions in competitive pathways that deviate the substrate, intermediates, or product. Figure adapted from (LIN; TAO, 2017).

Biocatalysts can be designed by deleting competitive pathways, overexpressing genes encoding endogenous enzymes and/or inserting exogenous genes into the host genome to produce non-native products. The principle is funneling the host metabolic framework by eliminating non-essential reactions but preserving the cell metabolism.(LEE; KIM, 2015)(KO

et al., 2020). Additionally, molecular biology tools are crucial to developing industrial strains without selective markers (e.g., antibiotic resistance) and avoiding the utilization of expensive inducers (e.g., IPTG). Antibiotic resistance markers can represent a route to spread antibiotic resistance to bacterial communities in the environment and should be avoided in strains applied for industrial scale production. Therefore, metabolic engineering through molecular biology tools is a powerful approach to design efficient microbial biocatalysts. Some important molecular tools include the polymerase chain reaction (PCR), deletion cassettes, gene cloning, replicative/integrative vectors, etc. Also, genome editing tools such as Zinc finger nucleases, TALENs, and the revolutionary CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) technology are applied in the metabolic engineering of biocatalysts. These techniques are utilized to insert/delete genes, cause site-specific mutation, up/downregulation of gene expression, remove selective markers, and multiplexed gene editing in which genes are edited simultaneously (for details for each technique see (SCHMIDT; SIEBERT, 2001)) .

Process scale-up is the main limition for the utilization of microorganisms for production in industrial scale. However, some commercial products have been succesfully synthetized by engineered microorganisms at the industrial level, including 1,4-butanediol (BURGARD et al., 2016), β-farnesene (BENJAMIN et al., 2016); artemisinin (PADDON et al., 2013), among others as reviewed by (WEHRS et al., 2019) and (KO et al., 2020). Some examples given in these reviews include lactic acid, polylactic acid, polyhydroxyalkanoates, succinic acid, isoprene, L-lysine, L-glutamate, vitamin B-12, and others. Therefore, industrial bioprocesses are a reality, and advances in the biotechnology field lead to a new era in which biocatalysts can be designed to fit the production process, synthesizing desired products from renewable substrates. In this thesis, two microorganisms were engineered through molecular biology tools to bypass the bottlenecks related to hyaluronic acid (HA) and vanillylamine (VA) production.

1.4 Two case studies: production of HA and VA by microbial biocatalysis

Here, the concept of microbial biocatalysis was applied to develop engineered strains to produce HA and VA as alternative methods to obtain these products. Both fermentative and biotransformation processes were applied for these biocatalysts to produce HA and VA, respectively. Molecular biology tools were applied to insert the genes necessary to produce HA

in a non-conventional yeast *Ogataea* (*Hansenula*) *polymorpha* and to produce VA from two lignin-derived substrates in *Pseudomonas putida*.

HA is a high-value biopolymer widely applied in medical and cosmetic products. Alternative approaches are being developed to obtain HA since fermentative processes using pathogenic microorganisms (mainly *Streptococcus* spp.) are still utilized for its production, although engineered *B. subtilis* strains are already applied for the synthesis of commercial HA (HyaCare®; Novozymes). Here, the production of HA by a fermentative process was demonstrated for the first time in the non-conventional yeast *O. polymorpha*. The synthesis of HA was achieved by inserting two genes (*hasA* and *hasB*) into the genome of *O. polymorpha* using an integrative vector. Additionally, as a proof-of-concept, a genetic switch was implemented using a serine-integrase to regulate both genes expression. The production of HA by *O. polymorpha* is reported in Chapter 2 based on Paper I (MANFRÃO-NETTO; GOMES; PARACHIN, 2019) and Paper II (MANFRÃO-NETTO et al., 2021a).

The second part of this work was related to VA's production from two lignin-derived substrates (vanillin and ferulic acid) in *P. putida* using biotransformation representing a novel eco-friendly route to produce aromatic amines. These compounds are important platform chemicals to produce Active Pharmaceutical Ingredients (APIs). In the case of VA, it is utilized to obtain capsaicin, which is applied as an antimicrobial and anticarcinogenic agent, painkiller, and in the treatment of different metabolic disorders. The production of VA by *P. putida* was achieved by whole-cell biocatalysis utilizing cells overexpressing a gene encoding an amine transaminase (ATA). Additionally, a growth-based approach for the evaluation of ATA activity against VA was implemented. The method was applied to validate one exogenous enzyme activity and three endogenous ATAs from *P. putida*. The description of the screening method, the *in silico* analysis of three endogenous ATA, and the VA production are reported in Chapter 3 based on Paper III (MANFRÃO-NETTO et al., 2021b).

2. Chapter 2

Production of Hyaluronic acid by engineered Ogataea polymorpha strains

2.1 Introduction

2.1.1 Hyaluronic Acid: Importance and features

Hyaluronic acid is a glycosaminoglycan composed by units of UDP-glucuronic acid (UDP-GlcUA) and UDP-N-acetyl-glucosamine (UDP-GlcNAc) disaccharides ligated by glycosidic bonds β-1,4 and β-1,3 (Figure 2). This polymer has versatile applications in the pharmaceutical and cosmetic industry, being used in cosmetics, eye drops, surgery, scar treatment, as well as controlled drug delivery systems. Besides, it is applied in treatments of osteoarthritis and utilized in medical areas such as otolaryngology, dermatology, and rheumatology (HUYNH; PRIEFER, 2020). The vast HA utilization occurs due to its rheological properties, viscoelasticity, and biocompatibility, influenced by the molecular weight (MW) of the HA polymer (MARCELLIN; STEEN; NIELSEN, 2014). Thus, the concentration and ligation of the precursors and the polymer chain's elongation are limiting steps for the synthesis and bioactivity of HA. Due to this wide range of applications in health and aesthetics, this biopolymer has a great market potential (DE OLIVEIRA et al., 2016). According to Grand View Research Inc., it is expected that the global HA market reaches USD 16.6 billion by 2027, with a CAGR of 8.1% between 2020 and 2027 (https://www.grandviewresearch.com/industry-analysis/hyaluronic-acid-market; Report ID: 978-1-68038-333-1).

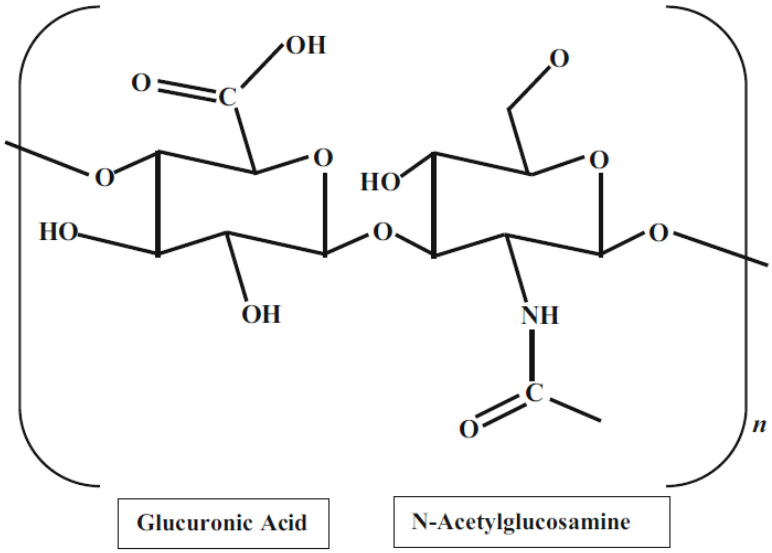


Figure 2. Molecular structure of an HA monomer with two pathways Glucuronic Acid and N-acetylglucosamine. This figure was obtained from (SZE; BROWNLIE; LOVE, 2016).

This biopolymer was first described and isolated in the 1930s from bovine vitreous humor (PALMER; MEYER, 1934). HA is present in soft tissues of vertebrates, algae, mollusks, and the bacteria capsule (SCHIRALDI; GATTA; ROSA, 2010). In humans, it is found throughout the body in different amounts, being abundant in connective tissues. High concentrations of HA in the synovial fluid help to lubricate and reduce joints' impacts (DOVEDYTIS; LIU; BARTLETT, 2020). One of this biopolymer's main medical applications is treating osteoarthritis, a degenerative joint disease that affects approximately 15 million people in Brazil (HILLMAN et al., 2017). In medical treatments for patients with osteoarthritis, HA supplementation, usually in injections, helps relieve pain and symptoms. Treatments using HA with MW of 0.5-2 MDa require five weekly injections, while those with MW of 3.5-6 MDa need only three weekly injections (MARCELLIN; STEEN; NIELSEN, 2014), demonstrating how chain size affects its therapeutic properties.

2.1.2 Obtaining Hyaluronic Acid

The sources to obtaining HA are animal tissues (rooster crests, animal vitreous humor, and synovial fluid) and fermentation with bacteria. However, the extraction of HA from animal tissues has been replaced for fermentative processes using bacteria. The increased demand for "*Greener Products*" and the difficulty to recover high-purity HA from animal extractions make bacterial production the preferred method for HA production. However, it can be obtained from animal tissues (MURADO et al., 2012). Therefore, microorganisms capable of synthesizing HA are a viable alternative for its production.

The synthesis of HA involves the enzyme hyaluronan synthase (HAS), first identified and purified from *Streptococcus pyogens* (DEANGELIS; PAPACONSTANTINOU; WEIGEL, 1993). It is an integral membrane enzyme capable of polymerizing the two precursors in the cytosol, elongating the polymer chain, and releasing it into the extracellular space, being present in mammalian cells, amphibians, and bacteria (TSEPILOV; BELODED, 2015). The HAS enzymes are divided into Class I and II. Class I enzymes are found in bacteria and vertebrates, while Class II has been described only in *Pasteurella multocida* (DEANGELIS, 1999). Class I HAS have multiple transmembrane domains, while Class II does not have any transmembrane domains, anchoring itself on the cell membrane from the C-terminal portion of its chain (WEIGEL; HASCALL; TAMMI, 1997). Nevertheless, HAS from both classes are encoded by the gene *hasA*.

For each HA mol produced, three molecules of Adenosine triphosphate (ATP) and two molecules of uridine triphosphate (UTP) are required (GOMES et al., 2019). Additionally, two moles of glucose and one mol of acetyl-CoA are also necessary for its synthesis. This demand is related to the precursors' synthesis, which involves different genes (Figure 3). These genes are commonly named as *has* due to the *has* operon dedicated to HA synthesis in *Streptococcus* spp. (SZE; BROWNLIE; LOVE, 2016).

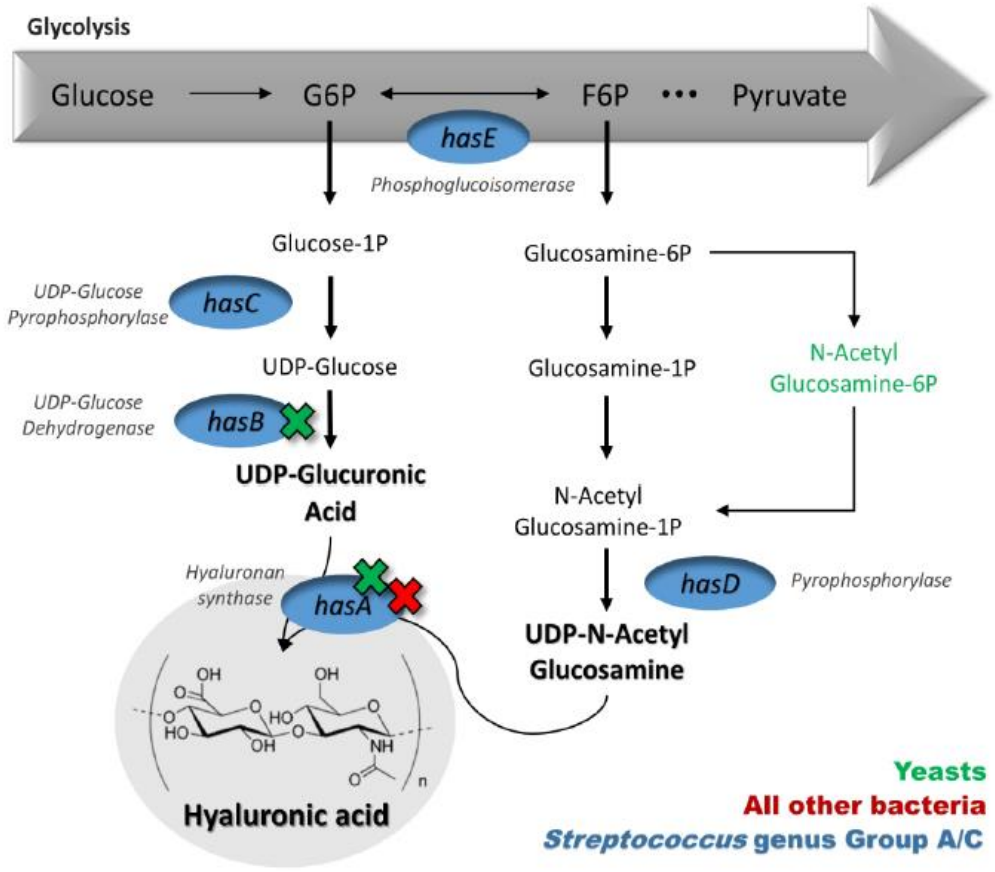


Figure 3. The microbial metabolic pathway of HA synthesis of natural producers is highlighted in blue and shows the HAS operon. The crosses indicate the lacking genes for HA production for bacteria (red) and yeast (green). Bacteria in general, except *Streptococcus* genus from the group A/C, lacks only the *hasA*, while yeasts lack *hasA* and *hasB*. Also, yeasts use N-Acetyl Glucosamine-6-Phosphate as a precursor to synthesize UDP-NAcetylglucosamine-1-Phosphate. **Reference**: Manfrão-Netto et al. 2021 (Paper IV).

As mentioned above, the microbial biosynthesis of HA has replaced the extraction from animal tissues and represents an effective and safe alternative to obtain this biopolymer. HA is produced in industrial scale from fermentation with *Streptococci*, mainly using *S. zooepidemicus*, since the early 1980s (LIU et al., 2011). However, the natural producers of HA,

such as *P. multocida* and the *Streptococci* are pathogens and fermentative processes using these microorganisms require subsequent purification processes to eliminate toxins. An alternative is the utilization of genetically modified non-pathogenic microorganisms to produce HA.

2.1.3 Strategies for production of HA in heterologous systems

Recently, the production of HA by heterologous systems is gaining attention due to the advances in the application of molecular biology tools. The implementation of novel genetic tools and their impact on HA production was the study of **Paper IV**. Gram-positive bacteria are the main microorganisms utilized as host platforms for the heterologous production of HA. Among those, *Bacillus subtilis*, *Corynebacterium glutamicum*, and *Lactococcus lactis* stand out as higher HA producers. The Gram-negative bacterium *Escherichia coli* is also commonly employed for HA production. Although heterologous producers represent a viable alternative for HA production, further metabolic engineering is required to reach titers of 6-7 g/L, similar to the natural producers used in industrial processes (BOERIU et al., 2013).

Many of these approaches are related to the deletion or downregulation of competitive pathways such as glycolysis, pentose phosphate pathway (PPP), or lactate pathway. In those cases, the genes *zwf* (encoding Glucose-6-Phosphate Dehydrogenase), *pfkA* (encoding 6-phosphofructokinase), and *ldh* (encoding lactate dehydrogenase) are targeted for deletion or downregulation. Therefore, the main bottlenecks for HA production in heterologous systems are the competitive pathways, which deviate carbon (primarily glucose) and energy. Additionally, improving the availability of both precursors through the overexpressing of *has* genes (*hasB*, *hasC*, *hasD* and/or *hasE*) is also a strategy employed to improve HA production in heterologous systems. Other strategies include utilizing cell membrane engineering, which is usually related to increases in cardiolipin content and/or distribution in the membrane (WESTBROOK et al., 2018). This approach is based on the intrinsic relation between this phospholipid and the HAS enzyme activity (TLAPAK-SIMMONS et al., 1998)(TLAPAK-SIMMONS et al., 1999). Thus, several metabolic engineering strategies are available in the scientific literature regarding HA production improvement by heterologous systems.

The application of such molecular biology tools leads to efficient host platforms producing HA titers similar or even higher than the natural producers. For instance, in *B. subtilis*, the downregulation of *pfkA* gene concomitant with the overexpression of the genes from the *has* operon led to the production of 19.38 g/L of HA low MW (6.62 kDa) using a 3-L

bioreactor (JIN et al., 2016). For *L. lactis,* the deletion of *ldh* gene with the overexpression of *hasB* and *hasD* generated a strain capable of producing 3.03 g/L of HA with MW of 1.4 MDa (KAUR; JAYARAMAN, 2016). In *C. glutamicum*, the overexpression of the *hasB* gene from *C. glutamicum* (*cgugdA2*) and the phosphoglucosamine mutase encoding gene from *P. putida* KT2440 (*ptglmM*) concomitant with the deletion of two genes (*cg0420* and *cg0424*) encoding putative glycosyltransferases related to the synthesis of extracellular polysaccharide generated a strain capable of producing 74.1 g/L of HA with MW of 53 kDa in a 5-L bioreactor in fed-batch mode (WANG et al., 2020). To the best of our knowledge, this is the highest HA titer reached so far by microorganisms. These are a few examples of the application of metabolic engineering to improve the HA synthesis in heterologous systems.

Regarding genetically modified yeasts to produce HA, only a few host platforms were developed. Most yeasts lack the UDP-glucose dehydrogenase enzyme encoded by the *hasB* gene, requiring the introduction of this gene and *hasA* for the synthesis of HA. A recombinant strain of *Komagataella phaffii* (commonly known as *Pichia pastoris*) was developed to produce HA (JEONG; SHIM; KIM, 2014). For this, all genes from the *has* operon (Figure 3) were introduced in yeast in different combinations, generating strains capable of producing HA with MW ranging from 1.2-2.5 MDa and titer of 0.8-1.7 g/L. The highest titer of 1.7 g/L was reached by the strain named EJP-CD containing the gene *hasA* and *hasB* both from *X. laevis* (named *xhasA2* and *xhasB*, respectively) as well as the natives *hasC* and *hasD* from *K. phaffii.* The other yeast-based platform developed for HA production was the non-conventional yeast *Kluyveromyces lactis* (GOMES et al., 2019). The *hasA* gene from *P. multocida* and the *hasB* gene from *X. laevis* were introduced in *K. lactis* genome, generating the strain BAP, which produced 1.89 g/L of HA with 2.097 MDa. To date, these are the only examples of genetically modified yeasts for HA production. However, the *DG42* gene from *Xenopus* sp. (*hasA*) was introduced in *S. cerevisiae*, and the DG42 enzyme was able to synthesize HA *in vitro* with MW ranging from 1-10 MDa (DEANGELIST; ACHYUTHAN, 1996). Here, the potential of the methylotrophic yeast *O. polymorpha* was evaluated as a host platform for HA production.

2.1.4 General characteristics of *Ogataea polymorpha*

The non-conventional yeast *O. polymorpha* is commonly used as a host platform due to its unique features, as discussed in Paper I. These characteristics include the capacity to metabolize several carbon sources such as glycerol, glucose, xylose, and cellobiose

(RYABOVA; CHMIL; SIBIRNY, 2003). In addition, it is capable of using methanol as a carbon source, which allowed the isolation of methanol-inducible promoters such as the pMOX (methanol oxidase), pFMD (formaldehyde dehydrogenase), and pDHAS (dihydroxyacetone synthase). During methanol induction, the peroxisomes that contain the enzymes necessary to metabolize methanol increase in size and number, filling almost the whole cytosol. Due to this characteristic, *O. polymorpha* is a model organism for studies of peroxisomes and methanol metabolism (VAN DER KLEI; VEENHUIS, 2002). These promoters are considered strong and are usually chosen to control the expression of heterologous genes (Table 1, Paper I), once in the presence of methanol, a high-level of the recombinant protein can be reached. *O. polymorpha* is a thermotolerant yeast, and can be cultivated at temperatures ranging between 30 to 50 °C. This phenotype is beneficial regarding the synthesis of mammalian proteins, especially those needing 37 °C to preserve their activity (VAN DIJK et al., 2000). The glycosylation pathway is present in *O. polymorpha*, allowing proper folding of recombinant eukaryotic proteins. Unlike other yeasts, it adds less sugar residues to the protein, preventing hyperglycosylation of recombinant proteins (KIM et al., 2006). A summary of all *O. polymorpha* features is illustrated in Figure 1 from Paper I.

There are three parental strains of *O. polymorpha* with distinct origins that are often used to produce recombinant proteins. The DL-1 strain (NRRL-Y-7560; ATCC26012) was isolated and characterized in soil samples (LEVINE; COONEY, 1973). The strain CBS4732 (CCY38-22-2; ATCC34438, NRRL-Y-5445), on the other hand, was isolated from irrigated soils in Pernambuco, Brazil (MORAIS; MAIA, 1959). These two strains are mainly utilized in the industrial processes. The strain NCYC495 (CBS1976; ATAA14754, NRLL-Y-1798) is commonly employed for laboratory research, especially for peroxisome studies, and it was isolated in Florida, USA (WICKERHAM, 1951). *O. polymorpha* was formerly named *Hansenula polymorpha*, however, analysis of phylogeny demonstrated that *H. polymorpha* constitutes two different species, *O. polymorpha* and *O. parapolymorpha* (KURTZMAN; ROBNETT, 2010); (SUH; ZHOU, 2010). The strains NCYC495 and CBS4732 are closely related and represent *O. polymorpha*, while DL-1 is phylogenetically distant, and it was reclassified as *O. parapolymorpha*. Previously, all strains were named *H. polymorpha*, but after the three strains' genome sequence, it was verified that these were two species. Nevertheless, both species share all the characteristics mencioned above. The genome sequences from these strains are available: DL-1 (RAVIN et al., 2013), NCYC495 (RILEY et al., 2016), and CBS4732 (RAMEZANI-RAD et al., 2003).

2.1.5 *O. polymorpha as a Microbial Cell Factory*

Several studies have focused on the genetic modification of *O. polymorpha* to develop efficient host platform production recombinant proteins, chemicals, and other compounds as reviewed by (HOLLENBERG; GELLISSEN, 1997); (STÖCKMANN et al., 2009)(Paper I). Advances in molecular biology tools, optimization of transformation protocols, and cultivation strategies in bioreactors led to the development of industrial processes based on *O. polymorpha* to produce biopharmaceuticals. Nowadays, three commercially available hepatitis B vaccines are produced using antigens produced by fermentation processes using *O. polymorpha*: HepavaxGene® (Johnson & Johnson), Gen Vax B® (Serum Institute of India) and Biovac-B® (Wockhardt)(http://www.dynavax.com/about-us/dynavax-gmbh/). Also, biopharmaceuticals successfully produced by engineered *O. polymorpha* strains and commercially available include hirudin (Thrombexx®, Rhein Minapharm), insulin (Wosulin®, Wockardt), and IFNa-2a Reiferon® (Rhein Minapharm) (GELLISSEN et al., 2005). Examples of other substances produced by recombinant *O. polymorpha* strains are reported in Table 1 from Paper I; however, none of these products are available commercially.

Features of *O. polymorpha* as a microbial cell factory for developing industrial processes include fermentation in high-cell density cultures, the capacity to use low-cost substrate as carbon source, and an available defined synthetic media, and established strategies for cultivation in large-scale bioreactors (JENZELEWSKI, 2005). Usually, these strategies are based on the utilization of methanol-inducible promoters controlling the expression of genes encoding recombinant proteins. These promoters are induced by methanol and can be derepressed in the presence of glycerol and limited glucose concentrations (SUPPI et al., 2013)(DUSNY; SCHMID, 2016). Methanol-free processes using those promoters represent a feasible alternative to control the production of recombinant proteins since methanol is harmful to human health, and its high flammability represents a risk of explosion. For more detailed information regarding strategies for the cultivation of *O. polymorpha* using either methanol or glycerol and/or glucose, the reader is directed to Paper I.

It is worth mentioning that the development of metabolically engineered strains is a key step to generate efficient microbial host platforms. The availability of molecular biology tools and established protocols for transformation are the main bottlenecks during the development of genetically modified microorganisms. Several genomic editing tools are available for *O.*

polymorpha, including CRISPR (Clustered Regularly Interspaced Palindromic Repeats) technology, and different protocols to deliver plasmids were optimized in the last years.

2.1.6 Molecular tools available for *O. polymorpha* genetic manipulation

O. polymorpha has genome editing tools available for genetic manipulation and well-established protocols for efficient transformation (Paper I). These protocols include electroporation (FABER et al., 1994), protoplasts (TIKHOMIROVA et al., 1988), and lithium acetate (SOHN et al., 1999). Electroporation is the preferred method with an efficacy of 1.7 x 10^6 clones/µg of plasmidial DNA versus 2 to 3 x 10^4 clones/µg of DNA using the protoplasts protocol. Some genomic editing tools have been described for genetic manipulation of *O. polymorpha*, whether to introduce exogenous genes or make deletions of endogenous genes. Episomal plasmids have been reported to be mitotically unstable in this yeast (BOGDANOVA; AGAFONOV; TER-AVANESIAN, 2000) and, consequently, are not suitable for the development of industrial strains. These episomal plasmids contain sequences of autonomous replication of *O. polymorpha* (HARS) derived from subtelomeric regions (SOHN et al., 1996). However, HARS sequences do not guarantee stability for circular plasmids across cell divisions. Thus, integrative plasmids are more suitable for the genetic manipulation of *O. polymorpha*.

A set of integrative plasmids applied to introduce genes in *O. polymorpha* genome is the pHIP series (SARAYA et al., 2012). Several promoters and selective markers (Table 3) are available in different pHIP plasmids. They have an easy terminology in which the letter indicates the selective marker, and the number indicates the promoter. For example, the pHIPH4 plasmid has the pAOX promoter (also named pMOX) designated by the number 4 and the letter "H" indicating hygromycin B as the selective marker. The available promoters in the pHIP series include pAOX, pAMO, pTEF, pTEF2, pCAT, pDHAS, pADH1, among others. For selection, auxotrophic and dominant markers, including *Sc-LEU2*, *Hp-URA3*, *Hp-ADE11*, *Hp-MET6*, *Sh-ble* (zeocin), *Sn-nat1* (nourseotricin), *Kp-hph* (hygromycin B), and *Tn-KanMX* (G-418/gentamycin). The pHIP plasmids are summarized in Table 3, and they are available upon request (https://www.rug.nl/research/molecular-cell-biology/research/the-hansenula-polymorpha-expression-system).

Table 3. The series of plasmids pHIP and their respective features.

Plasmid	*Promoter*	*Marker*
pHIPH4	*pAOX*	Hygromycin
pHIPN4	*pAOX*	Nourseotricin
pHIPN7	*pTEF*	Nourseotricin
pHIPZ7	*pTEF*	Zeocin
pHIPH15	*pDHAS*	Hygromycin
pHIPZ15	*pDHAS*	Zeocin
pHIPN18	*pADH1*	Nourseotricin
pHIPZ18	*pADH1*	Zeocin

The deletion of genes in *O. polymorpha* is performed through homologous recombination via deletion cassettes or using the CRISPR/Cas9 system (Paper I). The first approach consists of utilizing deletion cassettes containing a selective marker flanked by homologous regions to a target gene. Thus, by homologous recombination, the target gene is replaced by the marker. The disruption frequencies obtained have values of approximately 35% with homology arms from 500 bp to 1000 bp through this approach. The disruption frequency can be improved utilizing a strain with the *KU80* gene deleted (**Table 2**; Paper I). This locus' deletion increases repair by homologous recombination instead of non-homologous extremity union (NHEJ), ensuring specific integration in the target locus. Utilizing the same deletion cassettes with 250 bp homologous arms, the deletion efficiency was 88% in the Δ*yku80* strain versus 31% in the wild-type (SARAYA et al., 2012). This result demonstrates that using the Δ*yku80* strain is more advantageous for deleting genes in *O. polymorpha*. Lastly, the CRISPR/Cas9 technology is available as a powerful tool for genome editing in *O. polymorpha*, including insertion and disruption of genes (**Table 3**, Paper I). Three independent CRISPR/Cas9 systems were developed for *O. polymorpha* (NUMAMOTO; MAEKAWA; KANEKO, 2017); (JUERGENS et al., 2018); (WANG et al., 2018). Details for each system are reported in Paper I.

Novel molecular biology tools are constantly being developed to improve the genetic manipulation of microorganisms. Recently, integrases are being applied to the design of genetic switches to control gene expression (GOMIDE et al., 2020). Ints are site-specific recombinases that recognize specific DNA sequences designated as attachment (*att*) sites and cause rearrangement of DNA fragments based on those sequences. Different types of DNA rearrangements can be performed by integrase depending on the type and position of the *att*

sites utilized (details in (MERRICK; ZHAO; ROSSER, 2018). For instance, genetic elements (promoter, coding sequence, and/or terminator) flanked by the *att* sites *attP* and *attB* are fliped at 180° by the act of an integrase resulting in the sites *attR* and *attL* (Figure 4). One of the most important characteristics of these integrases is their unidirectional activity since the DNA fragments can only be reverted to the previous state with the activity of accessory protein named recombination directionality factor (RDF). Ints' utilization allows the design of logic gates to control the expression or inactive a gene in specific conditions (MERRICK; ZHAO; ROSSER, 2018). Therefore, Ints are powerful synthetic biology tools to design microbial cell factories to produce desired products under determined conditions since the expression/inactivation of genes is conditioned to the integrase activity. Here, serine-type phage integrase-13 (Int13)(YANG et al., 2014) was selected based on previous studies demonstrating its functionality in different eucaryotic systems, including HEK 293T cells, bovine fibroblast, and plant protoplasts (GOMIDE et al., 2020).

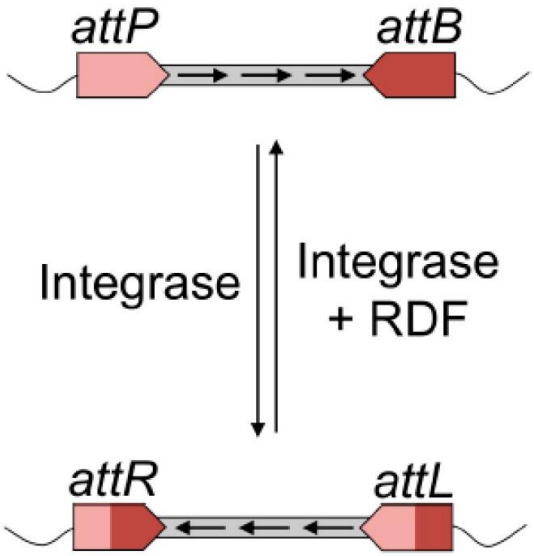


Figure 4. Example of the utilization of integrases to cause DNA rearrangement. In this case, the *attP* and *attB* sites are utilized, and by the action of an integrase, the flanked DNA fragment is flipped at 180°. Adapted from (MERRICK; ZHAO; ROSSER, 2018).

Therefore, the availability of different molecular biology tools, well-established protocols for transformation and cultivation in bioreactors, and distinct features highlight the potential of *O. polymorpha* as a reliable microbial cell factory to produce high-value products, including HA. Here, *O. polymorpha* was engineered for HA production by inserting the *hasB* gene for UDP-GlcUA synthesis concomitantly with the HAS encoding gene. Two microbial HAS, from *S. zooepidemicus* (*hasAs*) and *P. multocida* (*hasAp*) were chosen and integrated into *O.*

polymorpha genome separately. For the first time, it was also assessed the possibility to develop a genetic switch using a serine-integrase in *O. polymorpha* representing a novel synthetic biology tool for this non-conventional yeast (Paper II).

2.2 RESULTS AND DISCUSSION

Details regarding the methodology applied in this section are described in **Material and Methods** from Paper II, including the cloning strategy (*2.2. Construction of Plasmids and Strains*), protocols for preparation of electrocompetent cells (*2.4. Electrocompetent Cell Preparation and Transformation*), scanning electron microscopy (SEM) analysis (*2.7. Analysis through Scanning Electron Microscopy*) and HA quantification (*2.8. HA Quantification*). Furthermore, a summary of all strains constructed in this study is represented in Figure 5, including the promoters and the selective markers utilized for all EMB strains.

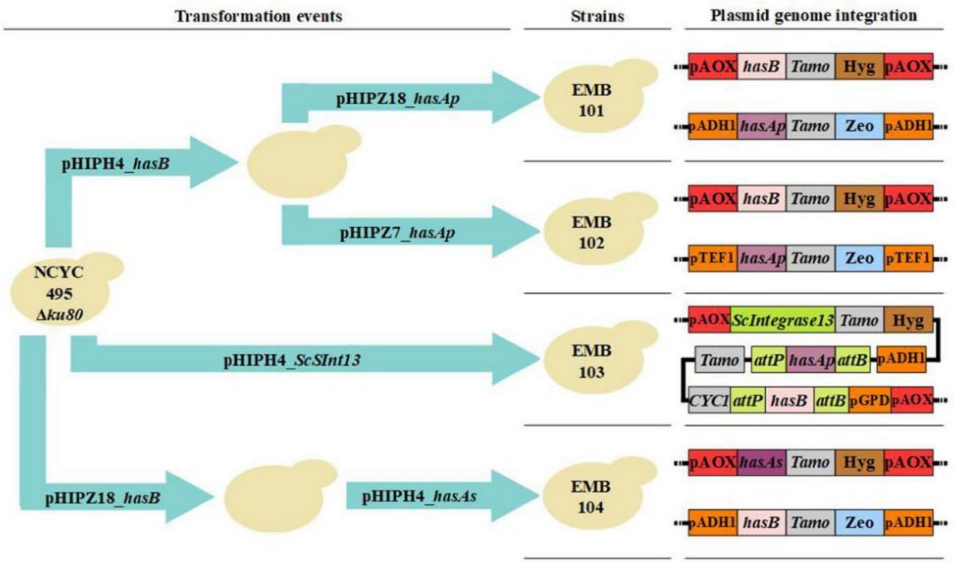


Figure 5. *O. polymorpha* strains constructed in this study for HA production (Paper II). Each blue arrow represents an event of transformation. The name inside the blue arrows corresponds to the plasmids utilized for the respective transformation. The *O. polymorpha* strains were named EMB101, EMB102, EMB103, and EMB104, and the respective genotype is represented in the last panel "Plasmid genome integration".

2.2.1 Construction of *O. polymorpha* strains to produce HA

In this work, four strains were constructed (Figure 7), inserting both *hasA* and *hasB* genes into *O. polymorpha* NCYC495 Δ*yku80* strain (SARAYA et al., 2012), which is designated here as the wild-type (WT) strain. The genes *hasB* from *X. laevis* and *hasA* from *P. multocida* (named as *hasAp*) were selected since both genes encoded active enzymes in other non-conventional

yeast *K. lactis* constructed for HA production (GOMES et al., 2019). The *hasA* from *S. zooepidemicus* (named as *hasAs*) is commonly utilized to generate heterologous strains for HA production, and it was selected to evaluate the expression of this gene in *O. polymorpha* as well as to compare the HA titers obtained with the cells expressing *hasAp* gene. The Figure 5 summarizes the genotype for each strain. Both *hasA* and *hasB* were confirmed by colony PCR as showed in the Supplementary Material from Paper II. The figures S1, S3, S4, and S5 represent the agarose gels confirming the presence of *hasA* and *hasB* in the strains EMB101, EMB102, EMB103, and EMB104, respectively.

In the case of EMB101, it was not possible to detect the *hasAp* gene by colony PCR after successive passing in YPD agar plates containing zeocine (selection of *hasAp*) and hygromicin (selection of *hasB*) (Figure 6). This strategy of successive passing in selective medium was applied previously in *O. polymorpha* (SOHN et al., 1996). This phenomenon was not investigated further in this work since it is out of the current study's scope. The *hasAp* was cloned in the pHIPZ18, which is widely applied to study peroxisome biogenesis in *O. polymorpha* (AKŞIT et al., 2018), and no report for its utilization for heterologous production was described so far. Therefore, the EMB101 was discarded from the experiments related to HA production. However, investigating the possible mechanisms involved in the loss of *hasAp* gene could help understand the stability of heterologous genes integrated into the locus ADH1 in *O. polymorpha*.

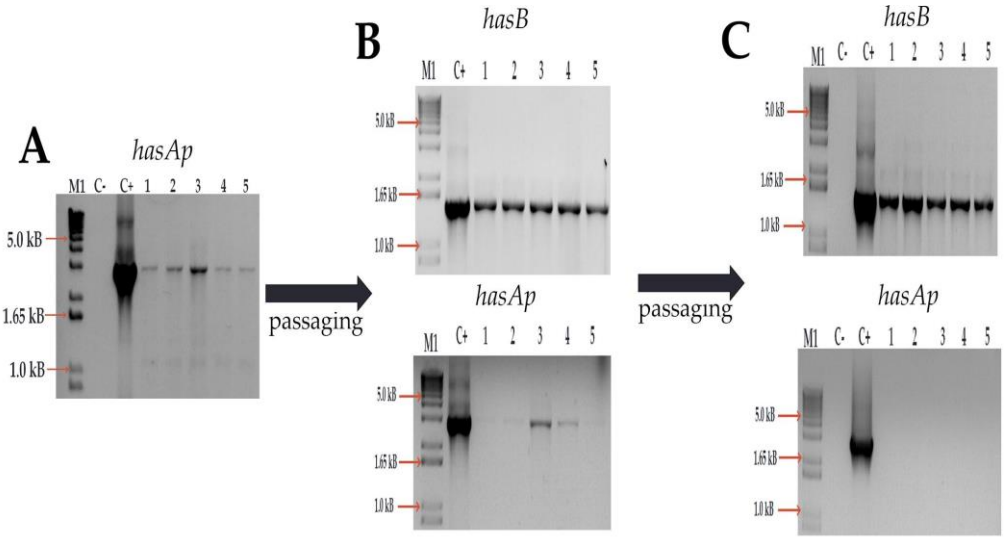


Figure 6. Colony PCRs to verify the *hasAp* (2966 bp) and *hasB* (1505 bp) genes stability in the genome of EMB101 strain after three successive passages on YPD agar plates supplemented with zeocin and hygromycin. Five colonies were selected after the transformation and plating on YPD plate containing both antibiotics for the verification of *hasB* and *hasAp* stability. In all PCRs reactions performed, the *O. polymorpha* NCYC495 *yku80*

was used as a negative control (C-) and the pHIPH4_*hasB* and pHIPZ18_*hasAp* plasmids were used as positive controls (C+) for the of *hasB* and *hasAp*, respectively. The fragments amplified at the height of 1.5 kb in the upper gels correspond to the *hasB* gene and the fragments in the height of 3.0 kb at the bottom gels correspond to the *hasAp* gene. The black arrows indicate a passaging to another YPD plate supplemented with zeocin and hygromycin. The agarose content was 1% and the ladder used was 1 kb Plus DNA Ladder (M1; Invitrogen). (A) Colony PCR for the *hasAp* gene of the first passaging on selective medium, (B) Colony PCR for the *hasB* (upper gel) and *hasP* (bottom gel) genes of the second passaging on YPD plate with both antibiotics. (C) Colony PCR for the *hasB* (upper gel) and *hasAp* (bottom gel) genes of the third passaging on a YPD plate with both antibiotics.

For EMB103 strain, a unidirectional genetic switch in which the expression of *hasAp* and *hasB* is controlled by the activity of a serine integrase was evaluated. To control the expression of both genes, the synthetic plasmid pHIPH4_*ScSInt13* containing a codon-optimized gene for *S. cerevisiae* encoding Int13 under control of the methanol inducible promoter pAOX and both *hasAp*/*hasB* genes was inserted in *the O. polymorpha* NCYC495 Δ*yku8*0 strain genome. The Int13 recognizing sites *attB* and *attP* were inserted flanking both genes *hasAp* and *hasB* that were synthesized in reverse complement orientation. After the addition of methanol, the Int13 is produced, acting in the *attB* and *attP* sites and causing a 180° rotation in the DNA fragment inserted between these two sites (in this case, the *has* genes). The resulting fragment contains the sites *attL* and *attR* sequences, and both *has* genes in the ORF with the promoter and termination allowing their expression. Figure 7A illustrates the action of Int13 in the system employed here, and Figure 7B shows the agarose gels demonstrating the correct rotation of *hasAp* and *hasB* after methanol induction confirmed by PCR. More details about the system are described in the Figure S4 legend from Paper II.

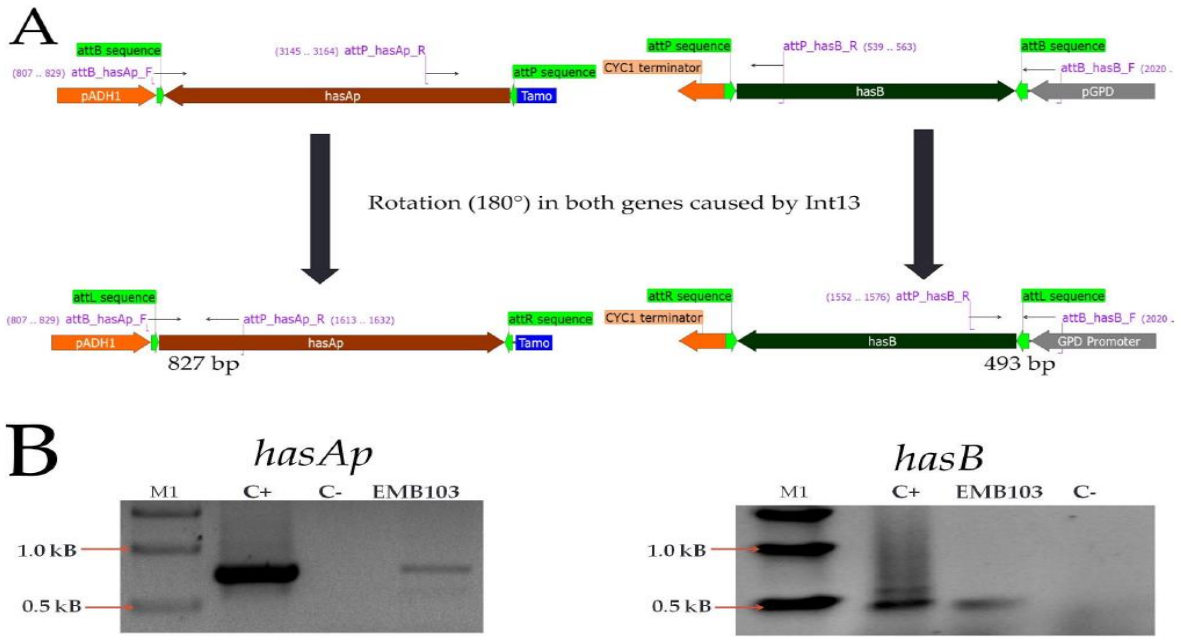


Figure 7. Scheme representing the genetic switch constructed to control the expression of both *hasB* and *hasAp* by a serine integrase (details described in Paper II). (A) The gene encoding the Int13 codon-optimized for *S. cerevisiae* is regulated by the promoter pAOX which is inducible by methanol. Thus, the addition of methanol

leads to the production of Int13 that recognizes the sites *attB* and *attP* flanking both has genes synthesized in reverse complement orientation. The action of Int13 causes a rotation of 180° in both genes resulting in two different flanking sequences named *attL* and *attR*. In the final, both genes are in ORF with promoter and terminator and can be properly transcripted. The correct gene orientation as well the formation of *attL* sequence could be evaluated by PCR using the pair of primers (the letters in purple indicate the name and the annealing positions of the primers utilized) attB_hasB_F and attP_hasB_R for *hasB* (resulting fragment of 493 bp), and attB_hasAp_F and attP_hasAp_R for *hasAp* (resulting fragment 827 bp). When the genes are in the initial orientation, the primer annealing fails, and no amplification occurs. The black arrows represent the primer orientation. (B) Confirmation by PCR of both genes rotation in the EMB103 strain after methanol induction. For all PCRs reactions, the genomic DNA was utilized as the template. The gel in the left shows the fragment (827 bp) containing the *attL* sequence and the *hasAp* flipped and in the right the fragment for *hasB* (493 bp). The *O. polymorpha* NCYC495 *yku80* was used as a negative control (C-) for the PCR reactions. The pHIPZ18_*hasAp* plasmid which contains the *hasAp* in ORF was utilized as positive controls (C+) for the PCR reactions of *hasAp*. The plasmid PKLAC2-BP constructed previously for our group (GOMES et al., 2019) was utilized as the control positive for the PCR reaction of *hasB* once all plasmids constructed in this work has the *hasB* gene controlled by endogenous promoters of *O. polymorpha* instead of pGDP promoter from *S. cerevisiae*.

The strains EMB102, EMB103, and EMB104 bearing both *hasA* and *hasB* genes confirmed by colony PCR were evaluated regarding HA production using a qualitative method by SEM analysis and a quantitative approach using the carbazole method.

2.2.2 Evaluating the HA production by *O. polymorpha* strains through SEM analysis

Firstly, the production of HA was assessed by SEM analysis since a capsule-like layer is observed. Also, this method was previously employed as a qualitative method to evaluate HA production in *K. lactis* (GOMES et al., 2019) and *L. lactis* (HMAR et al., 2014). A capsular layer was also observed in *C. glutamicum* engineered for HA production, but in this work, a phase-contrast microscopy analysis was employed to evaluate the capsule-like structure around the cells (WANG et al., 2020). Therefore, microscopy analysis is a feasible approach to detect the HA production in heterologous hosts.

The EMB102, EMB103, and EMB104 strains were prepared for HA production, separating the cultivation into two steps: (i) Biomass production and (ii) HA production. The first step also includes the induction of *has* genes by adding methanol to the medium, which induces the expression of genes regulated by the promoter pAOX. A growth curve of the cells during the biomass production is provided in Figure 8.

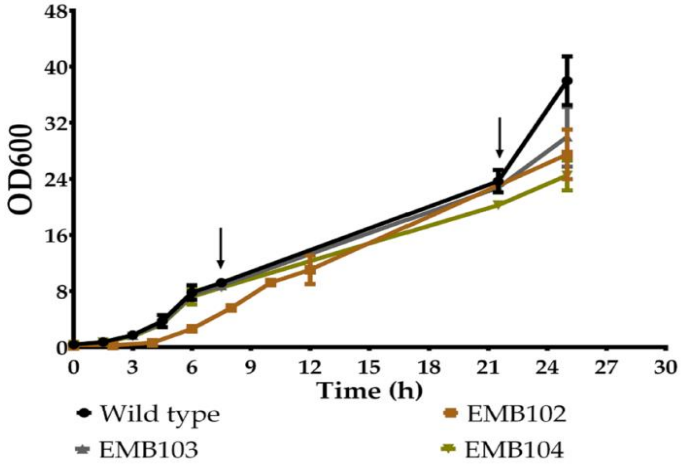


Figure 8. Growth profile of EMB102, EMB103, and EMB104 and *O. polymorpha* NCYC495 *yku80* (WT)strain grown on 1-L baffled shake flasks containing 100 mL of YPD medium and kept at 37 ◦C under agitation (200–250 rpm). The black arrows represent the addition of methanol to the medium in a final concentration of 1%. Means and standard deviations were calculated from two biological replicates.

The entire high-cell density cultures ($OD600_{nm}$ ≈ 30-40) were transferred to fresh YPD for HA production, and after 24, h the cultivation was ended. The cells were prepared for the SEM analysis, and the images are shown in Figure 4 and Figure 5 from Paper II. A discrete difference was observed between the WT strain and the strains EMB102 and EMB104, regarding the cell aggregation (Figure 4A, Paper II) and cell surface (Figure 4B, Paper II). In contrast, the EMB103 cells are different from the other strains. The cells from EMB103 show a higher level of aggregation that appears to be related to a substance that sticks the cells together (Figure 4, Paper II). Besides, the cell surface of EMB103 seems to be covered by a substance absent in the WT strain (Figure 5, Paper II). The same phenotype was observed by the *K. lactis* strains producing heterologous HA (GOMES et al., 2019), indicating that the substance observed on the surface of EMB103 cells is consistent with HA. Therefore, the production of HA by EMB103 is clear, while for the strains EMB102 and EMB104, the SEM results were inconclusive regarding the HA production. Furthermore, the successful synthesis of HA by EMB103 demonstrated the application of a genetic switch using a Int13 in *O. polymorpha*, representing a novel synthetic biology tool for this non-conventional yeast. Next, the production of HA by all strains was evaluated using a medium optimized for HA production (IM et al., 2009).

2.2.3 Production of HA by *O. polymorpha* using an optimized medium

The strains EMB102, EMB103, and EMB104 were evaluated regarding the HA production using a medium optimized for HA production. The same medium was also

employed in the methylotrophic yeast *K. phaffii* for HA production (JEONG; SHIM; KIM, 2014b). The high cell density cultures were obtained as described above (Figure 8), including the addition of methanol for the induction of genes under the control of the pAOX promoter. the cells were transferred to the optimized medium instead of YPD. The cells were cultivated in 1-L baffled shake flasks containing 100 mL of the optimized medium and grown at 37 °C/200 rpm for 48 h. To avoid glucose starvation, glucose was added to 2% during the cultivation based on the $OD600_{nm}$ measured. Samples were taken at 24 h for SEM analysis, and all culture broth was prepared for HA quantification. For this, the cells must be treated with SDS (addition of 0.1% SDS in equal volume) that removes the HA capsule present on the cell surface. Subsequently, the HA is precipitated and purified from the culture broth, allowing its quantification by the carbazole method (adapated from BITTER; MUIR, 1962)(details in the section **Material and Methods**: *2.8. HA Quantification* from Paper II). Utilizing this approach, it was possible to detect HA in the supernatant of all *O. polymorpha* constructed (Table 4).

Table 4. The HA titer obtained by EMB strains after 48 h of cultivation in an optimized medium.

Strains	HA titer (µg/mL)	Final OD_{600nm}
WT	NA	21.00 ± 1.41 [a]
EMB102	151.20 ± 13.04 [a,b]	27.00 ± 0.00 [a]
EMB103	197.76 ± 5.66 [b]	27.50 ± 0.71 [a]
EMB104	123.20 ± 26.56 [a]	21.50 ± 3.54 [a]

NA: Not applied. Equal letters indicate no statistical difference ($p < 0.05$). The statistical analysis applied here is detailed in the section **Material and Methods**: *2.6. Statistical Analysis and Data Presentation* from Paper II and in the legend of Table 2 from Paper II. Means and standard deviations were calculated from two biological replicates.

The highest HA concentration detected in the broth occurred in the strain EMB103, although these values were statistically significant only comparing this strain with EMB104. Nevertheless, the production of HA by all *O. polymorpha* strains developed was achieved, corroborating that this yeast is a suitable platform host for heterologous production of this high-value biopolymer. It is worth mentioning that these values are well below those obtained by other heterologous producers. Besides, the values are inferior comparing with other yeasts, including *K. lactis* (1.89 g/L) (GOMES et al., 2019) and *K. phaffii* (1.7 g/L) (JEONG; SHIM; KIM, 2014), although in those examples, the cultivations were performed in bioreactors which

allows a better control of fermentation parameters including pH, culture feeding, temperature, agitation, and dissolved oxygen (DO) that needed to be adjusted for an efficient HA production (JOHNS; GOH; OEGGERLI, 1994; LIU et al., 2009; PAN et al., 2015). Thus, several parameters could be tested to improve the HA production by engineered *O. polymorpha* strains. In this perspective, the cultivation of EMB102, EMB103, and EMB104 in bioreactors is crucial to understand the impact of such parameters in HA synthesis. It represents a goal for a follow-up study aiming to improve HA titers obtained by *O. polymorpha* constructed here. Additionally, further metabolic engineering approaches to delete or downregulation genes (e.g., *zwf* and *pfka*) from competitive pathways as well as the overexpression of genes from the HA pathway (Figure 3) are potential targets to improve the HA titers obtained by EMB102, EMB103, and EMB104 (details in the section **Results and Discussion**: *3.4. HA Quantification* from Paper II).

The SEM analysis also confirmed the production of HA by EMB102, EMB103, and EMB104 (Figure 9). In these assays, the samples were taken after 24 h of cultivation in the optimized medium, and the cells were treated with SDS previously the preparation for the SEM. Since the SDS removes the capsular structure around the cells, it was evaluated if any difference is observed between the cells treated with SDS and without the treatment. The treatment consists of adding 0.1% SDS in equal volume to the culture broth and keeping the samples at room temperature under agitation for 10 min to remove the HA capsule from the cells. This method represents a novel approach to confirm HA production by microorganism producing low titers of this biopolymer.

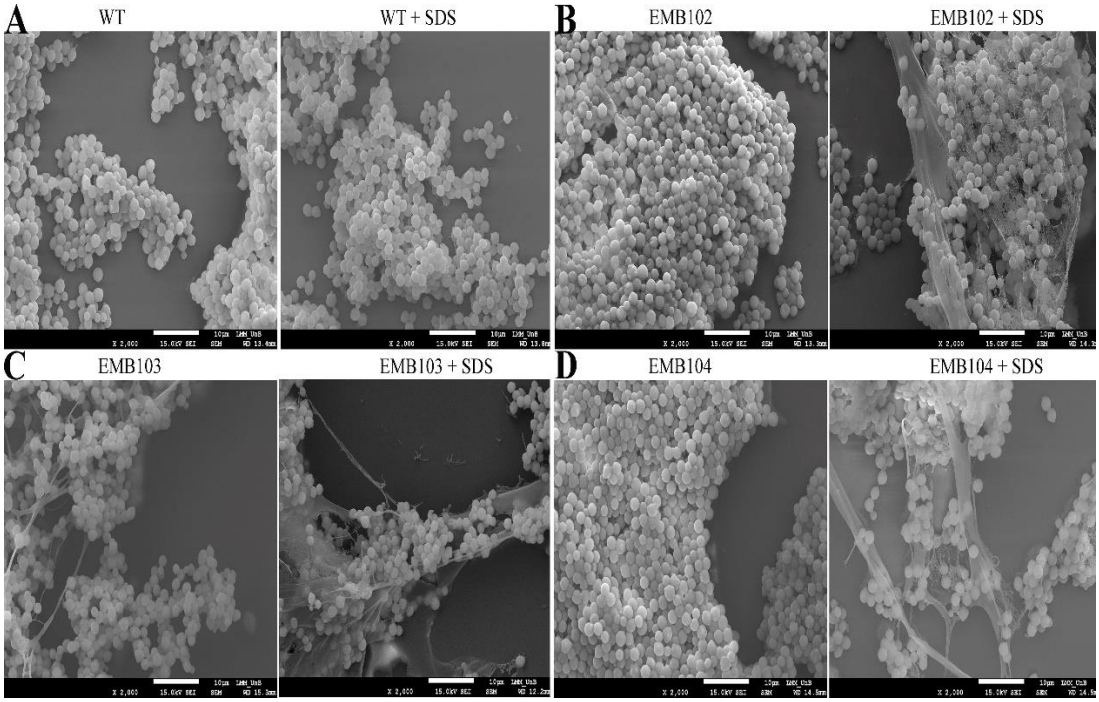


Figure 9. SEM analysis of *O. polymorpha* strains treated with SDS and without the treatment to remove the HA capsule from the cells. The names above each image correspond to the respective strain, and the nomenclature "+SDS" indicates that the cells were treated with SDS. (A) WT (left) and WT+SDS (right); (B) EMB102 (left) and EMB102 + SDS (right); (C) EMB103 (left) and EMB103 + SDS (right); (D) EMB104 (left) and EMB104 + SDS (right). The images were acquired in the magnitude of 2, 000X.

Regarding the images from the cells treated with SDS, it is evident the production of HA by EMB102 (Figure 9B), EMB103 (Figure 9C), and EMB104 (Figure 9D), corroborating the results from the HA quantification in the culture broths. The HA is visualized in these images as a net-like structure observed only after the treatment with SDS. The same structure is also visualized in EMB103 without the treatment (Figure 9C, panel on the left). Probably the HA is removed from the cells but retained in cell agglomerates allowing the visualization of these structures. Additionally, these results also indicate that HA extraction was not efficient, and optimization regarding the SDS treatment (e.g., time of treatment and/or concentration) is necessary to remove the HA capsule completely from the cell surface.

2.3 Conclusion

The production of HA by engineered strains of *O. polymorpha* was demonstrated. The insertion of both *hasA* and *hasB* genes into the genome of this yeast generated strains that produce HA. For the *hasA* gene, both *hasAp* from *P. multocida* and *hasAs* from *S. zooepidemicus* led to the production of HA by the cells, enabling the implementation of different strategies for heterologous production of HA in *O. polymorpha*. Additionally, a

genetic switch using a serine-type integrase-13 (Int13) to control the expression of two genes concomitantly was successfully implemented. In this work, this system was employed to control both *hasAp* and *hasB* by adding methanol to the culture medium. To the best of our knowledge, this is the first report of a genetic switch applied for gene regulation in *O. polymorpha*.

The synthesis of HA was confirmed by a qualitative approach as well as a quantitative method. The first applied an SEM analysis observing a capsule-like structure on the cell surface and a net-like structure connecting the engineered strains' cells. The HA production was confirmed by purifying the HA from the supernatant and quantifying it by the carbazole method. For efficient production, a medium that was previously reported as an optimized medium for HA production by microorganisms as utilized. Indeed, the engineered *O. polymorpha* strains in this medium allowed the production of HA for all strains.

The EMB103 produced the highest HA titer (197.76 μg/mL), which is considered a low concentration of HA compared with other heterologous producers, including the yeasts *K. phaffii* and *K. lactis*. However, the production can be improved by either optimizing the fermentation parameters (e.g., DO, aeration, pH) or by genetic engineering to funnel the metabolic framework of *O. polymorpha* for HA production. Furthermore, the EMB103 strain has both *has* genes regulated by constitutive promoters, while for EMB102 and EMB104, the methanol-inducible promoter pAOX was utilized to control the expression of *hasB* and *hasAs*, respectively. Thus, the strategy for methanol induction during the step of biomass production is a potential target to be considered for improving HA production in these strains. Nevertheless, as a proof-of-principle, the production of HA was demonstrated by *O. polymorpha*, although the process still needs optimization to improve the production.

3. Chapter 3

Production of aromatic amines from lignin monomers by Pseudomonas putida strains

3.1 Introduction

3.1.1 Lignin valorization through whole-cell biocatalysis

Lignin, the second most abundant carbon-based bio-material on the planet, has great potential to become a major source of building-block chemicals currently obtained from fossil origins (GILLET et al., 2017)(HALDAR; PURKAIT, 2020). The pulp and paper industry is responsible for the production of approximately 50 million tons/year of lignin (SUN et al., 2018b), which due to lack of economically sustainable alternatives is incinerated to produce energy, leading to an increase in the production of greenhouse gas and wasting its potential as raw material. Lignin is isolated from plant biomass by different physicochemical or enzymatic treatments producing different technical lignins, a valuable source of different platform chemicals (Figure 10).

The development of novel technology for efficient bio-refining of this recalcitrant and heterogeneous aromatic polymer into building-block chemicals may, however, change the current practice and could improve the economy of the forest-based industries (SCHUTYSER et al., 2018). The bio-refining and utilization of different aromatic compounds derived from depolymerized technical lignin are usually referred to as *Lignin Valorization* (Figure 10). Many platform chemicals obtained by petroleum refining can be produced through lignin depolymerization by chemocatalytic or enzymatic treatments. However, the depolymerization of technical lignin into monomers is still a challenge, and different aromatic compounds can be obtained depending on the technical lignin utilized or the treatment employed. The bulky chemicals obtained can be used to produce high-value compounds by chemical synthesis and/or microbial biotransformation.

Utilizing microorganism's natural pathways or metabolically engineering, it is possible to develop host platforms to produce chemicals of interest in a greener way from technical lignins. The interplay between lignin depolymerization and one-pot biocatalysis has been explored recently for lignin transformation into a high-value product, such as nylon (DUUREN et al., 2020)(KOHLSTEDT et al., 2018). Therefore, the possibility to produce high-value products from renewable feedstocks through whole-cell biocatalysis is the milestone for the genetic engineering of microorganisms (HALDAR; PURKAIT, 2020). Several compounds can be obtained from monolignols, including aromatic amines.

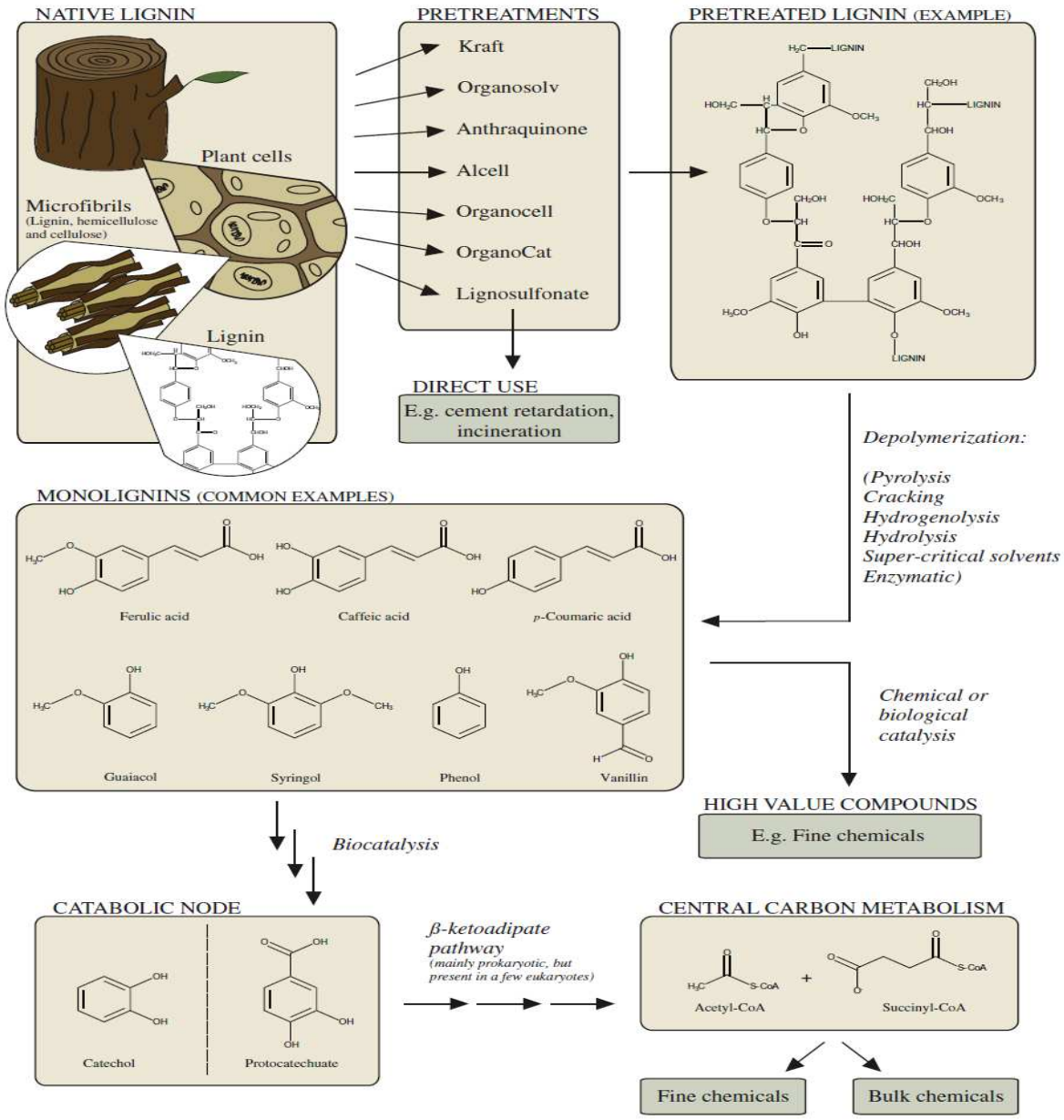


Figure 10. Pipeline for lignin valorization starting from plant raw material. Figure obtained from (ABDELAZIZ et al., 2016).

3.1.2 Aromatic amines: applications and obtention

Aromatic amines are important chemical building blocks for the production of various commodity products (KELLY et al., 2018), including pharmaceuticals (Figure 11) and functionalized materials with applications in the automobile, aerospace, building, and health industries (FROIDEVAUX et al., 2016). Most aromatic amines are derived from platform chemicals obtained from petroleum refining, and production routes from renewable resources are highly desirable (SHELDON; WOODLEY, 2018). A promising process strategy is to

combine chemo- and bio-catalytic methods, where the role of the first method is to fractionate and convert biomass to renewable platform chemicals suitable for further enzymatic or microbial bio-refining into aromatic amines. Many lignin-derived platform chemicals are suitable for further refinement into aromatic amines (BLONDIAUX et al., 2019). To obtain amines from carbonyl compounds, amine transaminase (ATA) can replace the traditional method using transition metal catalysis. It is now the preferred approach in the pharmaceutical industry (KELLY et al., 2020).

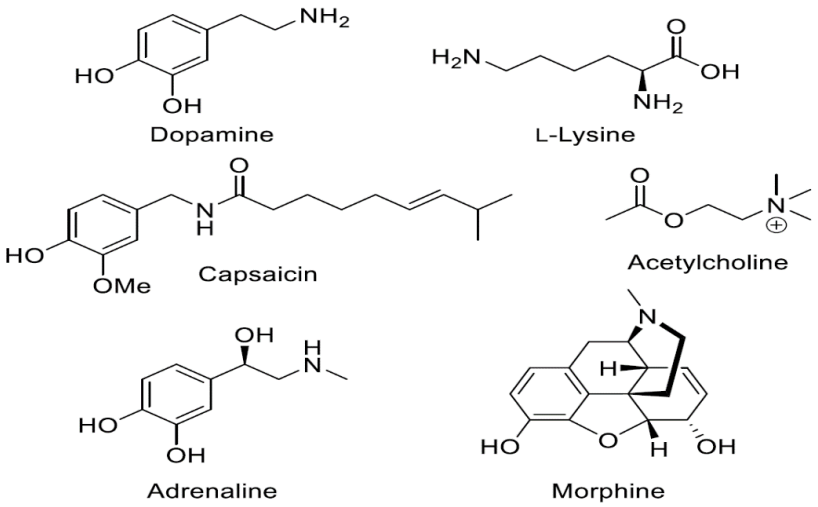


Figure 11. Examples of important amine compounds. Figure obtained from (ANDERSON, 2017)**.**

3.1.3 ATAs features and their potential for lignin valorization

Amine forming reactions have large potential to be integrated with existing production routes, resulting in process optimization to reach target products. In light of this, industrial lignocellulosic residues represent an important source of different platform chemicals suitable for transformation into aromatic amines through biocatalysis (Figure 10). These compounds can be reached through a transamination reaction substituting a carbonyl group with an amino group in an aldehyde, carboxylic acid, or ketone producing an amine. This reaction is performed by amine transaminases (EC 2.6.1.x), which are pyridoxal 5'-phosphate (PLP)-dependent enzymes catalyzing the reversible transamination through a ping-pong bi-bi mechanism (PATIL et al., 2018). The amino group is transferred from an amino donor to an amino acceptor with the corresponding amine's formation and releasing a by-product (Figure 12). For instance, when alanine is the amino donor, pyruvate is produced as the by-product, and for isopropylamine as the amino donor, acetone is formed instead (GOMM; O'REILLY, 2018).

The optimal conditions for the reaction (temperature and pH), stereoselectivity, the spectrum of substrates recognized, as well as the enzymatic affinity for those substrates diverge across the ATAs. Furthermore, the binding strength to PLP seems to affect its activity, and has been studied to improve enzymatic performance (PADROSA et al., 2019). The emergence of novel ATAs especially from metagenomics libraries evidences its importance for the production of amines (COSCOLÍN et al., 2019)(PAWAR; HALLAM; YADAV, 2018).

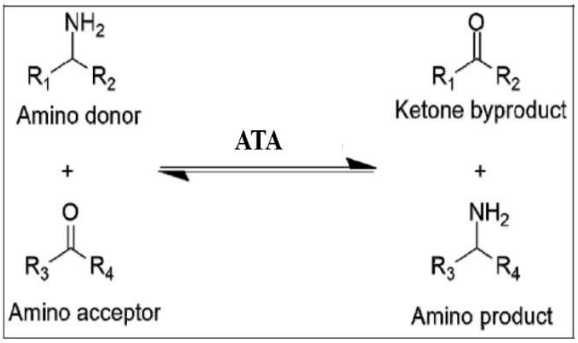


Figure 12. Transamination reaction of an Amino acceptor by transferring an amino group from an Amino Donor production a by-product (in this a ketone) and an amine. ATAs perform the reversible reaction. Figure adapted from (PATIL et al., 2018).

The ATAs are classified based on the relative position of the functional group to be transferred. The α-transaminases (α-TAs) act in groups at α position while the ω-TA recognizes groups on non-α positions (STEFFEN-MUNSBERG et al., 2015). Nevertheless, unfavorable equilibrium and product inhibition are disadvantageous features shared between ATAs, affecting the reaction yield and representing the main bottleneck to their utilization at industrial level. Several approaches have been developed recently to bypass this issue, as previously discussed (GUO; BERGLUND, 2017). Among those, the enzymatic removal of co-products and co-substrate (amine donor) utilization in excess are the most employed. *In vivo* strategies for co-product removal often involve cascade reactions that have dual functionality: (1) remove the co-product, and (2) recycle the co-substrate and the cofactors (WU et al., 2016). For instance, reactions with ATAs that use alanine as an amine donor are favored in the presence of the enzyme alanine dehydrogenase (AlaDH) that performs the conversion of pyruvate to alanine using ammonium as the cofactor. Indeed, concomitant expression of an ATA encoding gene from *C. violaceum* (*CV-ATA*) and *AlaDH* gene from *B. subtilis* efficiently drive the reaction to product formation (ZHOU et al., 2018)(WU; LIU; LI, 2017). The CV-ATA enzyme is commonly used for transamination reactions due to its intrinsic characteristics as the activity

on a wide range of substrates, including lignin-derived compounds (DU et al., 2014). One example of aromatic amine that can be obtained by biocatalysis of monolignols is vanillylamine (VA).

Regarding lignin-derived aromatic amines, VA has obtained significant attention, due to its potential to be produced from vanillin, which is obtained from depolymerized technical lignin, through a reversible transamination reaction. It is applied to produce polyepoxides (MOGHEISEH; KARIMIAN; KHOSHSEFAT, 2020) and as a building block chemical to produce capsaicinoids (ANDERSON et al., 2014). For epoxy polymers synthesis, it is utilized as a hardener to provide the thermo-mechanical properties (FACHE et al., 2015). Also, VA is an intermediate in the synthesis of bioactive capsaicinoids (e.g., capsaicin), responsible for the pungent sensation in chili pepper fruits (*Capsicum* sp.). It may have value as antimicrobial (MARINI et al., 2015) and anticarcinogenic agent (XU et al., 2018), painkiller (FATTORI et al., 2016), and in the treatment of different metabolic disorders (PANCHAL; BLISS; BROWN, 2018). VA can be synthesized from vanillin using enzymatic conversion catalyzed by an ATA. CV-ATA is known to convert vanillin into VA through a transamination reaction (KAULMANN et al., 2007)(DU et al., 2014). The ATA from *Capsicum chinense* (CC-ATA) was characterized *in vitro* as a vanillin aminotransferase capable of achieving the transamination of VA to vanillin, but not the opposite reaction (WEBER et al., 2014). According to the authors, the production of VA from vanillin using the isolated enzyme and L-alanine as amine donor failed due to the thermodynamical equilibrium favoring the reverse reaction ($\Delta G0' = +2.55$ kcal/mol for vanillin to VA). Therefore, VA may, in principle, be obtained by whole-cell transamination of vanillin by a biocatalyst with engineered ATA activity.

It is worth mentioning that lignin-derived substrates are phenolic aromatic compounds with considerable cytotoxicity requiring a robust host platform for feasible biotransformation of depolymerized technical lignin. Among those platforms, the *P. putida* is emerging as a robust biocatalyst for converting depolymerized technical lignin.

3.1.4 *P. putida*: a versatile microorganism for lignin valorization

The bacterium *P. putida* has recently obtained significant attention as a promising microbial platform for biotransformation of various substrates into high-value compounds (Figure 13), as recently reviewed by (WEIMER et al., 2020). It is a gram-negative bacterium

with a high potential value for lignin valorization due to its physiological robustness, tolerance to aromatic compounds, and the presence of several well-characterized funneling pathways for assimilation of lignin-derived compounds (BRINK et al., 2019). For metabolic engineering in *P. putida,* several synthetic biology tools for markerless (and scarless) deletion of genes have been reported previously (WIRTH; KOZAEVA; NIKEL, 2020)(APARICIO et al., 2020), including a CRISPR/Cas9 system for genetic manipulation (SUN et al., 2018). Plasmids from the Standard European Vector Architecture (SEVA) repository are commonly utilized for gene expression in *P. putida* (SILVA-ROCHA et al., 2013). Therefore, all these features consolidate *P. putida* as an efficient host platform for whole-cell biocatalysis of depolymerized lignin to obtaining values products (Figure 13).

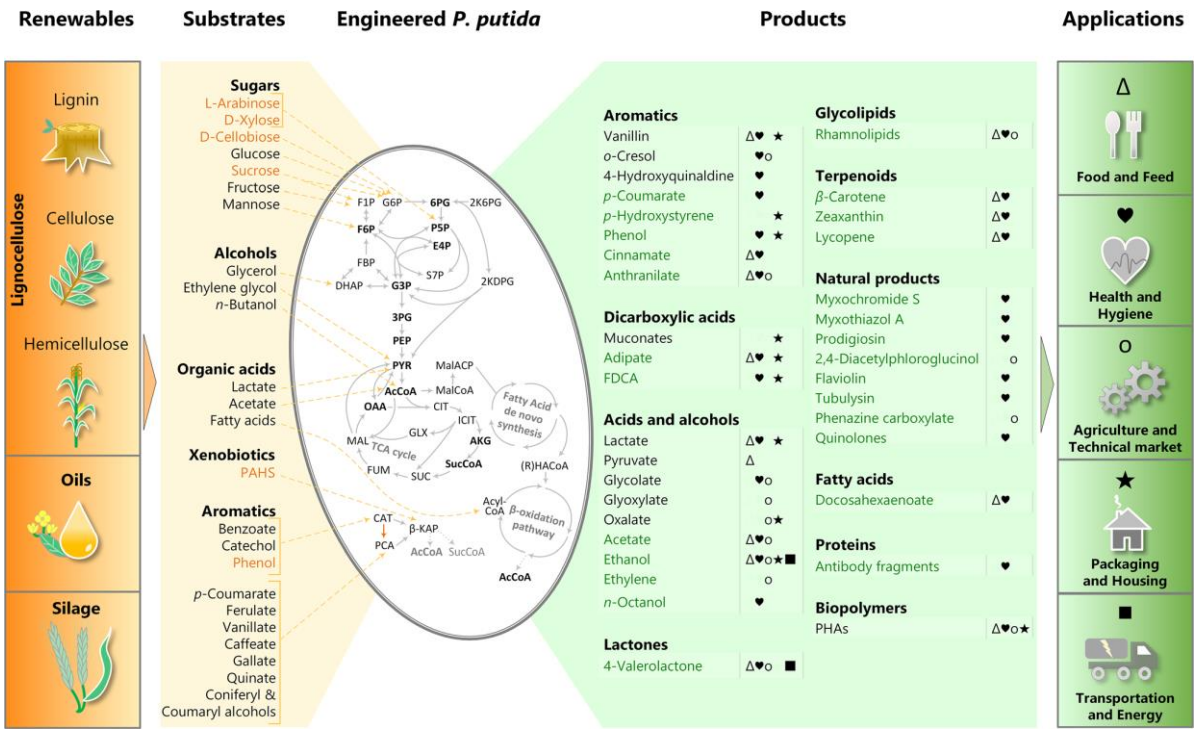


Figure 13. Overview of engineered *P. putida* strains constructed utilizing lignocellulose biomass to produce several products. Figure obtained from (WEIMER et al., 2020).

P. putida has previously been metabolically engineered to accumulate vanillin from ferulic acid (GRAF; ALTENBUCHNER, 2014). Increase the level of enzymes responsible for converting ferulic acid and prevent vanillin metabolism are two crucial points to generate a strain accumulating vanillin. Thus, the genes encoding feruloyl-CoA synthetase (*fcs*) and enoyl-CoA hydratase/aldolase (*ech*) responsible for producing vanillin from ferulic acid were overexpressed, substituting their endogenous promoters for the strong IPTG inducible P*tac*

promoter. Also, genes encoding enzymes related to vanillin metabolism into vanillic acid were targeted for disruption. The corresponding genes for vanillin dehydrogenase (*vdh*), two aldehyde dehydrogenases (PP_2680 and PP_0545), and a benzaldehyde dehydrogenase (PP_1948) were markerless deleted, generating the strain GN442 (GRAF; ALTENBUCHNER, 2014). This strain still maintains the conversion of vanillin to vanillyl alcohol by endogenous oxidoreductase. In a following study, the *calA* gene encoding a putative coniferyl alcohol dehydrogenase (PP_2426) was deleted in the *P. putida* GN442 strain (GARCÍA-HIDALGO et al., 2020). Bioconversion assays using ferulic acid as substrate demonstrated that the GN442ΔPP_2426 strain showed a higher vanillin yield (mol/mol) than GN442 (82% versus 69%, respectively), and this difference was related to lower production of vanillyl alcohol as the by-product. Thus, the *P. putida* GN442ΔPP_2426 strain is a promising candidate to produce VA from lignin-derived substrates.

In this study, *P. putida* strains were developed for perform whole-cell transamination to produce VA from the lignin monomers ferulic acid and vanillin (Figure 14). For this, two genes encoding CC-ATA and CV-ATA enzymes were cloned in *P. putida* KT2440 and GN442ΔPP_2426 strain (GARCÍA-HIDALGO et al., 2020). The strains KT2440 expressing *CC-ATA* gene (TMB-JH001), and *CV-TA* (TMB-JH002) were used to implement a growth-based approach to evaluate transaminase activity against VA. On the other hand, GN442ΔPP_2426 harboring *CC-TA* gene (TMB-JH003) and GN442ΔPP_2426 expressing *CV-ATA* (TMB-JH004) were used for whole-cell bioconversion using vanillin as a substrate to obtain VA. Furthermore, three native ATAs from *P. putida* KT2440 were identified using the protein sequence from CV-ATA and CC-ATA as query sequences. The cloning and activity characterization against VA of these three native ATAs were objects of study from Paper III. Lastly, the effect of co-expressing AlaDH encoding gene from *B. subtilis* for the regeneration of L-alanine during the production of VA from both vanillin and ferulic acid was evaluated. Although the production of VA was limited by product re-assimilation and the equilibrium towards vanillin, it was demonstrated the potential to produce lignin-derived amines by whole-cell biocatalysis from vanillin and ferulic acid using a *P. putida* strain with reduced background activity against vanillin (GN442ΔPP_2426).

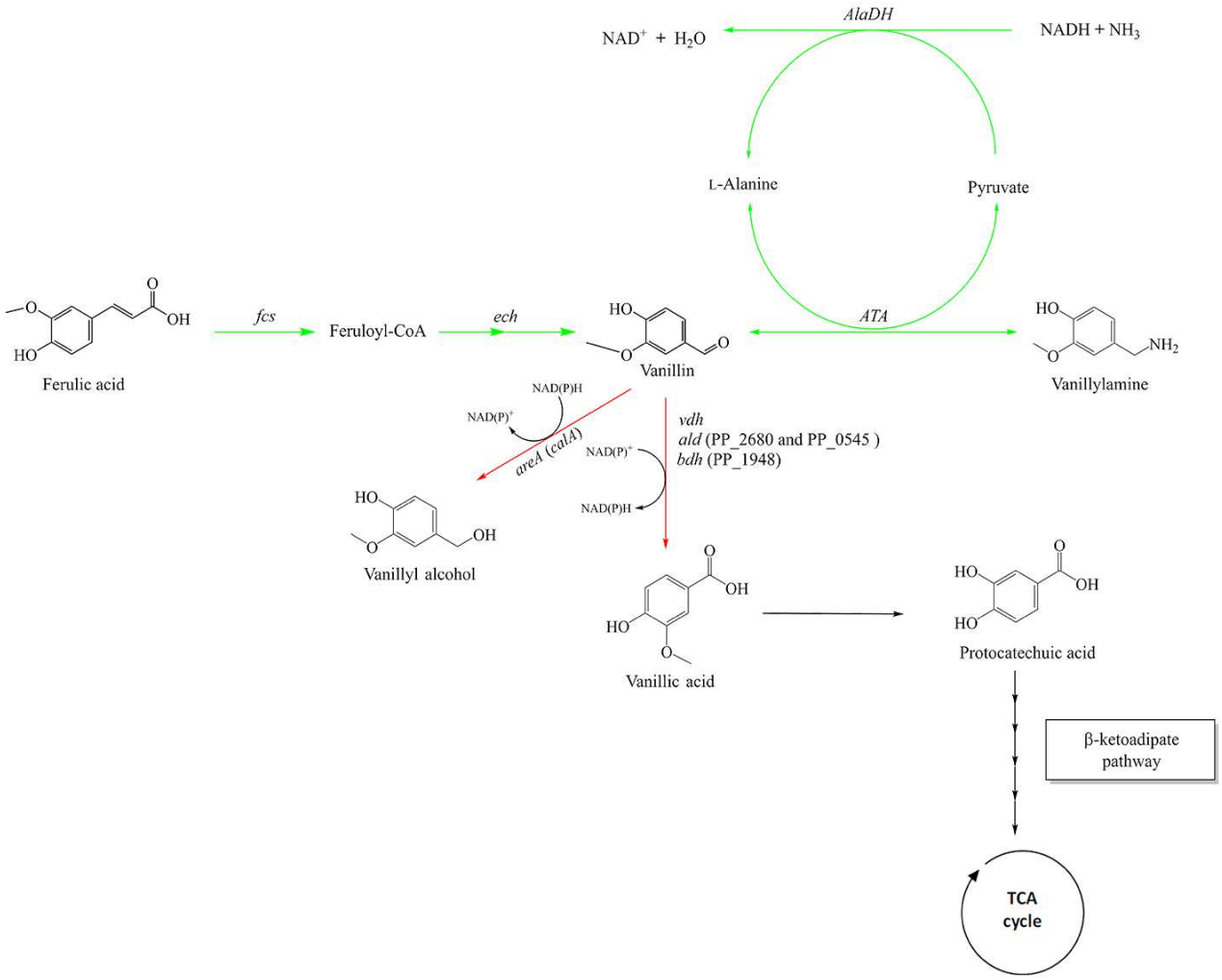


Figure 14. Engineered VA production pathway from ferulic acid using the *P. putida* strain GN442ΔPP_2426 (GRAF; ALTENBUCHNER, 2014)(GARCÍA-HIDALGO et al., 2020). The green arrows correspond to the overexpressed genes, and the red arrows to the deleted genes. Details in figure 1 legend from Paper III.

3.2 Results and Discussion

For details regarding the methodology applied in this section, including cloning strategy and protocols for whole-cell bioconversion assays, seen Experimental procedures from Paper III. Briefly, the growing-cells assays were performed in an M9 medium supplemented with glucose for cofactors regeneration (e.g., NADH) and the substrate for the bioconversion (vanillin or VA). IPTG was added to induce the different ATA encoding genes and streptomycin to prevent plasmid loss by the cells. The resting-cell bioconversions was adapted from (GRAF; ALTENBUCHNER, 2014). The cells were pregrown on LB containing streptomycin and grown until an OD~1, 5 mM. IPTG was added to induce the ATA encoding genes' expression. The culture was grown further to produce biomass and transferred to a

sodium phosphate buffer (50 mM, pH 7.2) supplemented with the substrate for the bioconversion (vanillin or ferulic acid) L-alanine as the amine donor. The concentration of substrates utilized for the bioconversion, the OD of cells utilized as biocatalysts, and the sampling points are provided in the respective figure legend. Additionally, a summary of all strains constructed in this study is represented in Figure 15, including the plasmid and strains employed.

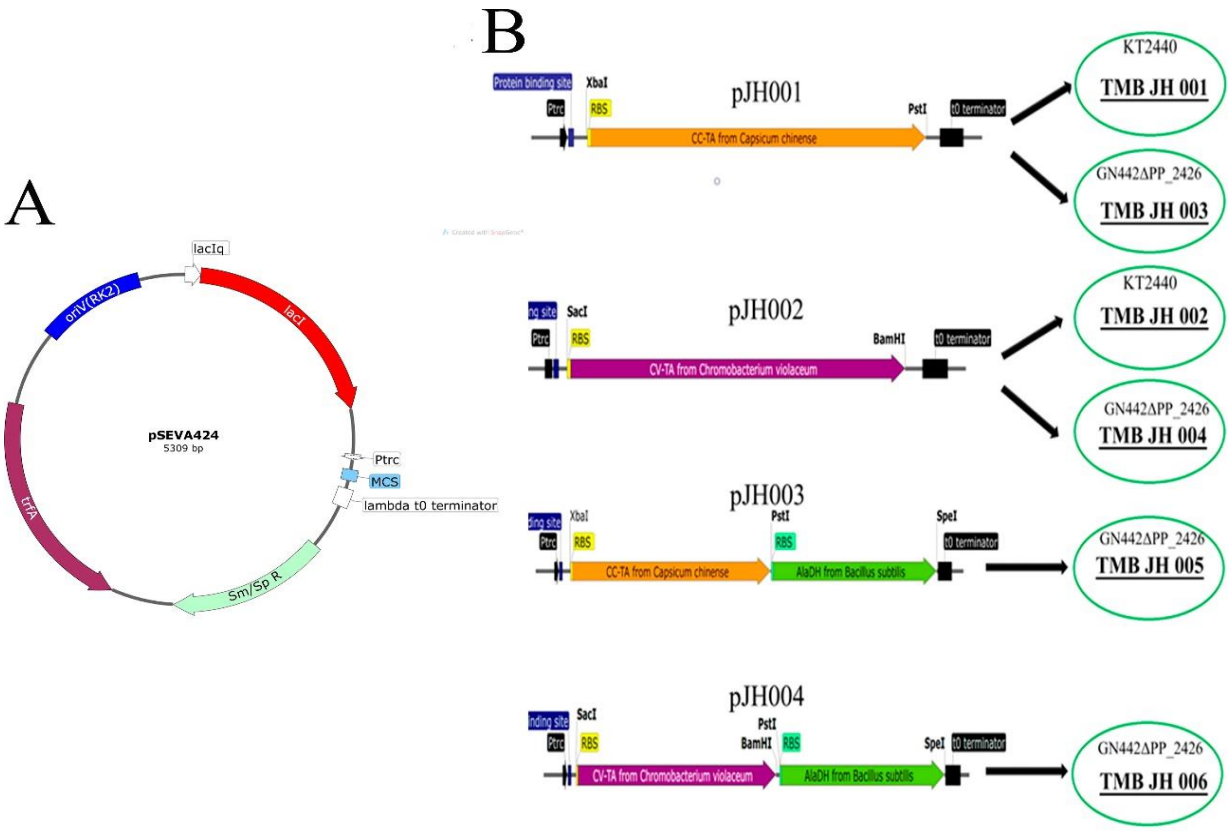


Figure 15. Summary of all constructed plasmids and strains employed in this study. **A**: The empty pSEVA424 plasmid utilized as the backbone to construct the plasmid pJH001 (*CC-ATA* gene), pJH002 (*CV-ATA* gene), pJH003 (*CC-ATA* + *AlaDH* genes), and pJH004 (*CV-ATA* + *AlaDH* genes). **B**: Plasmids (left) and Strains (right) constructed in this work. The names of the plasmids are shown above the arrow representing the genes. Orange arrow: *CC-ATA* gene; Purple arrow: *CV-ATA* gene; Green arrow: *AlaDH* gene. The names of the constructed strains are shown in bold and subscripted.

3.2.1 Determining the background activity against VA in *P. putida* KT2440 and identification of three endogenous ATAs

The potential for using ATAs in the pharmaceutical industry to produce platform amines justifies the search for novel enzymes with different kinetic properties and with activity on a

broad spectrum of substrates. Approaches have been developed to facilitate the screening for enzymes with desired transaminase activity from metagenomic libraries (PAWAR; HALLAM; YADAV, 2018) (COSCOLÍN et al., 2019), but some of these methods are still time-consuming or require high-cost reagents and/or equipment to be applied. For instance, the evaluation of the substrate spectrum can be performed by analyzing the presence of each product by HPLC after a transamination reaction using the purified enzyme or crude cell extract for the biocatalysis (BAUD et al., 2017) or colorimetric approaches using 2-(4-nitrophenyl)ethan-1-amine (R$ 144,8/g; consulted 19.02.2021 on Merck website) or o-xylylenediamine (R$ 250,2/g; consulted: 19.02.2021 on Merck website). These compounds are applied as the amine donor, and a color change occurs if the transamination reaction is achieved (details in (BAUD et al., 2015). Taking this into account, a method was proposed based on growth in agar plates to investigate the ability of transamination against VA using *P. putida* as a model organism. It was hypothesized that to use VA as a carbon source, the cells first convert this compound to vanillin via transamination, metabolized by *P. putida* (Figure 14). Thus, it is possible to evaluate transamination activity against VA only by observing colonies on the agar plate after a transformation with a plasmid bearing an ATA encoding gene. However, background transaminase activity is undesirable for this methodology and leads to inaccurate and unreliable conclusions. Thus, the background activity against VA in this bacterium was investigated since natural isolates of *P. putida* were reported to metabolize this substrate (FLAGAN; LEADBETTER, 2006). In addition, *Pseudomonas* spp. can utilize several amines as a source of carbon/nitrogen, as revised in (LUENGO; OLIVERA, 2020).

For this, the presence of genes in *P. putida* KT2440 genome encoding putative ATAs recognizing VA as substrate and the ability of *P. putida* KT2440 to metabolize VA were evaluated. Firstly, the activity against VA of endogenous ATAs were evaluated in *P. putida* KT2440 through a BLASTp analysis (DALAL; ATRI, 2014) utilizing the protein sequence from different ATAs as query sequences: CV-ATA (GenBank: WP011135573.1, Swiss-Prot: Q7NWG4)(DU et al., 2014), and the CC-ATA (GenBank: AAC78480.1, Swiss-Prot: O82521)(WEBER et al., 2014), previously found to catalyze the reversible transamination reaction of vanillin into VA. Additionally, it was included the protein sequence from Pp-SpuC an ATA from other *P. putida* strain (taxid: 303), which is a putrescine transaminase recognizing a wide broad spectrum of substrates (GALMAN et al., 2017)(GALMAN et al., 2018). The three best hits (lowest e-value, and highest sequence homology) for each enzyme were selected as possible ATAs with putative transaminase activity against VA (Table 5).

Table 5. BLASTp analysis against *P. putida* KT2440 protein database using three ATA as query sequences.

ATA used for the BLASTp	GenBank Accession number	Locus tag	Query cover (%)	Identity (%)	e-value
Pp-SpuC	AAN70747.1	PP_5182 (*spuC-II*)	99	98.01	0.0
Pp-SpuC	AAN67793.1	PP_2180 (*spuC-I*)	99	84.99	0.0
Pp-SpuC	AAN68196.1	PP_2588	98	51.89	2e-170
CV-ATA	AAN70747.1	PP_5182 (*spuC-II*)	96	57.82	0.0
CV-ATA	AAN67793.1	PP_2180 (*spuC-I*)	97	56.05	0.0
CV-ATA	AAN68196.1	PP_2588	94	54.23	1e-169
CC-ATA	AAN68407.1	PP_2799	92	46.82	5e-144
CC-ATA	AAN69998.1	PP_4421	93	42.63	7e-128
CC-ATA	AAN70747.1	PP_5182 (*spuC-II*)	96	42.25	6e-119

The three ATAs with the highest similarity to Pp-SpuC and CV-ATA were selected as candidates for further investigations and denoted as Pp-SpuC-II (GenBank: AAN70747.1), Pp-SpuC-I (GenBank: AAN67793.1), and Pp-ATA (GenBank: AAN68196.1). The *in vivo* activity against VA of these three ATA was determined by overexpressing the genes encoding the respective enzymes in *P. putida* KT2440 (**Figure S3**, Paper III) and GN442ΔPP_2426 (**Figure S5**, Paper III) and evaluating the ability of the cells to metabolize VA, as well as to achieve the transamination of vanillin into VA (**Figure S6**, Paper III). The overexpression of all endogenous ATAs from *P. putida* KT2440 allowed VA production from vanillin in different levels with the highest titer reached by the cell overexpressing Pp-SpuC-II (**Table 1**, Paper III). Details about these three ATAs from *P. putida* KT2440 are discussed in **Results and discussion**

in the section "*Development of a growth-based method to measure transaminase activity towards vanillylamine in P. putida KT2440 and identification of ATAs"* from Paper III, including possible substrates recognized by these enzymes, specific transaminase activity against VA and titer/yield obtained from the transamination of vanillin. Altogether, these results indicate the ATAs recognizing VA are present in *P. putida* KT2440, and Pp-SpuC-II is the major ATA involved in VA metabolism in this bacterium. Thus, it is necessary to determine if these enzymes are expressed in suitable levels to allow this compound as the substrate to sustain cell growth. To investigate the VA metabolism in *P. putida* KT2440, this strain was grown in a liquid M9 medium supplemented with glucose, vanillin, or VA as the sole carbon source (Figure 16).

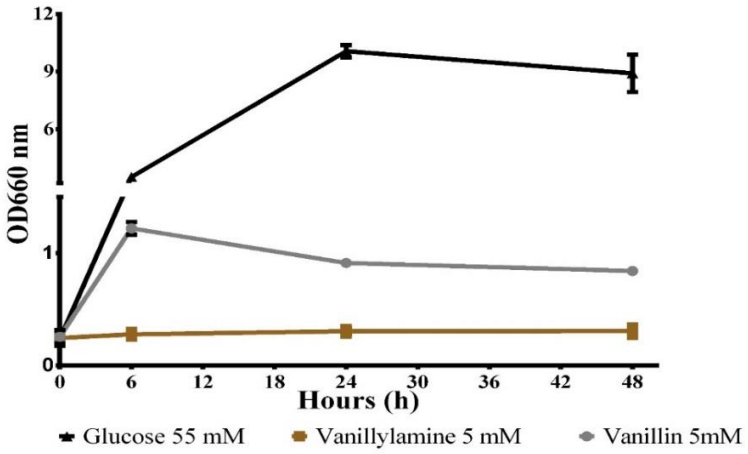


Figure 16. Growth profile of *P. putida* KT2440 wild-type strain after 48 h of cultivation on minimal liquid medium supplemented with glucose (10 g/L ≅ 55 mM), vanillin (5 mM), and vanillylamine (5 mM). The error bars represent the standard deviation (SD) from three biological replicates.

No growth was observed during the 48 h of cultivation with VA as the only carbon source, while significant growth was observed with both glucose (final OD 8.91±0.97) and vanillin (final OD 0.84±0.02). Therefore, the *P. putida* KT2440 strain lacks the endogenous ATAs lack activity towards VA in the applied conditions. Since it was demonstrated that the endogenous ATAs can recognize VA as substrate (**Table 1**, Paper III), the lack of growth is related to an insufficient level of enzymes to allow the utilization of VA. Therefore, the background activity of *P. putida* KT2440 against VA is not enough to enable the cells to metabolize VA, providing a suitable setting for evaluation of recombinant transaminases through a growth-based approach.

3.2.2 Evaluation of the transaminase activity against VA through a growth-based method

Having determined the lack of background transaminase activity against VA in *P. putida* KT2440, CC-ATA and CV-ATA activities were investigated by a growth-based method. In this approach, the gene encoding an ATA was overexpressed using the plasmid pSEVA424, which bears a selective marker for streptomycin resistance and IPTG-inducible promoter. Growth-based selection in the solid medium could be more time and resource-efficient than growth in the liquid medium since it would allow for simple integration into the transformation pipeline and reduce the need for colony-picking and microtiter plate cultivation. To investigate the possibility of selecting for ATAs on solid medium as an integrated part of the transformation protocol, the genes encoding the enzymes CC-ATA and CV-ATA were cloned into the pSEVA424 plasmid and the constructed plasmids were transformed into *P. putida* KT2440 using an electrotransformation protocol. The transformants were poured directly onto plates containing M9 medium (SAMBROOK; RUSSELL, 2001) with 1% (v/v) trace elements solution 100x (PFENNIG; LIPPERT, 1966) supplemented with 5 mM of VA as the sole carbon source and IPTG to inducible the expression of the ATA encoding gene. The plates were incubated at 30 °C for 3 days (Figure 17).

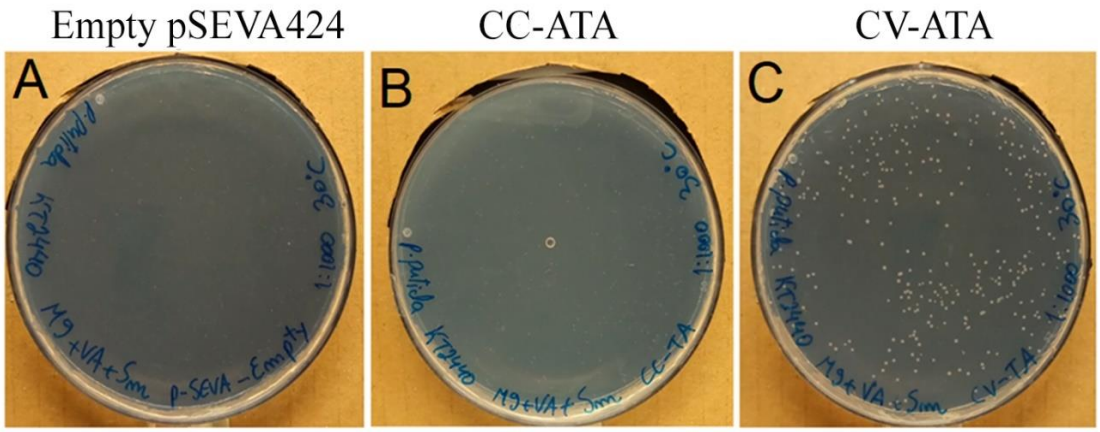


Figure 17. M9 agar plates after 3 days of incubation at 30° C with 5 mM of VA as the sole carbon source, streptomycin (100 µg/ml) to select positive clones and IPTG (1 mM) for induction of ATA encoding genes expression. **A**: cells with the empty pSEVA424 plasmid; **B**: cells expressing *CC-ATA* gene; **C**: cell expressing *CV-ATA* gene.

Clear colonies are visible in the plates containing the cells expressing *CV-ATA* gene (Figure 17C). The colonies on the plates with the transformants bearing the empty plasmid and the *CC-ATA* gene appear not to develop, and the cells expressing the *CV-ATA* gene, despite

then being noticed in the plate (Figure 17A and Figure 17B, respectively). This indicates that CV-ATA production allows cells to promptly metabolize VA for growth, developing better in the same incubation period than the cells without the enzyme. Furthermore, it was investigated the possibility of utilizing only VA to select transformants without the addition of antibiotics (Figure 18). In this case, the selection of positive colonies occurred due only to the VA's utilization for biomass production. To ensure that any apparent growth resulted from the cells using the substrate and not by the absence of the antibiotic, the wild-type cells (WT) were transformed without the empty pSEVA424 plasmid.

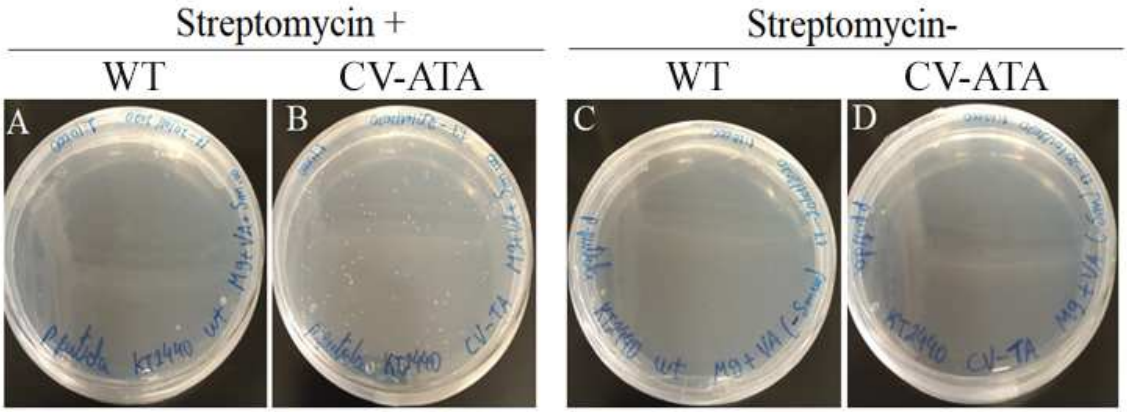


Figure 18. M9 agar plates incubated for 3 days at 30° C with 5 mM of VA as the sole carbon source and IPTG (1 mM) to induce *CV-ATA* gene expression. **A** and **C**: The *P. putida* KT2440 wild-type (WT) strain and **B** and **D**: The cells overexpression *CV-ATA* gene. **A** and **B**: Plates containing streptomycin (100 µg/ml) for selection and **C** and **D**: Plates without antibiotic with only VA for the selection.

The utilization of only VA for the selection prevented the cells expressing *CV-ATA* gene to grow on the plate (Figure 18D), indicating that the simultaneous presence of the substrate and antibiotic is mandatory for the efficiency of the screening method described here. The selection only by the substrate does not allow the plasmids to be maintained by the cells and preventing the utilization of VA to growth, indicating both substrate and antibiotic should be added to the agar plate for proper selection of positive clones carrying transaminase activity against VA. The lack of positive clones may also be related to an insufficient plasmid copy number, suggesting that a certain expression level of CV-ATA is needed to ensure high enough ATA activity to sustain cell growth. The correlation between antibiotic concentration and the number of plasmid copies was previously demonstrated (Begbie *et al.*, 2005; Lian *et al.*, 2016). Lastly, these results corroborate the incapability of *P. putida* KT2440 to utilize VA as the sole carbon source, as no colonies were visible on the plates with WT strain even without the

antibiotic (Figure 18A and Figure 18C). Thus, the formation of colonies on M9 agar plates with VA as the sole carbon source (Figure 17 and Figure 18B) occurred due to the expression of the *CV-ATA* gene, which allows *P. putida* to metabolize this substrate through a transamination reaction producing vanillin that is utilized by the cells to produce biomass.

To confirm that the growth is related to VA consumption, the cells were cultivated in a liquid M9 medium containing VA as the only carbon source. In this case, *P. putida* KT2440 was first transformed with the pSEVA424 plasmids carrying the *CC-ATA* and *CV-ATA* genes. Still, the transformants were plated on M9 supplemented with glucose instead of VA, allowing them to recover the clones with the *CC-ATA* gene once these cells are incapable of being selected on plates containing only VA as the carbon source (Figure 17. After transformation, positive clones were confirmed by colony PCR, and the cells carrying the *CC-ATA* gene were named as TMB-JH001 strain while *CV-ATA* as TMB-JH002. Both strains were cultivated for 24 hours in liquid M9 medium supplemented with 5 mM of VA as the sole carbon source, IPTG for the expression of ATA encoding genes, and streptomycin for plasmid maintenance (Figure 19).

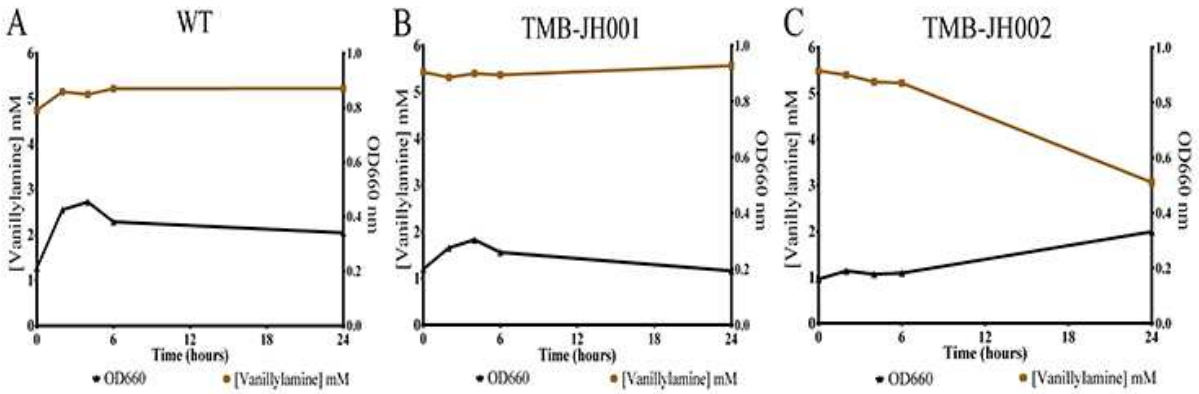


Figure 19. Confirmation of VA consumption by the *P. putida* KT2440 cells overexpressing *CC-ATA* (TMB-JH001) and *CV-ATA* (TMB-JH002) genes. The cells were cultured for 24 hours at 30 °C and 180 rpm in M9 medium with VA as the sole carbon source. **A**: *P. putida* KT2440 wild-type (WT); **B**: TMB-JH001 *P. putida* KT2440 strain overexpressing *CC-ATA* gene; **C**: TMB-JH002 *P. putida* KT2440 strain expressing *CV-ATA* gene. This experiment was performed in only one biological replicate. The initial OD was adjusted to 0.3-0.5. **Sampling points**: 0, 2, 4, 6, and 24 hours.

The VA concentration remained the same for the WT and TMB-JH001 strains during the 24 h of cultivation, even with an increase in OD during the first two hours of cultivation (Figure 19A and Figure 19B). The increase in OD due to VA utilization occurred only for TMB-

JH002 since a reduction in VA concentration was observed (Figure 19C). This result corroborates VA's utilization by the cells overexpressing *CV-ATA* gene and demonstrated the possibility to evaluate the transamination activity against VA using a growth-based system. The expression of the *CV-ATA* gene but not *CC-ATA* led to cell growth in M9 medium supplemented with VA as the sole carbon source, indicating that the screening method proposed distinguishes the activity from different ATAs. Additionally, the screening method can be extrapolated to other amines if the product is metabolizable by *P. putida* KT2440, although this was not evaluated in this work. Lastly, the method can be integrated with the transformation pipeline, facilitating screening for novel enzymes with desired ATA activity.

The overexpression of Pp-SpuC-II encoding gene in *P. putida* KT2440 also allowed the utilization of VA as the sole carbon source by the cells grown in M9 liquid medium (**Figure 2A**, Paper III), although this not occurred by the cells overexpressing Pp-SpuC-I and Pp-ATA encoding genes (**Figure S2**, Paper III). This result corroborates that the Pp-SpuC-II plays a key role in VA metabolism in *P. putida* KT2440. However, the utilization of this substrate depends on an adequate level of enzymes to sustain growth. Nevertheless, the overexpression in *P. putida* KT2440 of all endogenous ATAs encoding genes isolated here allowed the conversion of VA when the whole-cell bioconversion was performed in an M9 liquid medium containing glucose for cofactor regeneration (**Figure S3**, Paper III). Thus, the three endogenous ATA identified from *P. putida* KT2440 and the CV-ATA from *C. violaceum* were able to accomplish the transamination of VA to vanillin. Next, the possibility to produce VA from vanillin was evaluated, which is the unfavored reverse direction in the equilibrium between these two compounds ($\Delta G0' = +2.55$ kcal/mol).

3.2.3 Production of VA from vanillin through whole-cell transamination using a metabolically engineered *P. putida* strain

P. putida KT2440 utilizes vanillin as a carbon source. It thus carries several enzymatic activities required for its assimilation, including vanillin dehydrogenase and other oxidoreductases that catalyze the conversion of vanillin to the byproducts vanillic acid and vanillic alcohol (Figure 14). These reactions compete for the substrate with the activity of ATA. They can result in low transamination of vanillin. Therefore, the strain KT2440 is not a suitable platform for obtaining VA through whole-cell transamination. Instead, the ATAs encoding genes were overexpressed in the GN442ΔPP_2426 strain, which was previously engineered to

accumulate vanillin from ferulic acid and has lower oxidoreductase activity against vanillin than KT2440 (GARCÍA-HIDALGO et al., 2020). The transformation of GN442ΔPP_2426 with plasmids containing the *CC-ATA* and *CV-ATA* genes generated the strains TMB-JH003 and TMB-JH004, respectively (Figure 15). Also, GN442ΔPP_2426 strain was transformed with the empty pSEVA424 plasmid as the negative control. The presence of the genes was confirmed by colony PCR. The *in vivo* transaminase activity was first measured utilizing growing cells for the whole-cell bioconversion of VA to vanillin (Figure 20), which is the favored direction in this transamination reaction.

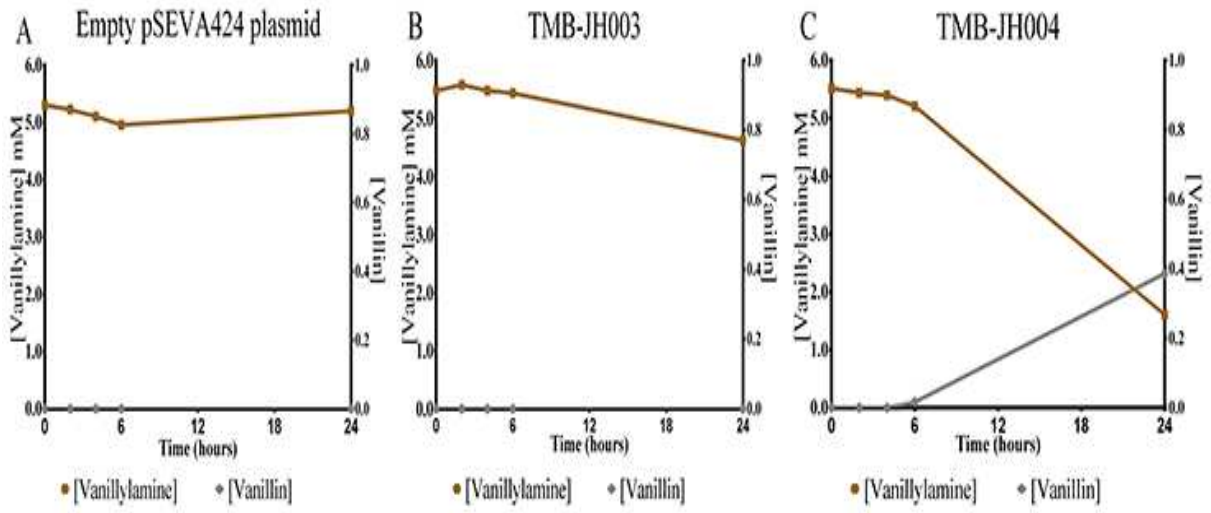


Figure 20. Confirmation of the transaminase activity against VA in the strains TMB-JH003 and TMB-JH004. The cells were cultured for 24 hours at 30 °C and 180 rpm in M9 medium with VA (5 mM) as the substrate for the bioconversion. Glucose (10 g/L) was added for cofactor regeneration and IPTG (1 mM) to induce ATA gene expression. **A**: *P. putida* GN442ΔPP_2426 carrying the empty plasmid; **B**: TMB-JH003: *P. putida* GN442ΔPP_2426 strain expressing *CC-ATA* gene; C: TMB-JH004 *P. putida* GN442ΔPP_2426 strain expressing *CV-ATA* gene. This experiment was performed in only one biological replicate. The initial OD was adjusted to 0.3-0.5. **Sampling points**: 0, 2, 4, 6, and 24 hours.

As in the plate assays, the enzyme CC-ATA could not carry out the transamination of VA (Figure 20B). The CV-ATA was also functional when expressed in GN442ΔPP_2426 since VA was successfully converted into vanillin (Figure 20C). The specific transaminase activity against VA was also analyzed by cell overexpressing Pp-SpuC-II, Pp-SpuC-I and Pp-ATA (**Table 1**, Paper III). For those assays, cells of GN442ΔPP_2426 overexpressing the three endogenous ATAs were utilized for the biocatalysis of VA in a growing-cells setup (glucose added to the M9 medium) (**Figure S5**, Paper III). The highest activities were reached by the cell overexpressing Pp-SpuC-II (0.281 ± 0.10 mmol h^{-1} OD^{-1}) and the strain TMB-JH004

(0.276 ± 0.09 mmol h^{-1} OD^{-1}). These results are in line with those obtained by the reference strain KT2440, indicating that the expression of ATA encoding genes in GN442ΔPP_2426 is also feasible to accomplish the transamination of VA to vanillin. Thus, it was demonstrated the capacity to produce vanillin from VA. The opposite reaction was investigated, which is the unfavored direction in the equilibrium between the two substrates, to produce VA from vanillin using whole-cell transamination (Figure 21). In this case, the OD from the cell culture was adjusted to 3 instead of 0.3-0.5 as in the experiments described above. The cells were grown previously in an M9 medium supplemented with glucose until an OD of 3 and then transferred to the M9 medium for the bioconversion containing vanillin as substrate and glucose for cofactor regeneration (Details in **Experimental procedures**; *Whole-cell bioconversions*, Paper III).

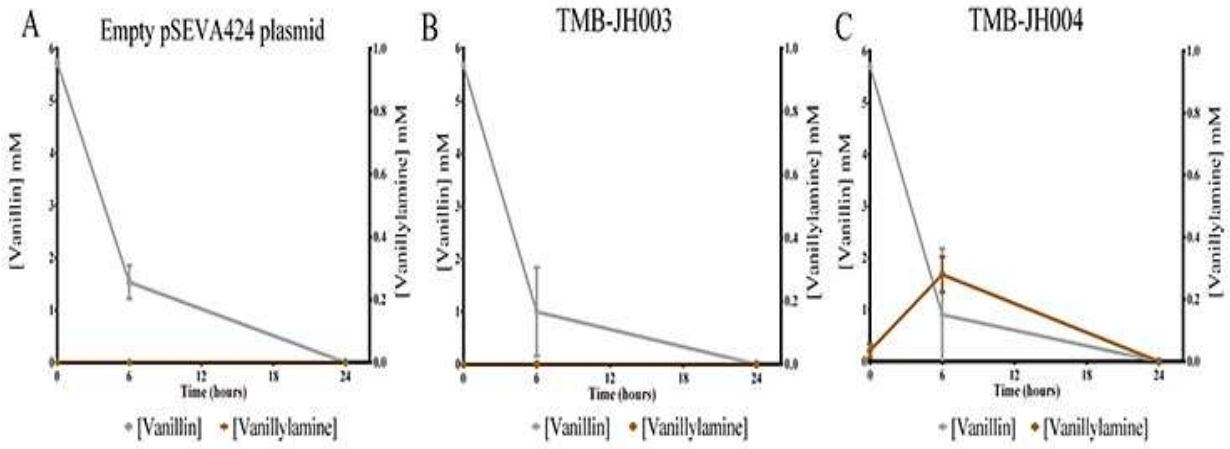


Figure 21. Whole-cell bioconversion of vanillin to VA using growing-cells of GN442ΔPP_2426 strain as biocatalysts. The assays were performed 24 hours at 30 °C and 180 rpm in M9 medium plus glucose (10 g/L). Vanillin (6 mM) was utilized as the substrate, alanine as the amine donor (50 mM) and IPTG (1 mM) was added to induce the CV-ATA gene expression. **A**: GN442ΔPP_2426 cells carrying the empty pSEVA424 plasmid; **B**: TMB-JH003 strain expressing *CC-ATA* gene; **C**: TMB-JH006 strain cell expressing both *CV-ATA* gene. The initial OD was adjusted to 3. Error bars indicate ± SD of two biological replicates. **Sampling points**: 0, 6, and 24 hours.

The TMB-JH004 strain produced 0.28 ± 0.06 mM of VA in 6 hours of bioconversion (Figure 21C), while no conversion occurred in the cells containing the empty plasmid and the TMB-JH003 strain expressing *CC-ATA* gene. The cells overexpressing CC-ATA also failed to achieve vanillin's reversible transamination to VA (Figure 21B). In a previous work (WEBER et al., 2014), the gene encoding *CC-ATA* was cloned in *E. coli*. The enzyme was characterized as vanillin aminotransferase activity (VAMT) since it was performed the *in vitro* transamination of VA into vanillin by CC-ATA. The expression of the *CC-ATA* gene in *P. putida* showed a

different result once both vanillin production from VA (Figure 20B) and the opposite reaction (Figure 21) were not achieved. Besides, no growth was observed for the cells expressing CC-ATA in the agar plates (Figure 17B) or liquid medium (Figure 19B). The possible explanations for the lack of CC-ATA against VA are the (1) genetic background of *P. putida* appears not to be suitable for the expression of the *CC-ATA* gene and production of the active enzyme or (2) an *in vivo* approach is not reliable to measure the ATA activity of CC-ATA against VA. The *in vitro* activity of CC-ATA against VA occurred only when the cells of *E. coli* BL21 were grown at 15 °C during the protein production step (WEBER et al., 2014), suggesting the optimization during the synthesis of CC-ATA in prokaryotic systems is required. The same group previously reported the kinetic resolution of (*rac*)-1-phenylethylamine to (R)-1-phenylethylamine using a recombinant *Saccharomyces cerevisiae* strain for whole-cell transamination through the expression of the same *CC-ATA* gene (WEBER; GORWA-GRAUSLUND; CARLQUIST, 2014), demonstrating the *in vivo* ATA activity of CC-ATA. Therefore, the expression of the *CC-ATA* gene and/or the protein misfolding in *P. putida* appears to be the consequence of the lack of ATA activity of CC-ATA. Nevertheless, since the VA production from CV-ATA was demonstrated, the CC-ATA was discarded from further investigations.

The VA production by growing-cells of TMB-JH004 was limited by product re-assimilation (Figure 21C) and by-product formation (not measured for Figure 21 but confirmed in **Figure 4B** and **4D** from Paper III), which may be related to the remainong activity of oxidoreductase in GN442ΔPP_2426. The production of cofactors (e.g., NADH) through the tricarboxylic acid cycle (TCA) due to glucose in the culture broth and, consequently, their utilization by other oxidoreductases explains the remaining background activity against vanillin in this strain. Potential targets for further metabolic engineering are the genes PP_3151, PP_5120, and PP_5258 encoding putative aldehyde dehydrogenase acting on vanillin as demonstrated by a proteomic analysis performed using *P. putida* KT2440 in response to this substrate (SIMON et al., 2014). Regarding the VA re-assimilation, the reaction reversibility may be impacting the transamination of vanillin and affecting the VA production. Indeed, the formation of vanillin from VA is the preferred direction of this reaction (ΔG0’ = +2.55 kcal/mol for vanillin to VA,WEBER et al., 2014). Thus, both hypotheses were evaluated.

To investigate the impact of glucose in vanillin's transamination, further bioconversions were performed using resting cells instead of growing cells. In this case, it was adapted the protocol applied by (GRAF; ALTENBUCHNER, 2014). First, the cells were pregrown in LB

medium, and IPTG (5 mM) was utilized to induce the ATA encoding gene expression in a high cell density culture. Posteriorly, the cells were washed and transferred to a sodium phosphate buffer containing the substrates for the transamination (vanillin and L-alanine). In this way, the cell growth and the whole-cell bioconversions are performed separately, and the biocatalysis is performed in a non-growth condition. This protocol was applied to investigate the capacity of CV-ATA, CC-ATA, and all three endogenous ATAs from KT2440 to produce VA from vanillin (**Figure S6**, Paper III).

VA was successfully measured in the culture broth of the cells overexpressing all endogenous ATAs from KT2440 in different levels: Pp-SpuC-II (0.700 ± 0.20 mM), Pp-SpuC-I (0.180 ± 0.01mM) and Pp-ATA (0.360 ± 0.01 mM) (**Table 1**, Paper III). The highest titer of 0.915 ± 0.30 mM was obtained by TMB-JH004 (**Figure S6D**, Paper III). For this strain, no vanillic acid and vanillyl alcohol were detected in the broth (**Figure S6D**, Paper III), corroborating that the presence of glucose leads to an increase in the background activity of oxidoreductases acting on vanillin. It is important to mention that a higher amount of vanillin (10 mM) was utilized as substrate for the resting-cell bioconversion than for the growing-cells (5 mM), explaining the difference between the VA titers obtained between these two bioconversions. Nevertheless, the lack of production of vanillic acid and vanillyl alcohol by TMB-JH004 demonstrated that the utilization of resting cells is recommended to avoid byproduct formation during the transamination of vanillin. The impact of byproducts formation during the production of VA from vanillin by ATAs is well discussed in the Section **Results and discussion**; *Overexpression of transaminases in GN442ΔPP_2426 enables whole-cell bioconversion of vanillin to VA* from Paper III.

Although the whole-cell transamination by resting cells of TMB-JH004 led to the reduction of byproducts production, the re-assimilation of VA was evident after 4 h of bioconversion (**Figure S6D**, Paper III). Furthermore, vanillin was not completely converted after 24 hours as occurred for the bioconversion applying growing cells of TMB-JH004 (Figure 21C). Only 1.72 mM of vanillin was converted into VA, indicating the process is still limited. This could be related to unfavorable reaction equilibrium towards the formation of VA due to inefficient removal of pyruvate produced during the transamination of vanillin or by low availability of intracellular L-alanine. In this perspective, the effect of the simultaneous production of CV-ATA and AlaDH from *B. subtilis* could improve the titers of VA. Pyruvate is the co-product of vanillin transamination when L-alanine is used as the amine donor (Figure

14). Thus, the utilization of AlaDH is an efficient strategy to drive the reaction to VA production since this enzyme utilizes pyruvate to produce L-alanine, removing the co-product to and recycle the amine donor. The *AlaDH* gene was cloned into the plasmid pJH002 resulting in the plasmid pJH004 utilized to transform GN442ΔPP_2426, resulting in the strain TMB-JH006 over-expressing both *CV-ATA* and *AlaDH* genes (confirmed by colony PCR) under the control of the same IPTG-inducible promoter. The whole-cell transamination of vanillin was compared utilizing growing-cells of TMB-JH006 and TMB-JH004 and a high cell density (OD= 10). For this bioconversion (**Figure 4**, Paper III), the M9 medium for the bioconversion was supplemented with the vanillin (5 mM) as the substrate, L-alanine (50 mM) as the amine donor, glucose (10 g/L) for NADH regeneration, and NH_4Cl (200 mM) to provide NH_3 for the AlaDH reaction.

The highest VA titer reached by TMB-JH004 was 0.70 ± 0.01 mM (**Figure 4D**, Paper III), while TMB-JH006 produced 0.84 ± 0.07 mM of VA (**Figure 4C**, Paper III). These values were obtained in the first 6 h of bioconversion, preceding the VA re-assimilation. This result indicates that concomitantly activity of CV-ATA and AlaDH in the conditions applied did not impact the re-assimilation of VA. The AlaDH reaction was further investigated by performing the bioconversion without L-alanine but in NH_4Cl (200 mM). In this case, the principle relies on the endogenous amine donors produced by the cells through the utilization of the ammonium added (**Figure 4A** and **4B**, Paper III). In the Section **Results and discussion**, *Effect of L-alanine regeneration by AlaDH on whole-cell bioconversion of vanillin to VA* from Paper III was discussed the impact of AlaDH in the transamination of vanillin.

The omission of both L-alanine and NH_4CL caused a lack of transamination activity against VA by TMB-JH004 (**Figure S7D**, Paper III). Still, the same did not occur with the addition of NH_4CL (**Figure 4B**, Paper III). Additionally, the re-assimilation of VA was abolished for both strains TMB-JH006 (**Figure 4A**, Paper III) and TMB-JH004 (**Figure 4B**, Paper III). This suggested that relying on the amine donor production by the cells providing only NH_4Cl is a feasible alternative for VA accumulation, especially by expressing both CV-ATA and AlaDH. Therefore, the intracellular levels of amines donors and NH_3 impact the production of VA through whole-cell transamination of vanillin. Optimization regarding the amine donor utilized and its concentration is needed for an efficient transamination of vanillin into VA. In addition, other parameters, such as the external supply of NADH, ammonium and/or PLP should be optimized for improving VA production. Nevertheless, as proof of concept, the

strains TMBJH004 and TMBJH006 demonstrated the potential to produce VA from vanillin through whole-cell transamination by either growing-cell or resting-cells. Lastly, it was evaluated the possibility of producing from ferulic acid using GN442ΔPP_2426.

3.2.4 Production of VA from Ferulic acid through whole-cell biocatalysis

Ferulic acid is a lignin-derivative from the coniferyl funneling pathway metabolized into vanillin by the enzymes Feruloyl-CoA-synthetase and Enoyl-CoA-hydratase/aldolase encoded in *P. putida* KT2440 by the genes *fcs* and *ech*, respectively (Figure 14). The strain GN442ΔPP_2426 was metabolically engineered to accumulate vanillin from ferulic acid (GRAF; ALTENBUCHNER, 2014)(GARCÍA-HIDALGO et al., 2020)**,** and it is a suitable host platform to produce VA from ferulic acid. As shown in Figure 5 from Paper III, both growing-cells of TMB-JH004 (0.235 ± 0.18 mM) and TMB-JH006 (0.473 ± 0.09 mM) were able to produce VA from ferulic acid. The production of the byproducts vanillic acid and vanillyl alcohol occurred during the conversion of ferulic acid to VA, corroborating that further metabolic engineering is necessary to develop a host platform to produce VA from both ferulic acid and vanillin. Nevertheless, the production VA was demonstrated from ferulic acid through microbial bioconversion although the process is still limited. To the best of our knowledge, this is the first time that *P. putida* as utilized as host platform to produce VA.

3.3 Conclusion

In this work, as a proof-of-concept, the production of a building-block amine by microbial biocatalysis from compounds obtained from depolymerized technical lignin was demonstrated. Employing a metabolically engineered *P. putida* strain designed to accumulate vanillin from ferulic acid (GN442ΔPP_2426 strain), it was possible to produce VA from two lignin-derived compounds ferulic acid and vanillin. These results demonstrated a feasible process for lignin valorization to produce aromatic amines by whole-cell transamination. Additionally, three endogenous ATA from *P. putida* KT2440 with activity against VA were identified through BLASTp analysis utilizing the protein sequence from CV-ATA, Pp-SpuC, and CC-ATA. The expression of all three endogenous ATA in GN442ΔPP_2426 led to VA production from vanillin in different levels depending on the ATA utilized.

The reduced background activity of endogenous ATAs against VA in *P. putida* KT2440 allowed the implementation of a growth-based approach to evaluate the transaminase activity

from different ATAs against this substrate. The method was incorporated into the transformation pipeline and applied to evaluate CV-ATA and CC-ATA activity against VA. It was demonstrated that only the cells expression *CV-ATA* gene could grow in agar plates containing VA as the sole carbon source. Furthermore, the heterologous expressions of CV-ATA and the endogenous Pp-SpuC-II encoding genes in *P. putida* KT2440 generated strains able to grow in VA as the sole carbon source, which is not metabolized promptly by this strain. These two enzymes showed the highest *in vivo* activity against VA. However, the expression of the other two endogenous ATA encoding genes identified in *P. putida* KT2440 also allowed the cells to convert VA. The Pp-SpuC-II appears to be the major ATA involved in VA metabolism in *P. putida* KT2440, although high levels are required to utilize this substrate.

The production of VA from vanillin by whole-cell transamination was limited by the formation of vanillic acid and vanillyl alcohol competing with the ATA activity for the substrate and by re-assimilation of the product, which may be related to an unfavorable equilibrium towards VA. The utilization of resting cells instead of growing cells led to reduced production of the byproducts, corroborating that GN442 ΔPP_2426 still retains oxidoreductase background activity against vanillin. The regeneration of L-alanine and pyruvate recycling by the enzyme AlaDH from *B. subtilis* abolished the re-assimilation of VA, but only when the transamination reaction relied on the endogenous amine donors produced by the cells. Optimizations regarding the concentration of amine donor, ammonium, and/or glucose (for NADH regeneration) utilized are necessary to improve the VA production from both vanillin and ferulic acid.

4. REFERENCES

ABDELAZIZ, O. Y. et al. Biological valorization of low molecular weight lignin. **Biotechnology Advances**, v. 34, n. 8, p. 1318–1346, 2016.

AKŞIT, A. et al. **Peroxisomal membrane contact sites in the yeast Hansenula polymorpha**. [s.l: s.n.].

ANASTAS, P.; EGHBALI, N. Green Chemistry: Principles and Practice. **Chem. Soc. Rev.**, v. 39, n. 1, p. 301–312, 2010.

ANDERSON, M. et al. Total synthesis of capsaicin analogues from lignin-derived compounds by combined heterogeneous metal, organocatalytic and enzymatic cascades in one pot. **Advanced Synthesis and Catalysis**, v. 356, n. 9, p. 2113–2118, 2014.

ANDERSON, M. **Amine Transaminases in Multi-Step One-Pot Reactions Stockholm 2017**. [s.l: s.n.].

APARICIO, T. et al. High-Efficiency Multi-site Genomic Editing of Pseudomonas putida through Thermoinducible ssDNA Recombineering. **iScience**, v. 23, n. 3, 2020.

BADENHORST, C. P. S.; BORNSCHEUER, U. T. Getting Momentum: From Biocatalysis to Advanced Synthetic Biology. **Trends in Biochemical Sciences**, v. 43, n. 3, p. 180–198, 2018.

BAUD, D. et al. A rapid, sensitive colorimetric assay for the high-throughput screening of transaminases in liquid or solid-phase. **Chemical Communications**, v. 51, n. 97, p. 17225–17228, 2015.

BAUD, D. et al. A metagenomics approach for new biocatalyst discovery: Application to transaminases and the synthesis of allylic amines. **Green Chemistry**, v. 19, n. 4, p. 1134–1143, 2017.

BEGBIE, S. et al. **The Effects of Sub-Inhibitory Levels of Chloramphenicol on pBR322 Plasmid Copy Number in Escherichia coli DH5α CellsJournal of Experimental Microbiology and Immunology (JEMI)**, 2005. Disponível em: <https://pdfs.semanticscholar.org/2b91/27bdc90f345e5956d8c5ffdb00ad81ec65f8.pdf>

BENJAMIN, K. R. et al. **Developing commercial production of semi-synthetic artemisinin, and of β-Farnesene, an Isoprenoid Produced by Fermentation of Brazilian SugarJournal of the Brazilian Chemical Society**, 2016.

BITTER, T.; MUIR, H. M. A modified uronic acid carbazole reaction. **Analytical biochemistry**, v. 4, p. 330–334, 1962.

BLONDIAUX, E. et al. Bio-based Aromatic Amines from Lignin-Derived Monomers. **ACS Sustainable Chemistry and Engineering**, v. 7, n. 7, p. 6906–6916, 2019.

BOERIU, C. G. et al. Production Methods for Hyaluronan. **International Journal of Carbohydrate Chemistry**, v. 2013, p. 1–14, 2013.

BOGDANOVA, A. I.; AGAFONOV, M. O.; TER-AVANESIAN, M. D. [Plasmid instability in

methylotrophic yeast Hansenula polymorpha: the capture of chromosomal DNA fragments by integrative plasmids]. **Molekuliarnaia biologiia**, v. 34, n. 1, p. 28–35, 2000.

BRINK, D. P. et al. Mapping the diversity of microbial lignin catabolism: experiences from the eLignin database. **Applied Microbiology and Biotechnology**, v. 103, n. 10, p. 3979–4002, 2019.

BURGARD, A. et al. Development of a commercial scale process for production of 1,4-butanediol from sugar. **Current Opinion in Biotechnology**, v. 42, p. 118–125, 2016.

COSCOLÍN, C. et al. Bioprospecting reveals class ω-transaminases converting bulky ketones and environmentally relevant polyamines. **Applied and Environmental Microbiology**, v. 85, n. 2, p. 1–20, 2019.

DALAL, A.; ATRI, A. An Introduction to Sequence and Series. **International Journal of Research**, v. 1, n. 10, p. 1286–1292, 2014.

DE CARVALHO, C. C. C. R. Whole cell biocatalysts: essential workers from Nature to the industry. **Microbial Biotechnology**, v. 10, n. 2, p. 250–263, 2017.

DE OLIVEIRA, J. D. et al. Genetic basis for hyper production of hyaluronic acid in natural and engineered microorganisms. **Microbial Cell Factories**, v. 15, n. 1, p. 1–19, 2016.

DEANGELIS, P. L. Molecular directionality of polysaccharide polymerization by the Pasteurella multocida hyaluronan synthase. **Journal of Biological Chemistry**, v. 274, n. 37, p. 26557–26562, 1999.

DEANGELIS, P. L.; PAPACONSTANTINOU, J.; WEIGEL, P. H. Isolation of a Streptococcus pyogenes gene locus that directs hyaluronan biosynthesis in acapsular mutants and in heterologous bacteria. **Journal of Biological Chemistry**, v. 268, n. 20, p. 14568–14571, 1993.

DEANGELIST, P. L.; ACHYUTHAN, A. M. Yeast-derived recombinant DG42 protein of Xenopus can synthesize hyaluronan in vitro. **Journal of Biological Chemistry**, v. 271, n. 39, p. 23657–23660, 1996.

DESSBESELL, L. et al. Global lignin supply overview and kraft lignin potential as an alternative for petroleum-based polymers. **Renewable and Sustainable Energy Reviews**, v. 123, n. February 2019, p. 109768, 2020.

DOVEDYTIS, M.; LIU, Z. J.; BARTLETT, S. Hyaluronic acid and its biomedical applications: A review. **Engineered Regeneration**, v. 1, n. October, p. 102–113, 2020.

DU, C. J. et al. Evaluation of CV2025 ω-transaminase for the bioconversion of lignin breakdown products into value-added chemicals: Synthesis of vanillylamine from vanillin. **Biocatalysis and Biotransformation**, v. 32, n. 5–6, p. 302–313, 2014.

DUSNY, C.; SCHMID, A. The MOX promoter in Hansenula polymorpha is ultrasensitive to glucose-mediated carbon catabolite repression. **FEMS Yeast Research**, v. 16, n. 6, p. 1–41, 2016.

DUUREN, J. B. J. H. et al. Limited life cycle and cost assessment for the bioconversion of lignin-derived aromatics into adipic acid. **Biotechnology and Bioengineering**, v. 117, n. 5, p. 1381–1393, 20 maio 2020.

FABER, K. N. et al. Highly-efficient electrotransformation of the yeast Hansenula polymorpha. **Current Genetics**, v. 25, n. 4, p. 305–310, 1994.

FACHE, M. et al. Amine hardeners and epoxy cross-linker from aromatic renewable resources. **European Polymer Journal**, v. 73, p. 344–362, 2015.

FATTORI, V. et al. Capsaicin: Current understanding of its mechanisms and therapy of pain and other pre-clinical and clinical uses. **Molecules**, v. 21, n. 7, p. 1–33, 2016.

FLAGAN, S. F.; LEADBETTER, J. R. Utilization of capsaicin and vanillylamine as growth substrates by Capsicum (hot pepper)-associated bacteria. **Environmental Microbiology**, v. 8, n. 3, p. 560–565, 2006.

FROIDEVAUX, V. et al. Biobased Amines: From Synthesis to Polymers; Present and Future. **Chemical Reviews**, v. 116, n. 22, p. 14181–14224, 2016.

GALMAN, J. L. et al. Biocatalytic transamination with near-stoichiometric inexpensive amine donors mediated by bifunctional mono- and di-amine transaminases. **Green Chemistry**, v. 19, n. 2, p. 361–366, 2017.

GALMAN, J. L. et al. Characterization of a Putrescine Transaminase From Pseudomonas putida and its Application to the Synthesis of Benzylamine Derivatives. **Frontiers in Bioengineering and Biotechnology**, v. 6, n. December, p. 1–11, 2018.

GARCÍA-HIDALGO, J. et al. Vanillin Production in Pseudomonas: Whole-Genome Sequencing of Pseudomonas sp. Strain 9.1 and Reannotation of Pseudomonas putida CalA as a Vanillin Reductase. **Applied and environmental microbiology**, v. 86, n. 6, 2020.

GARZÓN-POSSE, F. et al. Whole Cells as Biocatalysts in Organic Transformations. **Molecules**, v. 23, n. 6, 2018.

GELLISSEN, G. et al. New yeast expression platforms based on methylotrophic Hansenula polymorpha and Pichia pastoris and on dimorphic Arxula adeninivorans and Yarrowia lipolytica - A comparison. **FEMS Yeast Research**, v. 5, n. 11, p. 1079–1096, 2005.

GILLET, S. et al. Lignin transformations for high value applications: Towards targeted modifications using green chemistry. **Green Chemistry**, v. 19, n. 18, p. 4200–4233, 2017.

GOMES, A. M. V. et al. Heterologous hyaluronic acid production in kluyveromyces lactis. **Microorganisms**, v. 7, n. 9, 2019.

GOMIDE, M. S. et al. Genetic switches designed for eukaryotic cells and controlled by serine

integrases. **Communications Biology**, v. 3, n. 1, p. 1–14, 2020.

GOMM, A.; O'REILLY, E. Transaminases for chiral amine synthesis. **Current Opinion in Chemical Biology**, v. 43, n. Figure 1, p. 106–112, abr. 2018.

GOTTSCHALK, J. et al. Key factors for a one-pot enzyme cascade synthesis of high molecular weight hyaluronic acid. **International Journal of Molecular Sciences**, v. 20, n. 22, 2019.

GRAF, N.; ALTENBUCHNER, J. Genetic engineering of Pseudomonas putida KT2440 for rapid and high-yield production of vanillin from ferulic acid. **Applied Microbiology and Biotechnology**, v. 98, n. 1, p. 137–149, 2014.

GUAZZARONI, M.; SILVA-ROCHA, R.; WARD, R. J. Synthetic biology approaches to improve biocatalyst identification in metagenomic library screening. **Microbial Biotechnology**, v. 8, n. 1, p. 52–64, 13 jan. 2015.

GUO, F.; BERGLUND, P. Transaminase biocatalysis: Optimization and application. **Green Chemistry**, v. 19, n. 2, p. 333–360, 2017.

HALDAR, D.; PURKAIT, M. K. Lignocellulosic conversion into value-added products: A review. **Process Biochemistry**, v. 89, n. July 2019, p. 110–133, 2020.

HILLMAN, M. et al. , Adrielli Vilanova. p. 513–524, 2017.

HMAR, R. V. et al. Chromosomal integration of hyaluronic acid synthesis (has) genes enhances the molecular weight of hyaluronan produced in Lactococcus lactis. **Biotechnology Journal**, v. 9, n. 12, p. 1554–1564, 2014.

HOLLENBERG, C. P.; GELLISSEN, G. Production of recombinant proteins by methylotrophic yeasts. **Current Opinion in Biotechnology**, v. 8, n. 5, p. 554–560, out. 1997.

HUYNH, A.; PRIEFER, R. Hyaluronic acid applications in ophthalmology, rheumatology, and dermatology. **Carbohydrate Research**, v. 489, n. February, p. 107950, 2020.

IM, J. H. et al. Optimization of medium components for high-molecular-weight hyaluronic acid production by Streptococcus sp. ID9102 via a statistical approach. **Journal of Industrial Microbiology and Biotechnology**, v. 36, p. 1337–1344, 2009.

IVANKOVIĆ, A. Review of 12 Principles of Green Chemistry in Practice. **International Journal of Sustainable and Green Energy**, v. 6, n. 3, p. 39, 2017.

JENZELEWSKI, V. Fermentation and Primary Product Recovery. In: **Hansenula polymorpha**. Weinheim, FRG: Wiley-VCH Verlag GmbH & Co. KGaA, 2005. p. 156–174.

JEONG, E.; SHIM, W. Y.; KIM, J. H. Metabolic engineering of Pichia pastoris for production of hyaluronic acid with high molecular weight. **Journal of Biotechnology**, v. 185, p. 28–36, 2014a.

JEONG, E.; SHIM, W. Y.; KIM, J. H. Metabolic engineering of Pichia pastoris for production of hyaluronic acid with high molecular weight. **Journal of Biotechnology**, v. 185, p. 28–36, 2014b.

JIN, P. et al. Production of specific-molecular-weight hyaluronan by metabolically engineered Bacillus subtilis 168. **Metabolic Engineering**, v. 35, p. 21–30, 2016.

JOHNS, M. R.; GOH, L. T.; OEGGERLI, A. Effect of pH, agitation and aeration on hyaluronic acid production by Streptococcus zooepidemicus. **Biotechnology Letters**, v. 16, n. 5, p. 507–512, 1994.

JUERGENS, H. et al. **Genome editing in Kluyveromyces and Ogataea yeasts using a broad-host-range Cas9/gRNA co-expression plasmid**. [s.l: s.n.].

KAULMANN, U. et al. Substrate spectrum of ω-transaminase from Chromobacterium violaceum DSM30191 and its potential for biocatalysis. **Enzyme and Microbial Technology**, v. 41, n. 5, p. 628–637, 2007.

KAUR, M.; JAYARAMAN, G. Hyaluronan production and molecular weight is enhanced in pathway-engineered strains of lactate dehydrogenase-deficient Lactococcus lactis. **Metabolic Engineering Communications**, v. 3, p. 15–23, 2016.

KELLY, S. A. et al. **Application of ω-Transaminases in the Pharmaceutical IndustryChemical Reviews**, 2018.

KELLY, S. A. et al. Transaminases for industrial biocatalysis: novel enzyme discovery. **Applied Microbiology and Biotechnology**, 2020.

KIM, M. W. et al. Functional characterization of the Hansenula polymorpha HOC1, OCH1, and OCR1 genes as members of the yeast OCH1 mannosyltransferase family involved in protein glycosylation. **Journal of Biological Chemistry**, v. 281, n. 10, p. 6261–6272, 2006.

KO, Y. et al. Chem Soc Rev Tools and strategies of systems metabolic engineering for the development of microbial cell factories for chemical production. **Chemical Society Reviews**, 2020.

KOHLSTEDT, M. et al. From lignin to nylon: Cascaded chemical and biochemical conversion using metabolically engineered Pseudomonas putida. **Metabolic Engineering**, v. 47, p. 279–293, 2018.

KUHN, D. et al. Systems biotechnology - Rational whole-cell biocatalyst and bioprocess design. **Engineering in Life Sciences**, v. 10, n. 5, p. 384–397, 2010.

KURTZMAN, C. P.; ROBNETT, C. J. Systematics of methanol assimilating yeasts and neighboring taxa from multigene sequence analysis and the proposal of Peterozyma gen. nov., a new member of the Saccharomycetales. **FEMS Yeast Research**, v. 10, n. 3, p. 353–361, 2010.

LEE, S. Y.; KIM, H. U. Systems strategies for developing industrial microbial strains. **Nature Biotechnology**, v. 33, n. 10, p. 1061–1072, 2015.

LEVINE, D. W.; COONEY, C. L. Isolation and Characterization of a Thermotolerant Methanol-

Utilizing Yeast. **Applied Microbiology**, v. 26, n. 6, p. 982–990, 1973.

LIAN, J.; JIN, R.; ZHAO, H. Construction of plasmids with tunable copy numbers in Saccharomyces cerevisiae and their applications in pathway optimization and multiplex genome integration. **Biotechnology and Bioengineering**, v. 113, n. 11, p. 2462–2473, 2016.

LIN, B.; TAO, Y. Whole-cell biocatalysts by design. **Microbial Cell Factories**, v. 16, n. 1, p. 106, 13 dez. 2017.

LIU, L. et al. Comparative study on the influence of dissolved oxygen control approaches on the microbial hyaluronic acid production of Streptococcus zooepidemicus. **Bioprocess and Biosystems Engineering**, v. 32, n. 6, p. 755–763, 2009.

LIU, L. et al. **Microbial production of hyaluronic acid: current state, challenges, and perspectivesMicrobial Cell Factories**, 2011.

LUENGO, J. M.; OLIVERA, E. R. Catabolism of biogenic amines in Pseudomonas species. **Environmental Microbiology**, v. 22, n. 4, p. 1174–1192, 2020.

MANFRÃO-NETTO, J. H. C. et al. Evaluation of ogataea (Hansenula) polymorpha for hyaluronic acid production. **Microorganisms**, v. 9, n. 2, p. 1–16, 2021a.

MANFRÃO-NETTO, J. H. C. et al. Metabolic engineering of Pseudomonas putida for production of vanillylamine from lignin-derived substrates. **Microbial Biotechnology**, 2021b.

MANFRÃO-NETTO, J. H. C.; GOMES, A. M. V.; PARACHIN, N. S. Advances in Using Hansenula polymorpha as Chassis for Recombinant Protein Production. **Frontiers in Bioengineering and Biotechnology**, v. 7, n. May, p. 1–13, 2019.

MARCELLIN, E.; STEEN, J. A.; NIELSEN, L. K. Insight into hyaluronic acid molecular weight control. **Applied Microbiology and Biotechnology**, v. 98, n. 16, p. 6947–6956, 2014.

MARINI, E. et al. Antimicrobial and anti-virulence activity of capsaicin against erythromycin-resistant, cell-invasive group A streptococci. **Frontiers in Microbiology**, v. 6, n. NOV, p. 1–7, 2015.

MERRICK, C. A.; ZHAO, J.; ROSSER, S. J. Serine Integrases: Advancing Synthetic Biology. **ACS Synthetic Biology**, v. 7, n. 2, p. 299–310, 16 fev. 2018.

MOGHEISEH, M.; KARIMIAN, R.; KHOSHSEFAT, M. Vanillin-derived epoxy monomer for synthesis of bio-based epoxy thermosets: effect of functionality on thermal, mechanical, chemical and structural properties. **Chemical Papers**, n. 0123456789, 2020.

MURADO, M. A. et al. Optimization of extraction and purification process of hyaluronic acid from fish eyeball. **Food and Bioproducts Processing**, v. 90, n. 3, p. 491–498, 2012.

NUMAMOTO, M.; MAEKAWA, H.; KANEKO, Y. Efficient genome editing by CRISPR/Cas9 with a tRNA-sgRNA fusion in the methylotrophic yeast Ogataea polymorpha. **Journal of Bioscience**

and Bioengineering, v. 124, n. 5, p. 487–492, 2017.

PADDON, C. J. et al. High-level semi-synthetic production of the potent antimalarial artemisinin. **Nature**, v. 496, n. 7446, p. 528–532, 10 abr. 2013.

PADROSA, D. R. et al. Enhancing PLP-Binding Capacity of Class-III ω-Transaminase by Single Residue Substitution. **Frontiers in Bioengineering and Biotechnology**, v. 7, n. October, p. 1–13, 18 out. 2019.

PALMER, W. J.; MEYER, K. Polysaccharide of Vitreous Humor. **Ophthalmology**, p. 629–634, 1934.

PAN, N. C. et al. Effect of fermentation conditions on the production of hyaluronic acid by streptococcus zooepidemicus ATCC 39920. **Acta Scientiarum - Biological Sciences**, v. 37, n. 4, p. 411–417, 2015.

PANCHAL, S. K.; BLISS, E.; BROWN, L. Capsaicin in metabolic syndrome. **Nutrients**, v. 10, n. 5, p. 14–18, 2018.

PATIL, M. D. et al. Recent Advances in ω-Transaminase-Mediated Biocatalysis for the Enantioselective Synthesis of Chiral Amines. **Catalysts**, v. 8, n. 7, p. 254, 21 jun. 2018.

PAWAR, S. V.; HALLAM, S. J.; YADAV, V. G. Metagenomic discovery of a novel transaminase for valorization of monoaromatic compounds. **RSC Advances**, v. 8, n. 40, p. 22490–22497, 2018.

PFENNIG, N.; LIPPERT, K. D. Über das Vitamin B12-Bedürfnis phototropher Schwefelbakterien. **Archiv für Mikrobiologie**, v. 55, n. 3, p. 245–256, 1966.

PINTO, A.; CONTENTE, M. L.; TAMBORINI, L. Advances on whole-cell biocatalysis in flow. **Current Opinion in Green and Sustainable Chemistry**, v. 25, p. 100343, 2020.

RAMEZANI-RAD, M. et al. The Hansenula polymorpha (strain CBS4732) genome sequencing and analysis. **FEMS Yeast Research**, v. 4, n. 2, p. 207–215, 2003.

RAVIN, N. V et al. Genome sequence and analysis of methylotrophic yeast Hansenula polymorpha DL1. **BMC genomics**, v. 14, n. 837, 2013.

RILEY, R. et al. Comparative genomics of biotechnologically important yeasts. **Proceedings of the National Academy of Sciences**, v. 113, n. 35, p. 9882–9887, 2016.

RYABOVA, O. B.; CHMIL, O. M.; SIBIRNY, A. A. Xylose and cellobiose fermentation to ethanol by the thermotolerant methylotrophic yeast Hansenula polymorpha. **FEMS Yeast Research**, v. 4, n. 2 SPEC.ISS., p. 157–164, 2003.

SAMBROOK, J. ;; RUSSELL, D. W. ; **Sambrook & Russel - Molecular Cloning - Vol 1, 2, 3 - 3Rd Edition - c2000 Cshl Press.pdf**, 2001.

SARAYA, R. et al. Novel genetic tools for Hansenula polymorpha. **FEMS Yeast Research**, v. 12, n. 3, p. 271–278, 2012.

SCHIRALDI, C.; GATTA, A. LA; ROSA, M. DE. Biotechnological Production and Application of Hyaluronan. **Biopolymers**, 2010.

SCHMIDT, A.; SIEBERT, K. G. and Applications. v. LXX, p. 105–122, 2001.

SCHUTYSER, W. et al. Chemicals from lignin: An interplay of lignocellulose fractionation, depolymerisation, and upgrading. **Chemical Society Reviews**, v. 47, n. 3, p. 852–908, 2018.

SHELDON, R. A.; BRADY, D. The limits to biocatalysis: pushing the envelope. **Chemical Communications**, v. 54, n. 48, p. 6088–6104, 2018.

SHELDON, R. A.; WOODLEY, J. M. Role of Biocatalysis in Sustainable Chemistry. **Chemical Reviews**, v. 118, n. 2, p. 801–838, 2018.

SILVA-ROCHA, R. et al. The Standard European Vector Architecture (SEVA): A coherent platform for the analysis and deployment of complex prokaryotic phenotypes. **Nucleic Acids Research**, v. 41, n. D1, p. 666–675, 2013.

SIMON, O. et al. Comprehensive proteome analysis of the response of Pseudomonas putida KT2440 to the flavor compound vanillin. **Journal of Proteomics**, v. 109, p. 212–227, 2014.

SOHN, J. et al. A Novel Autonomously Replicating Sequence (ARS) for Multiple Integration in the Yeast Hansenula polymorpha DL-1. **Journal of Bacteriology**, v. 178, n. 15, p. 4420–4428, 1996.

SOHN, J. H. et al. A dominant selection system designed for copy-number-controlled gene integration in Hansenula polymorpha DL-1. **Applied Microbiology and Biotechnology**, v. 51, n. 6, p. 800–807, 1999.

STEFFEN-MUNSBERG, F. et al. Bioinformatic analysis of a PLP-dependent enzyme superfamily suitable for biocatalytic applications. **Biotechnology Advances**, v. 33, n. 5, p. 566–604, set. 2015.

STÖCKMANN, C. et al. Process development in Hansenula polymorpha and Arxula adeninivorans, a re-assessment. **Microbial Cell Factories**, v. 8, n. 22, 2009.

SUH, S. O.; ZHOU, J. J. Methylotrophic yeasts near Ogataea (Hansenula) polymorpha: A proposal of Ogataea angusta comb. nov. and Candida parapolymorpha sp. nov. **FEMS Yeast Research**, v. 10, n. 5, p. 631–638, 2010.

SUN, J. et al. Genome editing and transcriptional repression in Pseudomonas putida KT2440 via the type II CRISPR system. **Microbial Cell Factories**, v. 17, n. 1, p. 1–17, 2018a.

SUN, Z. et al. Bright Side of Lignin Depolymerization: Toward New Platform Chemicals. **Chemical Reviews**, v. 118, n. 2, p. 614–678, 2018b.

SUPPI, S. et al. Repression vs. activation of MOX, FMD, MPP1 and MAL1 promoters by sugars in Hansenula polymorpha: The outcome depends on cell's ability to phosphorylate sugar. **FEMS Yeast Research**, v. 13, n. 2, p. 219–232, 2013.

SZE, J.; BROWNLIE, J. C.; LOVE, C. A. Biotechnological production of hyaluronic acid: a mini review. **3 Biotech**, v. 6, n. 1, p. 1–9, 2016.

TIKHOMIROVA, L. P. et al. Transformation of methylotrophic yeast Hansenula polymorpha: cloning and expression of genes. v. 28, p. 343–351, 1988.

TLAPAK-SIMMONS, V. L. et al. The active streptococcal hyaluronan synthases (HASS) contain a single HAS monomer and multiple cardiolipin molecules. **Journal of Biological Chemistry**, v. 273, n. 40, p. 26100–26109, 1998.

TLAPAK-SIMMONS, V. L. et al. Purification and lipid dependence of the recombinant hyaluronan synthases from Streptococcus pyogenes and Streptococcus equisimilis. **Journal of Biological Chemistry**, v. 274, n. 7, p. 4239–4245, 1999.

TSEPILOV, R. N.; BELODED, A. V. Hyaluronic acid - An "old" molecule with "new" functions: Biosynthesis and depolymerization of hyaluronic acid in bacteria and vertebrate tissues including during carcinogenesis. **Biochemistry (Moscow)**, v. 80, n. 9, p. 1093–1108, 2015.

VAN DER KLEI, I. J.; VEENHUIS, M. Hansenula polymorpha: A Versatile Model Organism in Peroxisome Research. **Hansenula polymorpha**, p. 76–94, 2002.

VAN DIJK, R. et al. The methylotrophic yeast Hansenula polymorpha: A versatile cell factory. **Enzyme and Microbial Technology**, v. 26, n. 9–10, p. 793–800, 2000.

WANG, L. et al. Biotechnology for Biofuels Efficient CRISPR – Cas9 mediated multiplex genome editing in yeasts. **Biotechnology for Biofuels**, v. 277, n. 11, p. 1–16, 2018.

WANG, Y. et al. Eliminating the capsule-like layer to promote glucose uptake for hyaluronan production by engineered Corynebacterium glutamicum. **Nature Communications**, v. 11, n. 1, 2020.

WEBER, N. et al. Biocatalytic potential of vanillin aminotransferase from Capsicum chinense. **BMC Biotechnology**, v. 14, n. 1, p. 1–6, 2014.

WEBER, N.; GORWA-GRAUSLUND, M.; CARLQUIST, M. Exploiting cell metabolism for biocatalytic whole-cell transamination by recombinant Saccharomyces cerevisiae. **Applied Microbiology and Biotechnology**, v. 98, n. 10, p. 4615–4624, 2014.

WEHRS, M. et al. Engineering Robust Production Microbes for Large-Scale Cultivation. **Trends in Microbiology**, v. 27, n. 6, p. 524–537, 2019.

WEIGEL, P. H.; HASCALL, V. C.; TAMMI, M. Hyaluronan synthases. **Journal of Biological Chemistry**, v. 272, n. 22, p. 13997–14000, 1997.

WEIMER, A. et al. Industrial biotechnology of Pseudomonas putida: advances and prospects. **Applied Microbiology and Biotechnology**, 2020.

WESTBROOK, A. W. et al. Engineering of cell membrane to enhance heterologous production of hyaluronic acid in Bacillus subtilis. **Biotechnology and Bioengineering**, v. 115, n. 1, p. 216–231, 2018.

WIRTH, N. T.; KOZAEVA, E.; NIKEL, P. I. **Accelerated genome engineering of Pseudomonas putida by I-SceI—mediated recombination and CRISPR-Cas9 counterselectionMicrobial Biotechnology**, 2020.

WU, S. et al. Highly regio- and enantioselective multiple oxy- and amino-functionalizations of alkenes by modular cascade biocatalysis. **Nature Communications**, v. 7, n. 1, p. 11917, 14 set. 2016.

WU, S.; LIU, J.; LI, Z. Biocatalytic Formal Anti-Markovnikov Hydroamination and Hydration of Aryl Alkenes. **ACS Catalysis**, v. 7, n. 8, p. 5225–5233, 4 ago. 2017.

XU, S. et al. Capsaicin inhibits the metastasis of human papillary thyroid carcinoma BCPAP cells through the modulation of the TRPV1 channel. **Food and Function**, v. 9, n. 1, p. 344–354, 2018.

YANG, L. et al. Permanent genetic memory with >1-byte capacity. **Nature Methods**, v. 11, n. 12, p. 1261–1266, 26 dez. 2014.

ZHOU, Y. et al. Bioproduction of benzylamine from renewable feedstocks via a nine-step artificial enzyme cascade and engineered metabolic pathways. **ChemSusChem**, v. 11, n. 13, p. 2221–2228, 2018.

5. PARTICIPATION IN EVENTS

1) **MANFRAO NETTO, J. H. C.; RODRIGUES, K. A.; QUEIROZ, E. B.; PARACHIN, N. S.**

Event: XXII Simpósio Nacional de Bioprocessos e XIII Simpósio de Hidrólise Enzimática de Biomassas; **Inst.promotora**: UFU

Title: Hyaluronic Acid production in the yeast *Hansenula polymorph*a, 2019

2) **MANFRAO NETTO, J. H. C.; CARVALHO, L. S.; RODRIGUES, K. A.; PARACHIN, N. S.**

Event: International Specialized Symposium on Yeasts (ISSY34); **Inst.promotora**: IPATEC

Title: Development of a *Hansenula polymorpha* strain producing Hyaluronic Acid, 2018.

3) **MANFRÂO-NETTO, J.H.C.; RODRIGUES, K. A.; QUEIROZ, E. B.; PARACHIN, N. S.**

Event: VIII Simpósio do Programa de Pós-Graduação em Ciências Biológicas (Biologia Molecular); **Inst.promotora**: Universidade de Brasília

Title: Edição do genoma de Hansenula Polymorpha visando a identificação de etapas limitantes na produção de Ácido Hialurônico, 2018.

4) **MANFRÃO-NETTO, JOÃO HEITOR C.; RODRIGUES, K. A.; QUEIROZ, E. B.; PARACHIN, N. S.**

Event: 7th Brazilian Biotechnology Congress and 2nd Biotechnology Ibero-American Congress; **Inst.promotora**: Sociedade Brasileira de Biotecnologia

Title: Hyaluronic acid production in the methylotrophic yeast Hansenula polymorpha, 2018.

5) **MANFRAO NETTO, J. H. C.; CARVALHO, L. S.; PARACHIN, N. S.**

Event: VII Simpósio do Programa de Pós-Graduação em Ciências Biológicas (Biologia Molecular); **Inst.promotora**: Universidade de Brasília

Title: Estudo da fisiologia de Hansenula Polymorpha usando o sistema CRISPR/Cas para a identificação de etapas limitantes na produção de Ácido Hialurônico, 2017**.**

6. PRODUCTIONS

6.1 PAPER I

MANFRÃO-NETTO, J. H. C.; GOMES, A. M. V.; PARACHIN, N. S. Advances in Using Hansenula polymorpha as Chassis for Recombinant Protein Production. **Frontiers in Bioengineering and Biotechnology**, v. 7, n. May, p. 1–13, 2019.

Contribution: JM-N wrote the manuscript.

6.2 PAPER II

MANFRÃO-NETTO, J. H. C. et al. Evaluation of ogataea (Hansenula) polymorpha for hyaluronic acid production. **Microorganisms**, v. 9, n. 2, p. 1–16, 2021a.

Contribution: J.H.C.M.-N. constructed the strains, performed the growth curves, prepared the samples for scanning electron microscopy analysis, analyzed the data, and drafted the manuscript.

6.3 PAPER III

MANFRÃO-NETTO, J. H. C. et al. Metabolic engineering of Pseudomonas putida for production of vanillylamine from lignin-derived substrates. **Microbial Biotechnology**, 2021b.

Contribution: J.H.C.M.N. contributed to the design of the study, performed the bioinformatics analysis, design and performed the experiments, analysed the data and drafted the manuscript.

6.4 PAPER IV (SUBMITTED)

MANFRÃO-NETTO, J. H. C. et al. Genetic strategies for improving Hyaluronic Acid production in recombinant bacterial culture. **Journal of Applied Microbiology**, under review.

Contribution: JHCMN wrote and edited the manuscript.

frontiers in Bioengineering and Biotechnology

REVIEW
published: 01 May 2019
doi: 10.3389/fbioe.2019.00094

Advances in Using *Hansenula polymorpha* as Chassis for Recombinant Protein Production

*João Heitor Colombelli Manfrão-Netto, Antônio Milton Vieira Gomes and Nádia Skorupa Parachin**

Grupo Engenharia de Biocatalisadores, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília, Brazil

The methylotrophic yeast *Hansenula polymorpha*, known as a non-conventional yeast, is used for the last 30 years for the production of recombinant proteins, including enzymes, vaccines, and biopharmaceuticals. Although a large number of reviews have been published elucidating the applications of this yeast as a cell factory, the latest was released about 10 years ago. Therefore, this review aimed at summarizing available information on the use of *H. polymorpha* as a host for recombinant protein production in the last decade. Examples of chemicals and virus-like particles produced using this yeast also are discussed. Firstly, the aspects that feature this yeast as a host for recombinant protein production are highlighted including the techniques available for its genetic manipulation as well as strategies for cultivation in bioreactors. Special attention is given to the novel genomic editing tools, mainly CRISPR/Cas9 that was recently established in this yeast. Finally, recent examples of using *H. polymorpha* as an expression platform are presented and discussed. The production of human Parathyroid Hormone (PTH) and Staphylokinase (SAK) in *H. polymorpha* are described as case studies for process establishment in this yeast. Altogether, this review is a guideline for this yeast utilization as an expression platform bringing a thorough analysis of the genetic aspects and fermentation protocols used up to date, thus encouraging the production of novel biomolecules in *H. polymorpha*.

Keywords: *Hansenula polymorpha*, recombinant protein, methylotrophic yeast, genomic editing, bioprocess

OPEN ACCESS

Edited by:
C. Perry Chou,
University of Waterloo, Canada

Reviewed by:
Hyun Ah Kang,
Chung-Ang University, South Korea
Kostyantyn Dmytruk,
Institute for Cell Biology, National Academy of Sciences of Ukraine (NAN Ukraine), Ukraine
Nikolai Ravin,
Institute of Bioengineering, Research Center of Biotechnology of the Russian Academy of Sciences, Russia

****Correspondence:***
Nádia Skorupa Parachin
nadiasp@unb.br;
nadiasp@gmail.com

Specialty section:
This article was submitted to Synthetic Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *18 January 2019*
Accepted: *16 April 2019*
Published: *01 May 2019*

Citation:
Manfrão-Netto JHC, Gomes AMV and Parachin NS (2019) Advances in Using Hansenula polymorpha as Chassis for Recombinant Protein Production. Front. Bioeng. Biotechnol. 7:94. doi: 10.3389/fbioe.2019.00094

INTRODUCTION

Over the years, the use of unicellular microorganisms as cell factories to obtain recombinant proteins became consolidated (Kim et al., 2015a). Recombinant DNA techniques allow the introduction of foreign genes in a host organism for the production of heterologous proteins biologically actives. Within this context, the choice of the host organism is crucial, since the functionality, solubility, and activity of the protein must be preserved during its synthesis (Vaquero et al., 2015). Yeasts are commonly used for the heterologous production of proteins, especially those that require post-translational modifications for proper folding since these modifications occur less frequently in prokaryotes.

The *H. polymorpha* is commonly employed as an expression platform because of its unique characteristics (**Figures 1A–E**). It is thermotolerant and capable of growing at temperatures ranging from 30 to 50°C (**Figure 1B**). This capability is advantageous regarding mammalian protein

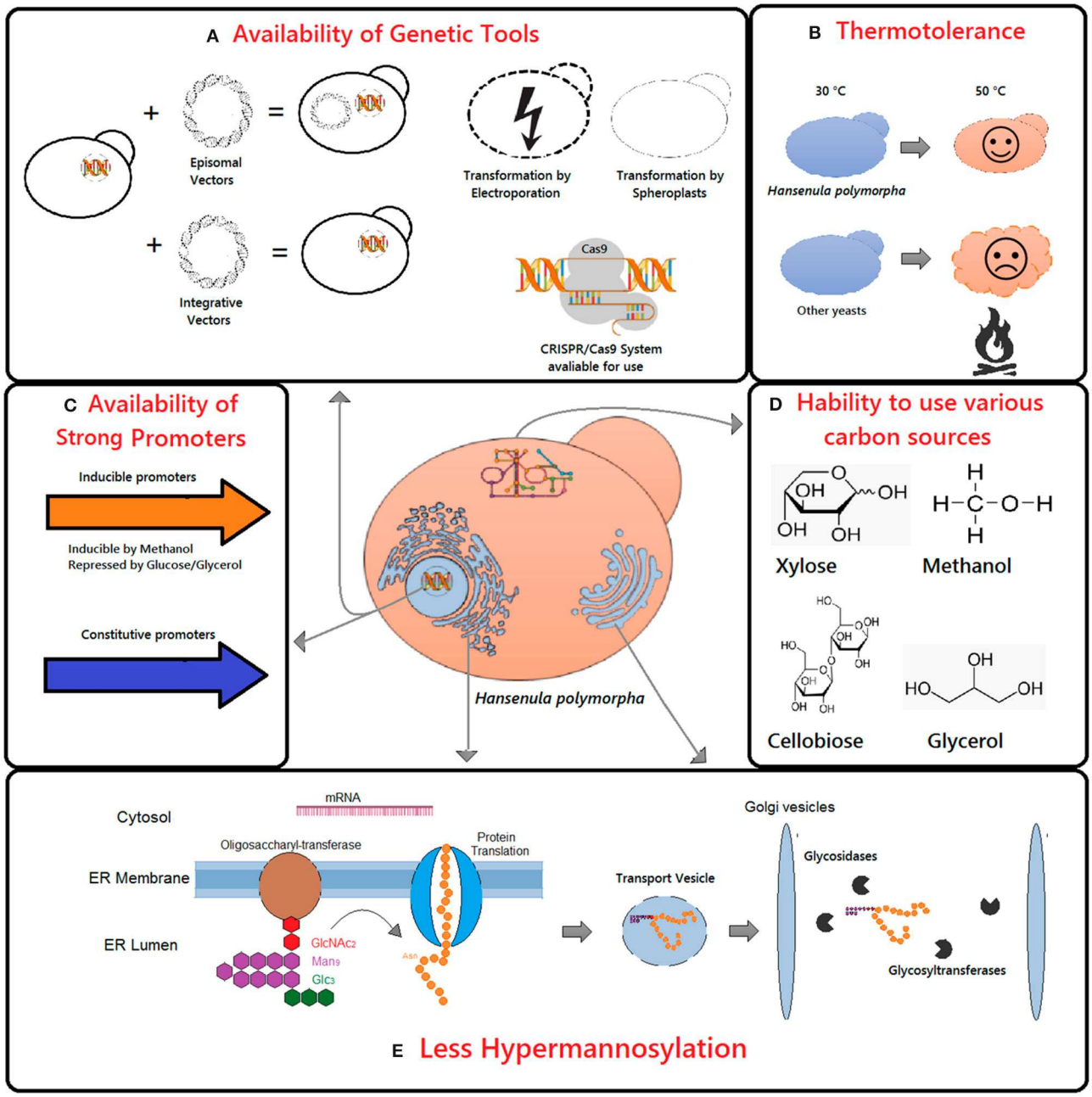


FIGURE 1 | Main advantages of *Hansenula polymorpha* as chassi for recombinant protein production include the availability of genetic tools **(A,C)**, thermotolerance **(B)**, ability to use various carbon sources **(D)**, and glycosylation pattern **(E)**.

production such as those requiring the 37°C temperature to preserve its biological activity (Van Dijk et al., 2000). Moreover, the presence of protein glycosylation pathway in *H. polymorpha* allows the production of eukaryotic recombinant proteins biologically active. Additionally, unlike other yeasts, it adds fewer sugar residues to the protein core, avoiding hyperglycosylation of recombinant proteins (**Figure 1E**). Finally, *H. polymorpha* is capable of using methanol as a carbon source which allowed the isolation of strong methanol inducible promoters (**Figure 1C**). Besides, it can utilize other carbon sources such as glycerol, glucose, xylose, and cellobiose (Ryabova et al., 2003) (**Figure 1D**).

Three parental strains with distinct origins of *H. polymorpha* are frequently used for recombinant protein production. The DL-1 strain (NRRL-Y-7560; ATCC26012) was isolated and characterized from soils samples (Levine and Cooney, 1973). The CBS4732 strain (CCY38-22-2; ATCC34438, NRRL-Y-5445) was isolated in irrigated soils in Pernambuco, Brazil (Morais and Maia, 1959). These two strains are mostly employed for industrial

use. Lastly, the NCYC495 strain (CBS1976; ATAA14754, NRLL-Y-1798) is commonly used in the laboratory and was isolated at Florida from concentrated orange juice (Wickerham, 1951). Phylogenetic analysis showed that *H. polymorpha* appears to be two different species: *Ogataea polymorpha and Ogataea parapolymorpha* (Kurtzman and Robnett, 2010; Suh and Zhou, 2010). The strain NCYC495 and CBS4732 are closely related and renamed as *O. polymorpha*, whereas DL-1 strain is phylogenetically distant and reclassified as *O. parapolymorpha*. To avoid misunderstanding, the nomenclature *H. polymorpha* will be used in this review once both species share all characteristics elucidated in **Figure 1**.

Various studies focused on genetically modifying *H. polymorpha* strains for the production of several recombinant proteins (Gellissen et al., 1992; Hollenberg and Gellissen, 1997; Stöckmann et al., 2009). Later on, the advances in genomic-editing tools, optimization of transformation and cultivation protocols have led to the industrial development of *H. polymorpha*-based processes for the production of various pharmaceuticals. Currently, three commercially available Hepatitis B vaccines are produced using antigens derived from fermentative processes with *H. polymorpha:* HepavaxGene® (Johnson & Johnson), Gen Vax B® (Serum Institute of India) and Biovac-B® (Wockhardt) (http://www.dynavax.com/about-us/dynavax-gmbh/). Moreover, biopharmaceuticals successfully produced in this yeast and already available in the market included hirudin (Thrombexx®, Rhein Minapharm), insulin (Wosulin®, Wockardt) and IFNa-2a Reiferon® (Rhein Minapharm) (Gellissen et al., 2005).

It is noteworthy that the last published review of bioprocess development at *H. polymorpha* was nearly 10 years ago (Stöckmann et al., 2009). Thus, this review brings up to date strategies and examples of using this yeast as a host for recombinant protein production. The focus will be given on the studies developed in the last decade and are summarized in **Table 1**. The relevance of this yeast for the production of recombinant proteins, especially those for human welfare, justifies this literature update. Besides, newly genomic tools developed in the past years which have improved genetic manipulation of *H. polymorpha* are also discussed.

WHY USE *H. POLYMORPHA* AS HOST FOR HETEROLOGOUS EXPRESSION?

The advantages of *H. polymorpha* for industrial processes comprise high-cell-density fermentation, capacity to utilize low-cost substrates, an established defined synthetic media, status GRAS (Generally Regarded As Safe) and consolidated strategies for cultivation in bioreactors (Jenzelewski, 2002). This yeast features genome-editing tools available for genetic manipulation (**Figure 1A**). An efficient protocol for transformation by electroporation has been described previously (Faber et al., 1994) as well as protocols for transforming protoplast ((Tikhomirova et al., 1988)). Among them, the electroporation method is more efficient than the protoplast, yielding $1.7 \times 10^6/\mu g$ plasmid DNA vs. 2 to $3 \times 10^4/\mu g$ DNA. The lithium acetate-dimethyl sulfoxide method has also been used tested (Sohn et al., 1999; Heo et al., 2003; Kim et al., 2015b). Furthermore, a method using nanoscale carriers for DNA delivery was employed for the transformation of *H. polymorpha* with twice efficiency of those obtained by electroporation and 15-fold for LiAc/DMSO method (Filyak et al., 2013). Moreover, three independent research groups have recently developed the CRISPR/Cas9 genome-editing tool for *H. polymorpha* (Numamoto et al., 2017; Juergens et al., 2018; Wang et al., 2018). Finally, the three strains of *H. polymorpha* had its genome sequenced, DL-1 (Ravin et al., 2013), NCYC495 (Riley et al., 2016), and CBS4732 (Ramezani-Rad et al., 2003).

Strategies for heterologous protein production in *H. polymorpha* take advantage of the yeast ability to grow in the presence of methanol. The methanol inducible promoters, formate dehydrogenase (*FMD*), and methanol oxidase (*MOX*) are the most utilized in genetic engineering strategies as it can be seen in **Table 1**. Shifting to methanol-feed led to upregulation of genes involved in its catabolism, for example the *FMD* gene was approximately 350-fold upregulated, while the *MOX* and *DHAS* genes were 17.3 and 19-fold upregulated when compared to growth on glucose (van Zutphen et al., 2010). Although an upregulation does not necessarily indicate a high promoter activity, other studies have shown that in the presence of methanol the *MOX* and *FMD* promoters present an enhanced activity (Amuel et al., 2000; Suppi et al., 2013).

The methanol-inducible promoters are not present only in *H. polymorpha* but in all methylotrophic yeasts. For instance, the well-known yeast *Pichia pastoris* (recently renamed as *Komagataella* sp.) is the yeast host more utilized for recombinant protein production. Although *P. pastoris* also has methanol-inducible promoters, the advantage of using *H. polymorpha* is that some of them are derepressed in the presence of glycerol which is less pronounced in *P. pastoris* (60–70% vs. 2–4% of induced levels) (Hartner and Glieder, 2006; Vogl and Glieder, 2013). In a comparative study, The Kunitz-type protease inhibitor (*KPI*) encoding gene was inserted in *H. polymorpha* and *P. pastoris* under the control of the alcohol oxidase, *AOX1* promoter (Raschke et al., 1996). For both yeasts, no mRNA encoding for KPI was detected when the cells were cultured in glucose as the carbon source but were abundant when induced by methanol. However, when cells grew in glycerol, it was possible to detect KPI only in *H. polymorpha*. Therefore, the derepression of methanol inducible promoters is a favorable feature of *H. polymorpha* over other methylotrophic yeasts.

Additionally, *H. polymorpha* is thermotolerant while *P. pastoris* is not (**Figure 1B**). The increase in temperature does not imply a higher yield of the recombinant protein but is relevant in industrial processes since it reduces microbial contamination and cooling costs (Abdel-Banat et al., 2010). Also, higher temperatures facilitate the implementation of Simultaneous Saccharification and Fermentation (SSF) since the thermal resistance allows the utilization of thermophilic hydrolases (Voronovsky et al., 2009). The *EGII* gene encoding endoglucanase II from *Trichoderma reesei* was produced in both methylotrophic yeasts, and the recombinant proteins were characterized. Although the secreted enzymes showed optimum activity at the same temperature (75°C), the one

TABLE 1 | Recombinant proteins produced in the last decade using *H. polymorpha* as host.

Protein	Maximum production	Promoter	Utilized carbon source	References
Human serum albumin (HSA)	5.8 g/L	*MOX*	Glycerol/Methanol	Youn et al., 2010
Heat shock protein gp96	≈150 mg/L	*FMD*	Methanol	Li et al., 2011
Ferritin (FTH1)	1.9 g/L	*FMD*	Glycerol/Methanol	Eilert et al., 2012
Bacteriocin enterocin A (EntA)	4.8 μg/mL	*TEF1* (*Arxula adeninivorans*)	Glucose	Borrero et al., 2012
Granulocyte colony stimulating factor (GCSF)	ND	*FMD*	Methanol	Talebkhan et al., 2016
Streptavidin (SAV)	≈751 mg/L	*FMD*	Methanol	Wetzel et al., 2016
Human parathyroid hormone (PTH) fragment 1–34	150 mg/L	*FMD*	Glycerol/Methanol	Mueller et al., 2013
Penicillin	1.1 μg/mL	*MOX*	Glucose/Methanol	Gidijala et al., 2009
Human papilomavirus 16 L1 Protein (HPV16L1)	78.6 mg/L	*MOX*	Methanol	Li et al., 2009
HPV type 16 L1-L2 chimeric protein (SAF)	132.10 mg/L	*MOX*	Methanol	Bredell et al., 2018
Rabies virus glycoprotein (RVG)	14.6 mg/L	*FMD*	Glycerol	Qian et al., 2013
Hepatitis B virus PreS2-S antigen	250 mg/L	*MOX*	Methanol	Xu et al., 2014
Human papillomavirus Type 52 L1 Protein (HPV 52 L1)	ND	*MOX*	Methanol	Liu et al., 2015
Rotavirus VP6 protein (RV VP6)	3350.71 mg/L	*MOX*	Methanol	Bredell et al., 2016
Hepatitis E virus-like particles (HEV VLPs)	1.0 g/L	*MOX*	Methanol	Su et al., 2017
Uricase from *Candida utilis*	52.3 U/mL	*MOX*	Methanol	Chen et al., 2008
Lipase from *Yarrowia lipolytica* (YlLip11)	1,144 U/L	*TEF1* (*Arxula adeninivorans*)	Glucose	Kumari et al., 2015
T4 lysozyme	0.49 g/L	*Not used*	Glycerol	Wang et al., 2011
Staphylokinase (SAK)	1,212 mg/L	*FMD*	Glycerol/Methanol	Moussa et al., 2012

ND, Not Determined.

produced by *H. polymorpha* was shown to be more thermostable while remaining active after incubation at high temperatures (60–80°C) (Akbarzadeh et al., 2013).

If the recombinant protein is to be produced under the control of a methanol inducible promoter, the cultivation in the bioreactor is performed in a two-step. Initially, the growth phase is performed using glucose or glycerol as a carbon source in batch cultures. Then, the induction phase is accomplished by feeding the bioreactor with methanol or methanol/glycerol mixtures that can be added continuously or in pulses. For instance, the production of ferritin in *H. polymorpha* using a methanol and glycerol mixture (4:1) during the induction phase resulted in 1.9 g L of the recombinant protein (**Table 1**) (Eilert et al., 2012). In the case of the rabies virus glycoprotein production, only glycerol was employed during the induction phase (Qian et al., 2013). Despite the low levels achieved (14.6 mg/L), this is an example that glycerol can be used in both the growth and induction phases of this yeast. Although these two strategies are feasible to control recombinant protein production, the most common strategy employed is induction using only methanol 0.5–1% (**Table 1**).

A disadvantage of using such methanol inducible promoters is their repression caused by the presence of glucose in the media, although the derepression has already been reported (Mayer et al., 1999). In this point of view, *H. polymorpha* strains deficient in glucose repression represent alternative platforms for recombinant protein production (Krasovska et al., 2007). These mutants have a knock out at the *GCR1* gene that encodes a hexose transporter with altered activity leading to several alterations in the cell metabolism, including the derepression of methanol-related genes in the presence of glucose (Stasyk et al., 2004). Thus, the methanol-induced promoters controlling the recombinant protein production might be induced by either methanol or glucose. The *H. polymorpha gcr1* mutants are commonly utilized as the bio-elements for Yeast-based biosensing. Some practical examples of these biosensors utilization include the detection of L-lactate (Smutok et al., 2007), urate (Dmytruk et al., 2011), formaldehyde (Sigawi et al., 2014), and D-lactate (Smutok et al., 2018).

Some *H. polymorpha*-based platforms utilize the strong constitutive promoter *GAP* instead of methanol-inducible ones (Heo et al., 2003). The use of this promoter for the construction of recombinant strains, as well as *gcr1* mutants, enables the production of proteins without the addition of methanol. Since this molecule is flammable and toxic, avoiding its use can be advantageous. Indeed, several studies developing *H. polymorpha* strains for bioethanol production utilize genes under the control of the *GAP* promoter (Kurylenko et al., 2014, 2018). Examples of the utilization of the *H. polymorpha GAP* promoter included the development of thermotolerant strains, improvement of xylose utilization and those capable of Simultaneous Saccharification and Fermentation (SSF) as recently reviewed (Dmytruk et al., 2017).

Another advantage of using *H. polymorpha* as a host for recombinant protein production is its glycosylation pattern (**Figure 1E**). The yeasts frequently hyperglycosylate recombinant proteins. However, the intensity and type of sugar added

dependent on both the organism and the sequence of the heterologous protein. The production of a recombinant glucose oxidase of *Aspergillus niger* was attempted using both yeasts *H. polymorpha* and *S. cerevisiae* (Kim et al., 2004). It has been shown that in *H. polymorpha* 27% less glycosylation was observed. Also, antibody anti-α1,3-mannose did not recognize the protein produced by *H. polymorpha* but was positive for that originated from *S. cerevisiae* which indicates that in *H. polymorpha* the recombinant protein was not immunogenic (Ballou, 1990). In the following years, efforts were made to develop engineered strains with human-pattern glycosylation (Kim et al., 2006; Oh et al., 2008; Cheon et al., 2012). These strains lack essential genes which encode enzymes for hypermannosylation pathways such as α-1,6-mannosyltransferase (Δ*hpoch1*) and dolichyl-phosphate-mannose dependent α-1,3-mannosyltransferase (Δ*hpalg3*) beside their has the human gene encoding β-1,2-N-acetylglucosaminyltransferase I (*GNTI*). The null mutants were able to produce human hybrid-type N-glycans (Cheon et al., 2012). Therefore, all the knowledge acquired about the physiology, metabolism, and genetics of this yeast enables its utilization as a host for heterologous protein production.

THE *H. POLYMORPHA* GENETIC ENGINEERING TOOLS

The viability of tools for fast and precise genome edition is crucial for the development of the expression platforms. Some methods for the genetic manipulation of *H. polymorpha* have already been described, both for gene introduction and deletion (**Figure 1A**). It has been previously reported that episomal plasmids are mitotically unstable in *H. polymorpha* (Bogdanova et al., 1995, 2000) and consequently are not suitable for the development of industrial strains. Episomal plasmids contain the *H. polymorpha* autonomous replicating sequences *(HARS*) derived from subtelomeric regions (Sohn et al., 1996). Nevertheless, the *HARS* sequences do not guarantee stability for circular plasmids. Thus, the prolonged incubation in selective medium forces the plasmid integration usually in the respective subtelomeric locus (Kim et al., 2003). Hence, the integrative plasmids are most suitable for genetic manipulation of *H. polymorpha.* Depending on the locus of integration it is possible to reach between 1 up to 100 copies into the genome (Agaphonov et al., 1999).

The *H. polymorpha* integration plasmids (pHIP) series have several promoters and selective markers (For detailed information see (Saraya et al., 2012). They show an easy terminology in which the letter indicates the selective marker and the number indicates the promoter. For example, the pHIPH15, the "HIP" means "*Hansenula* Integration plasmid" while the letter stands for the selective marker, in this example hygromycin. The number indicates the promoter utilized in the plasmid and "15" is for *DHAS*. There are fifteen *H. polymorpha* promoters available in the pHIP series (https://www.rug.nl/research/molecular-cell-biology/research/the-hansenula-polymorpha-expression-system), and the site-specific integration is driven by linearization in the promoter region of the plasmid. For selection, auxotrophic and dominant markers can be used. They include *Sc LEU2*, *Hp URA3*, *Hp ADE11*, *Hp MET6*, *Sh-ble* (Zeocine), *Sn-nat1* (Nourseothricin), *Kp-hph* (Hygromycin B), and *Tn-KanMX* (G418/Geneticin) (Saraya et al., 2012). Another possibility for the introduction of an exogenous gene in *H. polymorpha* is the wide-range vector CoMed™ system (Böer et al., 2007). This vector was designed to fit in many species of yeast, enabling to save time and effort during the cloning procedure. The vector contains *ARS* and rDNA sequences that drive the integration of the plasmid into *H. polymorpha* genome. Besides, it was constructed in modules flanked by recognizing sites of restriction enzymes that allows the exchange of expression cassettes and selective markers.

Gene deletion in *H. polymorpha* is reached through the construction of cassettes containing a homologous region of the gene to be deleted flanking 5′ and 3′ of the target locus. Usually, the cassette has an antibiotic as a selection marker such as zeocin or hygromycin. Although there are different approaches for gene deletion in *H. polymorpha*, the disruption frequencies are of approximately 35% with homology arms ranging from 500 to 1,000 bp (**Table 2**). The deletion cassettes can be constructed by single-step PCR as an adaptation of the protocol previously utilized in *S. cerevisiae* (Manivasakam et al., 1995). A one-step mediated-PCR method for gene disruption was also previously reported (Gonzalez, 1999). As a proof concept, the gene *YNR1* of *H. polymorpha* strain NCYC495 (*ura3*) was disrupted utilizing a construct bearing the *URA3* auxotrophic marker flanked by homologous regions to the target gene. The homologous arms 5‘ and 3‘ tested varied between 25 and 1,000 bp in size with the best results obtained with the larger homology (35%). Similar results were observed in the disruption of *MOX* gene (36%) with 1,000/1,000 homologous arm size (Gonzalez, 1999) (**Table 2**).

In another study, the deletion cassettes were designed containing flanking regions with different lengths for targeting *MOX* locus. In size, the homologous arms were tested between 30 and 250 bp. The deletion frequency for fragments of up to 50 bp was only ±12%. The larger homologous arm tested showed an incidence of 31% (**Table 2**). Another strategy adopted was the deletion of *ku80* aiming at increasing the deletion frequency (Saraya et al., 2012). This gene deletion increases the repair by homologous recombination (HR) instead of non-homologous end joining (NHEJ), ensuring site-specific integration. The Δ*yku80* strain was generated by replacing this gene with the *URA3* gene. For cassettes designed to target in *MOX* gene with Hygromycin B as selection marker and flanking regions approximately 250 bp, the deletion efficiency was 88% in Δ*yku80* strain vs. 31% in the wild-type (**Table 2**).

The limited number of available markers can impair multiple gene insertion or deletion. Therefore, recycling the selective markers using the Cre–*loxP* recombination technique has been applied in *H. polymorpha* (Qian et al., 2009; Agaphonov and Alexandrov, 2014). The gene *PMC1* encoding for the protein Calcium-transporting ATPase 2 was disrupted in the strain DL1-L (*leu2*) using auxotrophic marker *LEU2.* The *MOX* promoter drives the *CRE* gene expression, so when methanol was added into the cultivation medium, the cassette

TABLE 2 | Most common techniques used for gene deletion in *H. polymorpha*.

Target locus	Strain	Frequency of Deletion %	Method	Homologous arm size (5′/3′) bp	References
YNR1	NCYC495	35	Deletion cassette	1,000/1,000	Gonzalez, 1999
MOX	NCYC495	36	Deletion cassette	1,000/1,000	
MOX	NCYC495 Δ*yku80*	88	Deletion cassette	245/247	Saraya et al., 2012
MOX	NCYC495	31	Deletion cassette	245/247	
ALG3	NCYC495	35	Deletion cassette	491/520	Qian et al., 2009
ALG3	NCYC495	76	Co-transformation with single-stranded DNA	491/520	
ALG3	NCYC495	19	Deletion cassette	∼250/250	
ALG3	NCYC495	33	Co-transformation with single-stranded DNA	∼250/250	

contained the marker *LEU2* was excised from the genome. Furthermore, these clones were unable to grow in the presence of $CaCl_2$ corroborating the *PMC1* deletion. The combination of different approaches was utilized to increase the efficiency for gene deletion in *H. polymorpha* NCYC495 (Qian et al., 2009). The knock-out system uses a sticky-end polymerase chain reaction method for the construction of deletion cassettes, LiAc/single-stranded (SS)-DNA/PEG for cell transformation and l*oxP*-flanked selective markers for multiple deletions. The main advantage of this approach is the use of single-stranded DNA for co-transformation with deletion cassettes. Two genes were targeted, *URA5* and *ALG3*, encoding the orotate-phosphoribosyl transferase (OPRTase) and alpha-1,3-mannosyltransferase, respectively. The cassette bearing *loxP-kanMX-loxP* with ∼500 bp or ∼250 bp homologous arms were co-transformed with and without single-stranded DNAs. For gene *ALG3*, the frequency of homologous recombination increased from 19 to 33% with ∼250 bp homologous arms in the presence of single-stranded DNA (**Table 2**). When 500 bp homology regions were utilized, the frequency raised from 35 to 76% for co-transformation with single-stranded DNA (**Table 2**). The same pattern was observed for the *URA5* locus, for 250 bp 17% and 500 bp 31% without co-transformation while in the presence of single-stranded DNA were 32 and 73%, respectively (Qian et al., 2009).

At present, the CRISPR/Cas9 technologies have been applied in various organisms aiming at genome edition by the promise of being more rapid and low cost than previously available technologies [see details in Donohoue et al. (2018)]. Three studies already reported the use of CRISPR/Cas9 in *H. polymorpha* (**Table 3**). In the first one, the locus *ADE2* in the NCYC495 was targeted due to the red-phenotype that can be easily visualized upon successful deletion (Numamoto et al., 2017). For that, an episomal plasmid expressing both Cas9 and gRNA guided by the promoters *ScTEF1* and *OpSNR6*, respectively was introduced in *H. polymorpha*. In the absence of the *ADE2* DNA donor, the deletion frequency was around 10^{-3}. When a DNA donor with 60/60 bp of homologous arms for *ADE2* locus was co-transformed with Cas9 and $gRNA_{ADE2}$, it increased the efficiency by up to 47%. To evaluate the system efficiency in another locus, the loci *ADE8* and *PHO8* were disrupted resulting in 0.36 and 0.08% efficiencies, respectively. Aiming at improving the CRISPR/Cas9 system in this yeast, the promoters guiding Cas9 and gRNA were substituted for *OpTDH3* and *tRNA* promoters, respectively. The cells were then transformed with these plasmids into two combinations: *ScTEF1* (Cas9) and *tRNA* ($gRNA_{ADE2}$) promoters and *OpTDH3* (Cas9) and *tRNA* ($gRNA_{ADE2}$) promoters. Without the DNA donor, the first combination reached a frequency of 38% whiles the second one 45% for the locus *ADE2*. Finally, the improved system in that *OpTHD3* promoter guide the Cas9 expression and the *tRNA* promoter the gRNA expression was used for deletion of three genes from phosphate signal transduction (PHO) pathway, *PHO1*, *PHO11,* and *PHO84*. In this case, more than one gRNA was tested for each locus, and no DNA donor was utilized. Four gRNAs were designed for the *PHO1* gene, however, in 2 of them, the disruption efficiency was 0% while for another two 50, and 71% were observed. For the locus *PHO11*, one of three showed 0% of disruption and 17 and 30% for the other gRNAs. The last locus targeted *PHO84* obtained the same frequency of approximately 67% for both two gRNAs utilized (Numamoto et al., 2017).

In another study, the CRISPR/Cas9 was used for genome edition in both *Kluyveromyces lactis* and *Ogataea* sp. In the last one, the utilized strains were CBS4732 and DL-1 (Juergens et al., 2018). The *AaTEF1* promoter was used to guide a Cas9 variant with higher activity, known as improved Cas9 (iCas9) (Bao et al., 2015). The promoter *ScTDH3* was used to control the gRNA expression. The *ADE2* locus also was disrupted with the same gRNA for both strains *O. polymorpha* (CBS 4732) and *O. parapolymorpha* (DL-1). The disruption frequencies varied between the two species which may indicate that even in phylogenetically close species the efficiency of the CRISPR/Cas9 system may require adjustments. The red phenotype, expected for colonies whose *ADE2* was deleted, just appeared after prolonged incubation time with disruption rates of 9% for *O. polymorpha* while *O. parapolymorpha* showed ∼61 and 63% after 96 and 192 h of incubation, respectively. Surprisingly, the utilization of marker-free DNA donor did not affect the deletion efficiency in this case. Also, a multiple gene editing for multi-locus targeting was employed in this study. A plasmid carrying gRNAs targeting the loci *ADE2* and *YNR1* was utilized

TABLE 3 | CRISPR/Cas9 systems available for genomic editing in *H. polymoprha*.

Strain	Cas9 Promoter	Cas9 origin	sgRNA Promoter	gRNA tested per target	Cas9 and gRNA in the same plasmid?	Plasmid type	target locus	% deletion efficiency	DNA donor (5′/3′)bp	Reference
NCYC495	*TEF1* (*S. cerevisiae*)	*Homo sapiens*	*SNR6* (*H. polymorpha*)	1	Yes	Episomal	*ADE2*	10^{-3}	No	Numamoto et al., 2017
	TEF1 (*S. cerevisiae*)	*Homo sapiens*	*SNR6* (*H. polymorpha*)	1	Yes	Episomal	*ADE2*	47	Yes (60/60)bp	
	TEF1 (*S. cerevisiae*)	*Homo sapiens*	*SNR6* (*H. polymorpha*)	1	Yes	Episomal	*ADE8*	0.36	No	
	TEF1 (*S. cerevisiae*)	*Homo sapiens*	*SNR6* (*H. polymorpha*)	1	Yes	Episomal	*PHO8*	0.08	No	
	TEF1 (*S. cerevisiae*)	*Homo sapiens*	$tRNA^{CUG}$(*H. polymorpha*)	1	Yes	Episomal	*ADE2*	38	No	
	TDH3 (*H. polymorpha*)	*Homo sapiens*	$tRNA^{CUG}$(*H. polymorpha*)	1	Yes	Episomal	*ADE2*	45	No	
	TDH3 (*H. polymorpha*)	*Homo sapiens*	$tRNA^{CUG}$(*H. polymorpha*)	4	Yes	Episomal	*PHO1*	71	No	
	TDH3 (*H. polymorpha*)	*Homo sapiens*	$tRNA^{CUG}$(*H. polymorpha*)	3	Yes	Episomal	*PHO11*	30	No	
	TDH3 (*H. polymorpha*)	*Homo sapiens*	$tRNA^{CUG}$(*H. polymorpha*)	2	Yes	Episomal	*PHO84*	67	No	
DL-1	*TEF1* (*Arxula adeninivorans*)	iCas9	*TDH3* (*S. cerevisiae*)	1	Yes	Episomal	*ADE2*	63 ± 2	ND	Juergens et al., 2018
	TEF1 (*Arxula adeninivorans*)	iCas9	*TDH3* (*S. cerevisiae*)	1	Yes	Episomal	*ADE2* and *YNR1*	2-5	ND	
CBS4732	*TEF1* (*Arxula adeninivorans*)	iCas9	*TDH3* (*S. cerevisiae*)	1	Yes	Episomal	*ADE2*	9 ± 1	ND	
Laboratory strain CGMCC7.89	*TEF1* (*S. cerevisiae*)	iCas9	*SNR52* (*S. cerevisiae*)	1	No	Integrative	*LEU2*	58.33 ± 7.22	Yes (1,000/1,000)bp	Wang et al., 2018
	TEF1 (*S. cerevisiae*)	iCas9	*SNR52* (*S. cerevisiae*)	1	No	Integrative	*URA3*	65.28 ± 2.41	Yes (1,000/1,000)bp	
	TEF1 (*S. cerevisiae*)	iCas9	*SNR52* (*S. cerevisiae*)	1	No	Integrative	*ADE2*	62.18	Yes (1,000/1,000)bp	
	TEF1 (*S. cerevisiae*)	iCas9	*SNR52* (*S. cerevisiae*)	1	No	Integrative	*ADE2*	37.18	Yes (500/500)bp	
	TEF1 (*S. cerevisiae*)	iCas9	*SNR52* (*S. cerevisiae*)	1	No	Integrative	*URA3, LEU2* and *HIS3*	23.61	Yes (1,000/1,000)bp	
	MOX (*H. polymorpha*)	iCas9	*SNR52* (*S. cerevisiae*)	1	No	Integrative	*rDNA*	75	Yes (1,000/1,000)bp	

ND No difference was observed in deletion efficiency with or without HR donor.*

simultaneously for gene disruption. The gRNAs were placed in a tandem array spaced by a 20-bp linker. Although the deletion frequencies reached values between only 2 and 5%, this was the first report of multiple genes editing in *H. polymorpha* through CRISPR/Cas9 (Juergens et al., 2018).

Recently, an efficiently multiplex system was employed using a laboratory strain CGMCC7.89 (from the China General Microbiological Culture Collection Center) as host (**Table 3**). The plasmids carrying the Cas9 and gRNA were inserted into the genome to ensure their stability through up- and downstream homologous arms (∼1.5 kb). The Cas9 was integrated into the *MET2* locus while and the gRNA into the *ADE2* locus. The positive clones were selected by the respective auxotrophies. All transformations were performed with the presence of donor DNA consisting of a fragment containing the up- and downstream region of the target gene amplified from the *H. polymorpha* genome. Initially, the *LEU2* and *URA3* were targeted to test the CRISPR/Cas9 system. Successful deletion occurred with 58.33 ± 7.22% for *LEU2* and 65.28 ± 2.41% for *URA3*. The effect of the size of the homology arms was tested using the *ADE2* locus flanking regions from 50 to 1,000 bp. When 50 bp was utilized, the efficiency was 0%, for 500 and 1,000 bp the values obtained were 37.18 and 62.18%, respectively. Therefore, all editing templates used for subsequent experiments had 1,000/1,000 bp of homology. These results corroborate the hypothesis that for *H. polymorpha* at least 500 bp are required to ensure deletion efficiency. The competence for multiple simultaneous deletions was verified with an efficiency of 23.61% for three loci *URA3*, *LEU2,* and *HIS3*. Next, a template containing a *gfpmut3a* expression cassette flanked by up- and downstream homologous arms of the three loci, *URA3*, *LEU2,* and *HIS3*, were used separately as DNA donor to prove that gene insertion was feasible using CRISPR/Cas9. The efficiency of integration observed reached 66.7%, 66.7% and 62.50% for *HIS3*, *URA3,* and *LEU2*, respectively. After that, the *gfpmut3a* was replaced by genes encoding the proteins within the resveratrol metabolic pathway that resulted in maximum production of 4.7 mg/L. The effect of gene copy number required for resveratrol synthesis was tested by multiple integrations into rDNA regions using CRISPR/Cas9. In this strategy, the *TEF1* promoter that was driving the expression of the CAS9 encoding gene was replaced by the methanol inducible *MOX*. In the presence of the *gfpmut3a* cassette with homologous arms, its insertion in the rDNA region occurred with a frequency of 75% and 11 copies of *gfpmut3a* were quantified in one of the studied clones. Finally, this approach was used to introduce the three genes that enable resveratrol synthesis in this yeast. The maximum resveratrol production was 97.23 ± 4.84 mg/L in a strain containing nine copies of each gene (Wang et al., 2018).

EXAMPLES OF RECOMBINANT PROTEINS IN *H. POLYMORPHA*

The latest recombinant proteins that utilize *H. polymorpha* as host are summarized in **Table 1** although, up to this moment, none of the listed examples were performed at industrial scale. The first heterologous protein produced in *H. polymorpha* was the surface antigens from hepatitis B virus (Janowicz et al., 1991). Furthermore, *H. polymorpha,* has been extensively used for the production of Virus-like particle (VLP), which are viral proteins that can be used for the development of vaccines (Kumar and Kumar, 2019). Only in the last decade, *H. polymorpha* was used to produce VLPs for the hepatitis B virus (Li et al., 2011; Xu et al., 2014), hepatitis C virus (He et al., 2008), hepatitis E virus (Su et al., 2017), Human papillomavirus (Li et al., 2009; Liu et al., 2015; Bredell et al., 2018), rabies virus (Qian et al., 2013), and rotavirus (Bredell et al., 2016). Recently, a chimeric VLP was development for *H. polymorpha* (Wetzel et al., 2018). In this type of platform, viral proteins are produced heterologously as chimeric proteins containing antigens from different viruses, allowing the simultaneous production of different VLPs. For *H. polymorpha*, the developed chimeric VLPs contained viral proteins from bovine viral diarrhea virus, the classical swine fever virus, the feline leukemia virus, and the west Nile virus, using as scaffold the small surface protein of the duck hepatitis B virus (Wetzel et al., 2018).

Further examples of recombinant proteins using *H. polymorpha* as host include the human Parathyroid Hormone (PTH) and Staphylokinase (SAK). Here it is summarized how the recombinant strain was constructed, the process was optimized and scaled-up to 80 liters (Moussa et al., 2012; Mueller et al., 2013). The PTH is a hormone secreted by the parathyroid gland responsible for calcium homeostasis in the blood and osteogenesis. It is a glycoprotein, but it has already been shown that non-glycosylation does not affect its biological activity (Bisello et al., 1996). The N-terminal 1–34 fragment of this protein has an essential function as a receptor binding region and is employed for the treatment of osteoporosis (Cho et al., 2018). The integrative plasmid pFPMT-MFα containing the coding region for the 1–34 fragment of PTH fused with MFα signal was inserted into Δ*yps7* strain (deficient protease YPS7). The *FMD* promoter controlled the production of PTH, and the plasmid carried a *HARS* sequence for genomic integration (Kim et al., 2003). Positive clones for PTH production were selected for medium optimization on a micro-plate scale.

Using micro-scale assays, it has been possible to evaluate the influence of culture media on the PTH titers. At this stage, all standard complex media (YPD, YPG, YNB) in addition to two synthetic media Syn 6 and Syn 6-cp (supplemented with citrate and peptone) were tested. The best PTH concentration (∼40 mg/L) was achieved with the Syn6-cp medium. After that, the influence of three peptones (soy, wheat, and potato) in a 100 mL shake scale was investigated with the highest production of the recombinant protein (9.6 mg/L) was achieved using wheat peptone. Next step was performed in 250 mL baffled flasks with 100 mL of working volume in fed-batch mode. Manual applications after 20, 24, 42, 46, and 50 h of inducing-mixture containing a methanol/glycerol (1:1) and wheat peptone resulted in 25.4 mg/L of PTH. In the following steps, 300-mL bioreactors were utilized to test these parameters under controlled fermentation conditions. Cell cultures were pre-grown in glycerol 3% for 24 h. After that, the induction phase was performed with methanol pulsed every 6 h during 24 h leading

to the production of 68.3 mg/L of PTH. A similar result was obtained in 2 L reactors (67 mg/L) using the same strategy. Finally, 8 L and 80 L fermentation were performed using the optimum conditions. For that, the partial oxygen pressure (pO2) was set to 30%. The feeding strategy was changed from pulse-wise to feeding scheduled realized at a constant rate ranging from 20 to 60 g/L/h for 52 h. Using this fermentation strategy was obtained 120 mg/L and 150 mg/L PTH in 8 and 80 L, respectively (Mueller et al., 2013).

A similar approach has been used for the heterologous production of Staphylokinase (SAK), a biopharmaceutical with pro-fibrinolytic activity (Moussa et al., 2012). Recombinant strains producing an SAK THR164 variant (ThromboGenics NV) were developed and the process was scaled-up from micro-titer plates to 80 L. Two isoforms encoding the SAK protein, rSAK-1 and rSAK-2, were placed under control of the *FMD* promoter, in addition to containing the MFα1 sequence for secretion. The rSAK-2 had a substitution of the Thr-30 amino acid residue for an alanine. This is a recognition site for N-glycosylation and its mutation showed to affect the glycosylation pattern of the produced protein. The final plasmids bearing *URA3* auxotrophic marker and containing the genes encoding rSAK-1 and rSAK-2 were transformed into the strain RB11, a uracil-auxotrophic variant of CBS4732. After the selection, the positives clones were tested for rSAK-1 and rSAK-2 production in shake flasks using YPG (glycerol). The Western blot assay demonstrated that r-SAK1 and not r-SAK-2 was glycosylated. Taking into account that glycosylated SAK shows reduction in enzymatic activity, only strains producing rSAK-2 were further utilized for process optimization.

The micro-plate assays were employed to evaluate the pH, medium and feeding strategy effects on rSAK-2 production using SYN6 medium as basis. The screening of pHs ranging 4–8, twenty different peptones and two feeding strategies indicated that the highest production of rSAK-2 (90 mg/L) occurred in pH 6.5 with SYN6 supplemented with wheat peptone and induced by methanol. Further medium optimization was realized in 500 mL shake flasks with a working volume of 100 mL. At this stage, variations of synthetic medium SYN6 were tested. The removal of trace elements and vitamins of SYN6 yielded 200 mg/L of rSAK-2 after 48 h of fermentation, the best production achieved in shake flasks. This medium was designed as SYN6.46d. After that, the process was transferred to 300 mL bioreactors where two feeding solutions were evaluated. The first one, named as FEED_1, contained 20% yeast extract (w/v), 10% peptone (w/v), 5% glycerol (w/v), and 10% methanol (w/v). While FEED_2 in turn was composed by 20% peptone (w/v), 10% glycerol (w/v), 20% yeast extract (w/v) and 10% methanol (w/v). In constant feeding mode (4 mL of feeding solution per hour) using FEED_2, the yield of rSAK-2 was 423 mg/L at 48 h of fermentation, the highest amount reached so far. Finally two parameters were adjusted before scaling-up the process from 300 mL bioreactors. The air flow was setting to 2 L/min instead 0.5 L/min and the stirring speed was changed from 500 rpm to 800 rpm. Lastly, fermentations using the medium SYN6.46d in a constant fed-batch mode with the feeding solution FEED_2 were performed at bioreactors of 2, 8, and 80 L that yielded rSAK-2 amounts of 1,212, 1,081, and 1,109 mg/L, respectively (Moussa et al., 2012).

Although *H. polymorpha* is more frequently used as host for recombinant protein, the production of chemicals and fatty acids has also been described. The most extensive example discussed in the literature is the utilization of *H. polymorpha* to produce ethanol using xylose as carbon source. Up to now the best reported strain achieved 12.5 g/L using a strain where *CAT8* gene was disrupted. This gene encoding a transcriptional activator involved in the regulation of xylose metabolism in *H. polymorpha* (Ruchala et al., 2017). For an extensive review on this particular bioprocess the reader is directed to a recently published review (Dmytruk et al., 2017). Production of γ-linolenic acid was achieved in *H. polymorpha* by the introduction of *Mucor rouxii* Δ^6-desaturase gene under control of the *MOX* promoter (Khongto et al., 2010). Despite the utilization of a methanol-inducible promoter, the maximum γ-linolenic acid titers, 697 mg/L, was reached when glycerol was used as the carbon source. For the production of 1,3-propandiol, all six genes necessary for its biosynthesis were transferred from *Klebsiella pneumoniae* (Hong et al., 2011). All inserted genes were present in one single plasmid under the control of the *GAP* promoter. The resulting strains produced 2.4 g/L and 0.8 g/L of 1,3-Propandiol using glucose and glycerol as substrate, respectively (Hong et al., 2011). Similarly, the insertion of four genes were introduced in *H. polymorpha* enabling the synthesis of 5-hydroxyectoine (Eilert et al., 2013). Nevertheless, in this study all genes were cloned in different plasmids, where each plasmid had upstream- and downstream regions amplified from the yeast genome to direct the integration and a unique selection marker. The resulting strain was able to achieve 2.8 g/L of 5-hydroxyectoine using methanol as the carbon source. Recently, CRISPR/Cas9 was utilized to introduce three genes (*TAL*, *4CL*, and *STS*) required for the resveratrol synthesis in *H. polymorpha* (Wang et al., 2018). All genes were cloned into one expression cassette and integrated into the *rDNA* locus aiming at multiple copy integrations. The best producing strain contained 9 copies of each gene and was able to produce approximately 98 mg/L (Wang et al., 2018).

CONCLUSIONS

The development of new cell factories for the production of heterologous proteins that are scarce is the primary challenge of the 21st century. Advances in molecular genetics and cultivation techniques drive the growing number of new expression platforms. Among these, the yeast *H. polymorpha* stand out as host due to the presence of strongly methanol-inducible such as *MOX* and *FMD* promoters, the glycosylation pattern compatible with human glycoproteins, the thermotolerance capacity and it is able to use different carbon sources (**Figures 1A–E**). Furthermore, various genetic engineering tools and transformation protocols are established in this yeast. Nevertheless, the low frequency of homologous recombination in *H. polymorpha* delays the strain construction step. Many efforts were made to bypass this problem: utilization of *ku80*

strain, methods for construction of deletion cassettes and implementation of CRISPR/Cas9 technology.

Genome editing via CRISPR/Cas9 represents a powerful tool for genetic manipulation. It is possible to perform not only the deletion of endogenous genes from the organism but also the insertion of exogenous sequences into its genome. The three systems implemented in *H. polymorpha* allowed disruption or the introduction of exogenous genes. Furthermore, the utilization of episomal plasmids for CRISPR/Cas9 implementation in *H. polymorpha* required modifications into the initial strategy to enhance the deletion frequencies (**Table 3**). In both cases, adjustments to the developed approach increased the efficiency of the system: substitution of promoters (Numamoto et al., 2017) or prolonged incubation times to guarantee the activity of Cas9 (Juergens et al., 2018). Implementation of CRISPR/Cas9 through integrative plasmids guaranteed gene deletion rates >50% for different loci (Wang et al., 2018). Also, co-transformation with a DNA template to induce HR repair after slicing of Cas9 was more efficient than using only Cas9 and gRNA (**Table 3**). The CRISPR/Cas9 system was also efficient for the introduction of exogenous genes into *H. polymorpha* (Wang et al., 2018). It was possible to introduce a complete pathway for the synthesis of resveratrol in *H. polymorpha* in a single transformation event, representing a revolution in the genetic manipulation of this yeast. Thus, it is evident how the utilization of genome editing tools will reduce the time and cost of strain construction allowing a rapid introduction into the commercial scale.

Finally, the examples given for recombinant protein production can be extrapolated to the production of other molecules. Micro-scale cultivation is often used to optimize culture conditions by testing different mediums composition, pH and other factors that may affect the productivity of the desired protein. The best results are scaled up adding the feeding and oxygenation strategies. Altogether these features elucidated how *H. polymorpha* is a promising host for the establishment of various bioprocesses. This is reflected already by the number of products available in the market and by the pipeline of those that are in the optimization phase.

AUTHOR CONTRIBUTIONS

JM-N wrote the manuscript. AG prepared the figure. NP assisted with writing, editing, and finalizing the manuscript. All authors approved its publication.

ACKNOWLEDGMENTS

This work was supported by the Brazilian funding agencies of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) and FAP-DF.

REFERENCES

Abdel-Banat, B. M. A., Hoshida, H., Ano, A., Nonklang, S., and Akada, R. (2010). High-temperature fermentation: how can processes for ethanol production at high temperatures become superior to the traditional process using mesophilic yeast? *Appl. Microbiol. Biotechnol.* 85, 861–867. doi: 10.1007/s00253-009-2248-5

Agaphonov, M., and Alexandrov, A. (2014). Self-excising integrative yeast plasmid vectors containing an intronated recombinase gene. *FEMS Yeast Res.* 14, 1048–1054. doi: 10.1111/1567-1364.12197

Agaphonov, M. O., Trushkina, P. M., Sohn, J. H., Choi, E. S., Rhee, S. K., and Ter-Avanesyan, M. D. (1999). Vectors for rapid selection of integrants with different plasmid copy numbers in the yeast *Hansenula polymorpha* DL1. *Yeast* 15, 541–551. doi: 10.1002/(SICI)1097-0061(199905)15:7<541::AID-YEA392>3.0.CO;2-G

Akbarzadeh, A., Siadat, S. O. R., Zamani, M. R., Motallebi, M., and Barshan, T. M. (2013). Comparison of biochemical properties of recombinant endoglucanase II of *Trichoderma reesei* in methylotrophic yeasts, *Pichia pastoris* and *Hansenula polymorpha*. *Prog. Biol. Sci.* 3, 108–117. doi: 10.22059/PBS.2013.32100

Amuel, C., Gellissen, G., Hollenberg, C. P., and Suckow, M. (2000). Analysis of heat shock promoters in *Hansenula polymorpha*: the *TPS1* promoter, a novel element for heterologous gene expression. *Biotechnol. Bioprocess Eng.* 5, 247–252. doi: 10.1007/BF02942181

Ballou, C. E. (1990). Isolation, characterization, and properties of *Saccharomyces cerevisiae mnn* mutants with nonconditional protein glycosylation defects. *Methods Enzymol.* 185, 440–470. doi: 10.1016/0076-6879(90)85038-P

Bao, Z., Xiao, H., Liang, J., Zhang, L., Xiong, X., Sun, N., et al. (2015). Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in *Saccharomyces cerevisiae*. *ACS Synth. Biol.* 4, 585–594. doi: 10.1021/sb500255k

Bisello, A., Greenberg, Z., Behar, V., Rosenblatt, M., Suva, L. J., and Chorev, M. (1996). Role of glycosylation in expression and function of the human parathyroid hormone/parathyroid hormone-related protein receptor. *Biochemistry* 35, 15890–15895. doi: 10.1021/bi962111+

Böer, E., Steinborn, G., and Kunze, G. (2007). Yeast expression platforms. *Appl. Microbiol. Biotechnol.* 77, 513–523. doi: 10.1007/s00253-007-1209-0

Bogdanova, A. I., Agaphonov, M., and Ter-avanesyan, M. D. (1995). Plasmid reorganization during integrative transformation in *Hansenula polymorpha*. *Yeast* 11, 343–353.

Bogdanova, A. I., Agaphonov, M. O., and Ter-Avanesian, M. D. (2000). Plasmid instability in methylotrophic yeast *Hansenula polymorpha*: the capture of chromosomal DNA fragments by replicative plasmids. *Mol. Biol.* 34, 36–40.

Borrero, J., Kunze, G., Jiménez, J. J., Böer, E., Gútiez, L., Herranz, C., et al. (2012). Cloning, production, and functional expression of the bacteriocin enterocin A, produced by *Enterococcus faecium* T136, by the yeasts *Pichia pastoris*, *Kluyveromyces lactis*, *Hansenula polymorpha*, and *Arxula adeninivorans*. *Appl. Environ. Microbiol.* 78, 5956–5961. doi: 10.1128/AEM.00530-12

Bredell, H., Smith, J. J., Görgens, J. F., and Zyl, W. H. (2018). Expression of unique chimeric human papilloma virus type 16 (HPV-16) L1-L2 proteins in *Pichia pastoris* and *Hansenula polymorpha*. *Yeast* 35, 519–529. doi: 10.1002/yea.3318

Bredell, H., Smith, J. J., Prins, W. A., Görgens, J. F., and van Zyl, W. H. (2016). Expression of rotavirus VP6 protein: a comparison amongst *Escherichia coli*, *Pichia pastoris* and *Hansenula polymorpha*. *FEMS Yeast Res.* 16:fow001. doi: 10.1093/femsyr/fow001

Chen, Z., Wang, Z., He, X., Guo, X., Li, W., and Zhang, B. (2008). Uricase production by a recombinant *Hansenula polymorpha* strain harboring *Candida utilis* uricase gene. *Appl. Microbiol. Biotechnol.* 79, 545–554. doi: 10.1007/s00253-008-1472-8

Cheon, S. A., Kim, H., Oh, D. B., Kwon, O., and Kang, H. A. (2012). Remodeling of the glycosylation pathway in the methylotrophic yeast *Hansenula polymorpha* to produce human hybrid-type N-glycans. *J. Microbiol.* 50, 341–348. doi: 10.1007/s12275-012-2097-2

Cho, M., Han, S., Kim, H., Kim, K. S., and Hahn, S. K. (2018). Hyaluronate–parathyroid hormone peptide conjugate for transdermal treatment of osteoporosis. *J. Biomater. Sci. Polym. Ed.* 29, 793–804. doi: 10.1080/09205063.2017.1399001

Dmytruk, K., Kurylenko, O., Ruchala, J., Ishchuk, O., and Sibirny, A. (2017). "Development of the thermotolerant methylotrophic yeast *Hansenula*

polymorpha as efficient ethanol producer", in *Yeast Diversity in Human Welfare*, eds T. Satyanarayana and G. Kunze (Singapore: Springer), 257–282.

Dmytruk, K. V., Smutok, O. V., Dmytruk, O. V., Schuhmann, W., and Sibirny, A. A. (2011). Construction of uricase-overproducing strains of *Hansenula polymorpha* and its application as biological recognition element in microbial urate biosensor. *BMC Biotechnol.* 11:58. doi: 10.1186/1472-6750-11-58

Donohoue, P. D., Barrangou, R., and May, A. P. (2018). Advances in industrial biotechnology using CRISPR-Cas systems. *Trends Biotechnol.* 36, 134–146. doi: 10.1016/j.tibtech.2017.07.007

Eilert, E., Hollenberg, C. P., Piontek, M., and Suckow, M. (2012). The use of highly expressed FTH1 as carrier protein for cytosolic targeting in *Hansenula polymorpha*. *J. Biotechnol.* 159, 172–176. doi: 10.1016/j.jbiotec.2011.12.014

Eilert, E., Kranz, A., Hollenberg, C. P., Piontek, M., and Suckow, M. (2013). Synthesis and release of the bacterial compatible solute 5-hydroxyectoine in *Hansenula polymorpha*. *J. Biotechnol.* 167, 85–93. doi: 10.1016/j.jbiotec.2013.02.005

Faber, K. N., Haima, P., Harder, W., Veenhuis, M., and Ab, G. (1994). Highly-efficient electrotransformation of the yeast *Hansenula polymorpha*. *Curr. Genet.* 25, 305–310. doi: 10.1007/BF00351482

Filyak, Y., Finiuk, N., Mitina, N., Bilyk, O., Titorenko, V., Hrydzhuk, O., et al. (2013). A novel method for genetic transformation of yeast cells using oligoelectrolyte polymeric nanoscale carriers. *Biotechniques* 54, 35–43. doi: 10.2144/000113980

Gellissen, G., Janowicsz, Z. A., Weydemann, U., Melber, K., Strasser, A. W. M., and Hollenberg, C. P. (1992). High-level expression of foreign genes in *Hansenula polymorpha*. *Biotechnol. Adv.* 10, 179–189.

Gellissen, G., Kunze, G., Gaillardin, C., Cregg, J. M., Berardi, E., Veenhuis, M., et al. (2005). New yeast expression platforms based on methylotrophic *Hansenula polymorpha* and *Pichia pastoris* and on dimorphic *Arxula adeninivorans* and *Yarrowia lipolytica*—a comparison. *FEMS Yeast Res.* 5, 1079–1096. doi: 10.1016/j.femsyr.2005.06.004

Gidijala, L., Kiel, J. A. K. W., Douma, R. D., Seifar, R. M., van Gulik, W. M., Bovenberg, R. A. L., et al. (2009). An engineered yeast efficiently secreting penicillin. *PLoS ONE* 4:e8317. doi: 10.1371/journal.pone.0008317

Gonzalez, C. (1999). One-step, PCR-mediated, gene disruption in the yeast *Hansenula polymorpha*. *Yeast* 15, 1323–1329.

Hartner, F. S., and Glieder, A. (2006). Regulation of methanol utilisation pathway genes in yeasts. *Microb. Cell Fact.* 5, 1–21. doi: 10.1186/1475-2859-5-39

He, F., Joshi, S. B., Bosman, F., Verhaeghe, M., and Middaugh, C. R. (2008). Structural stability of hepatitis C virus envelope glycoprotein E1: effect of pH and dissociative detergents. *J. Pharm. Sci.* 98, 3340–3357. doi: 10.1002/jps.21657

Heo, J. H., Hong, W. K., Cho, E. Y., Kim, M. W., Kim, J. Y., Kim, C. H., et al. (2003). Properties of the *Hansenula polymorpha*-derived constitutive *GAP* promoter, assessed using an HSA reporter gene. *FEMS Yeast Res.* 4, 175–184. doi: 10.1016/S1567-1356(03)00150-8

Hollenberg, C. P., and Gellissen, G. (1997). Production of recombinant proteins by methylotrophic yeasts. *Curr. Opin. Biotechnol.* 8, 554–560. doi: 10.1016/S0958-1669(97)80028-6

Hong, W. K., Kim, C. H., Heo, S. Y., Luo, L. H., Oh, B. R., Rairakhwada, D., et al. (2011). 1,3-Propandiol production by engineered *Hansenula polymorpha* expressing *dha* genes from *Klebsiella pneumoniae*. *Bioprocess Biosyst. Eng.* 34, 231–236. doi: 10.1007/s00449-010-0465-z

Janowicz, Z. A., Melber, K., Merckelbach, A., Jacobs, E., Harford, N., Comberbach, M., et al. (1991). Simultaneous expression of the S and L surface antigens of hepatitis B, and formation of mixed particles in the methylotrophic yeast, *Hansenula polymorpha*. *Yeast* 7, 431–443. doi: 10.1002/yea.320070502

Jenzelewski, V. (2002). "Fermentation and primary product recovery," in *Hansenula polymorpha—Biology and Applications*, ed G. Gellissen (Wiley-VCH Verlag, Weinheim), 156–174.

Juergens, H., Varela, J. A., Gorter de Vries, A. R., Perli, T., Gast, V. J. M., Gyurchev, N. Y., et al. (2018). Genome editing in *Kluyveromyces* and *Ogataea* yeasts using a broad-host-range Cas9/gRNA co-expression plasmid. *FEMS Yeast Res.* 18:foy012. doi: 10.1093/femsyr/foy012

Khongto, B., Laoteng, K., and Tongta, A. (2010). Fermentation process development of recombinant *Hansenula polymorpha* for gamma-linolenic acid production. *J. Microbiol. Biotechnol.* 20, 1555–1562. doi: 10.4014/jmb.1003.03004

Kim, H., Thak, E. J., Lee, D., Agaphonov, M. O., and Kang, H. A. (2015b). *Hansenula polymorpha* Pmt4p plays critical roles in O-mannosylation of surface membrane proteins and participates in heteromeric complex formation. *PLoS ONE* 10:e0129914. doi: 10.1371/journal.pone.0129914

Kim, H., Yoo, S. J., and Kang, H. A. (2015a). Yeast synthetic biology for the production of recombinant therapeutic proteins. *FEMS Yeast Res.* 15, 1–16. doi: 10.1111/1567-1364.12195

Kim, M. W., Kim, E. J., Kim, J.-Y., Park, J.-S., Oh, D.-B., Shimma, Y., et al. (2006). Functional characterization of the *Hansenula polymorpha* HOC1, OCH1, and OCR1 genes as members of the yeast OCH1 mannosyltransferase family involved in protein glycosylation. *J. Biol. Chem.* 281, 6261–6272. doi: 10.1074/jbc.M508507200

Kim, M. W., Rhee, S. K., Kim, J. Y., Shimma, Y. I., Chiba, Y., Jigami, Y., et al. (2004). Characterization of N-linked oligosaccharides assembled on secretory recombinant glucose oxidase and cell wall mannoproteins from the methylotrophic yeast *Hansenula polymorpha*. *Glycobiology* 14, 243–251. doi: 10.1093/glycob/cwh030

Kim, S.-Y., Sohn, J.-H., Bae, J.-H., Pyun, Y.-R., Agaphonov, M. O., Ter-Avanesyan, M. D., et al. (2003). Efficient library construction by *in vivo* recombination with a telomere-originated autonomously replicating sequence of *Hansenula polymorpha*. *Appl. Environ. Microbiol.* 69, 4448–4454. doi: 10.1128/aem.69.8.4448-4454.2003

Krasovska, O. S., Stasyk, O. G., Nahorny, V. O., Stasyk, O. V., Granovski, N., Kordium, V. A., et al. (2007). Glucose-induced production of recombinant proteins in *Hansenula polymorpha* mutants deficient in catabolite repression. *Biotechnol. Bioeng.* 97, 858–870. doi: 10.1002/bit.21284

Kumar, R., and Kumar, P. (2019). Yeast-based vaccines: new perspective in vaccine development and application. *FEMS Yeast Res.* 19:foz007. doi: 10.1093/femsyr/foz007

Kumari, A., Baronian, K., Kunze, G., and Gupta, R. (2015). Extracellular expression of YlLip11 with a native signal peptide from *Yarrowia lipolytica* MSR80 in three different yeast hosts. *Protein Expr. Purif.* 110, 138–144. doi: 10.1016/j.pep.2015.02.016

Kurtzman, C. P., and Robnett, C. J. (2010). Systematics of methanol assimilating yeasts and neighboring taxa from multigene sequence analysis and the proposal of *Peterozyma* gen. nov., a new member of the Saccharomycetales. *FEMS Yeast Res.* 10, 353–361. doi: 10.1111/j.1567-1364.2010.00625.x

Kurylenko, O. O., Ruchala, J., Hryniv, O. B., Abbas, C. A., Dmytruk, K. V., and Sibirny, A. A. (2014). Metabolic engineering and classical selection of the methylotrophic thermotolerant yeast *Hansenula polymorpha* for improvement of high-temperature xylose alcoholic fermentation. *Microb. Cell Fact.* 13:122. doi: 10.1186/s12934-014-0122-3

Kurylenko, O. O., Ruchala, J., Vasylyshyn, R. V., Stasyk, O. V., Dmytruk, O. V., Dmytruk, K. V., et al. (2018). Peroxisomes and peroxisomal transketolase and transaldolase enzymes are essential for xylose alcoholic fermentation by the methylotrophic thermotolerant yeast, *Ogataea* (*Hansenula*) *polymorpha*. *Biotechnol. Biofuels* 11, 1–16. doi: 10.1186/s13068-018-1203-z

Levine, D. W., and Cooney, C. L. (1973). Isolation and characterization of a thermotolerant methanol-utilizing yeast. *Appl. Microbiol.* 26, 982–989.

Li, W., He, X., Guo, X., Zhang, Z., Zhang, B. (2009). Optimized expression of the L1 protein of human papilloma virus in *Hansenula polymorpha*. *Chin. J. Biotech.* 25, 1516–1523.

Li, Y., Song, H., Li, J., Wang, Y., Yan, X., Zhao, B., et al. (2011). *Hansenula polymorpha* expressed heat shock protein gp96 exerts potent T cell activation activity as an adjuvant. *J. Biotechnol.* 151, 343–349. doi: 10.1016/j.jbiotec.2010.12.006

Liu, C., Yao, Y., Yang, X., Bai, H., Huang, W., Xia, Y., et al. (2015). Production of recombinant human papillomavirus type 52 L1 protein in *Hansenula polymorpha* formed virus-like particles. *J. Microbiol. Biotechnol.* 25, 936–940. doi: 10.4014/jmb.1412.12027

Manivasakam, P., Weber, S. C., Mcelver, J., and Schiestl, R. H. (1995). Micro-homology mediated PCR targeting in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 23, 2799–2800.

Mayer, A. F., Hellmuth, K., Schlieker, H., Lopez-Ulibarri, R., Oertel, S., Dahlems, U., et al. (1999). An expression system matures: a highly efficient and cost-effective process for phytase production by recombinant strains of *Hansenula polymorpha*. *Biotechnol. Bioeng.* 63, 373–381.

Morais, J. O. F., and Maia, M. H. D. (1959). Estudos de microorganismos encontrados em leitos de despejos de caldas de destilarias de Pernambuco. II. Uma nova especie de *Hansenula*: *H. polymorpha*. *Anais de Escola Superior de Quimica de Universidade do Recife* 1, 15–20.

Moussa, M., Ibrahim, M., El Ghazaly, M., Rohde, J., Gnoth, S., Anton, A., et al. (2012). Expression of recombinant staphylokinase in the methylotrophic yeast *Hansenula polymorpha*. *BMC Biotechnol.* 12:96. doi: 10.1186/1472-6750-12-96

Mueller, F., Moussa, M., El Ghazaly, M., Rohde, J., Bartsch, N., Parthier, A., et al. (2013). Efficient production of recombinant parathyroid hormone (rPTH) fragment 1-34 in the methylotrophic yeast *Hansenula polymorpha*. *GaBI J.* 2, 114–122. doi: 10.5639/gabij.2013.0203.035

Numamoto, M., Maekawa, H., and Kaneko, Y. (2017). Efficient genome editing by CRISPR/Cas9 with a tRNA-sgRNA fusion in the methylotrophic yeast *Ogataea polymorpha*. *J. Biosci. Bioeng.* 124, 487–492. doi: 10.1016/j.jbiosc.2017.06.001

Oh, D. B., Park, J. S., Kim, M. W., Cheon, S. A., Kim, E. J., Moon, H. Y., et al. (2008). Glycoengineering of the methylotrophic yeast *Hansenula polymorpha* for the production of glycoproteins with trimannosyl core N-glycan by blocking core oligosaccharide assembly. *Biotechnol. J.* 3, 659–668. doi: 10.1002/biot.200700252

Qian, W., Aguilar, F., Wang, T., and Qiu, B. (2013). Secretion of truncated recombinant rabies virus glycoprotein with preserved antigenic properties using a co-expression system in *Hansenula polymorpha*. *J. Microbiol.* 51, 234–240. doi: 10.1007/s12275-013-2337-0

Qian, W., Song, H., Liu, Y., Zhang, C., Niu, Z., Wang, H., et al. (2009). Improved gene disruption method and Cre- loxP mutant system for multiple gene disruptions in *Hansenula polymorpha*. *J. Microbiol. Methods* 79, 253–259. doi: 10.1016/j.mimet.2009.09.004

Ramezani-Rad, M., Hollenberg, C. P., Lauber, J., Wedler, H., Griess, E., Wagner, C., et al. (2003). The *Hansenula polymorpha* (strain CBS4732) genome sequencing and analysis. *FEMS Yeast Res.* 4, 207–215. doi: 10.1016/S1567-1356(03)00125-9

Raschke, W. C., Neiditch, B. R., Hendricks, M., and Cregg, J. M. (1996). Inducible expression of a heterologous protein in *Hansenula polymorpha* using the alcohol oxidase 1 promoter of *Pichia pastoris*. *Gene* 177, 163–167. doi: 10.1016/0378-1119(96)00293-4

Ravin, N. V., Eldarov, M. A., Kadnikov, V. V., Beletsky, A. V., Schneider, J., Mardanova, E. S., et al. (2013). Genome sequence and analysis of methylotrophic yeast *Hansenula polymorpha* DL1. *BMC Genomics* 14:837. doi: 10.1186/1471-2164-14-837

Riley, R., Haridas, S., Wolfe, K. H., Lopes, M. R., Hittinger, C. T., Göker, M., et al. (2016). Comparative genomics of biotechnologically important yeasts. *Proc. Natl. Acad. Sci. U.S.A.* 113, 9882–9887. doi: 10.1073/pnas.1603941113

Ruchala, J., Kurylenko, O. O., Soontorngun, N., Dmytruk, K. V., and Sibirny, A. A. (2017). Transcriptional activator *Cat8* is involved in regulation of xylose alcoholic fermentation in the thermotolerant yeast *Ogataea* (*Hansenula*) *polymorpha*. *Microb. Cell Fact.* 16, 1–13. doi: 10.1186/s12934-017-0652-6

Ryabova, O. B., Chmil, O. M., and Sibirny, A. A. (2003). Xylose and cellobiose fermentation to ethanol by the thermotolerant methylotrophic yeast *Hansenula polymorpha*. *FEMS Yeast Res.* 4, 157–164. doi: 10.1016/S1567-1356(03)00146-6

Saraya, R., Krikken, A. M., Kiel, J. A. K. W., Baerends, R. J. S., Veenhuis, M., and van der Klei, I. J. (2012). Novel genetic tools for *Hansenula polymorpha*. *FEMS Yeast Res.* 12, 271–278. doi: 10.1111/j.1567-1364.2011.00772.x

Sigawi, S., Smutok, O., Demkiv, O., Gayda, G., Nitzan, Y., Vus, B., et al. (2014). Detection of waterborne and airborne formaldehyde: from amperometric chemosensing to a visual biosensor based on alcohol oxidase. *Materials* 7, 1055–1068. doi: 10.3390/ma7021055

Smutok, O., Dmytruk, K., Gonchar, M., Sibirny, A., and Schuhmann, W. (2007). Permeabilized cells of flavocytochrome b2 over-producing recombinant yeast *Hansenula polymorpha* as biological recognition element in amperometric lactate biosensors. *Biosens. Bioelectron.* 23, 599–605. doi: 10.1016/j.bios.2007.06.021

Smutok, O., Karkovska, M., Prokopiv, T., Kavetskyy, T., Sibirnyj, W., and Gonchar, M. (2018). D-lactate-selective amperometric biosensor based on the mitochondrial fraction of *Ogataea polymorpha* recombinant cells. *Yeast*. doi: 10.1002/yea.3372

Sohn, J., Choi, E., Kim, C., and Agaphonov, M. O. (1996). A Novel autonomously replicating sequence (ARS) for multiple integration in the yeast *Hansenula polymorpha* DL-1. *J. Bacteriol.* 178, 4420–4428.

Sohn, J. H., Choi, E. S., Kang, H. A., Rhee, J. S., Agaphonov, M. O., Ter-Avanesyan, M. D., et al. (1999). A dominant selection system designed for copy-number-controlled gene integration in *Hansenula polymorpha* DL-1. *Appl. Microbiol. Biotechnol.* 51, 800–807. doi: 10.1007/s002530051465

Stasyk, O. V., Stasyk, O. G., Komduur, J., Veenhuis, M., Cregg, J. M., and Sibirny, A. A. (2004). A Hexose Transporter Homologue Controls Glucose Repression in the Methylotrophic Yeast *Hansenula polymorpha*. *J. Biol. Chem.* 279, 8116–8125. doi: 10.1074/jbc.M310960200

Stöckmann, C., Scheidle, M., Dittrich, B., Merckelbach, A., Hehmann, G., Melmer, G., et al. (2009). Process development in *Hansenula polymorpha* and *Arxula adeninivorans*, a re-assessment. *Microb. Cell Fact.* 8:22. doi: 10.1186/1475-2859-8-22

Su, C., Li, L., Jin, Z., Han, X., Zhao, P., Wang, L., et al. (2017). Fermentation, purification and immunogenicity evaluation of hepatitis E virus-like particles expressed in *Hansenula polymorpha*. *Sheng Wu Gong Cheng Xue Bao* 33, 653–663. doi: 10.13345/j.cjb.160387

Suh, S. O., and Zhou, J. J. (2010). Methylotrophic yeasts near *Ogataea* (*Hansenula*) *polymorpha*: a proposal of *Ogataea angusta* comb. nov. and *Candida parapolymorpha* sp. nov. *FEMS Yeast Res.* 10, 631–638. doi: 10.1111/j.1567-1364.2010.00634.x

Suppi, S., Michelson, T., and Viigand, K. (2013). Repression vs. activation of *MOX* , *FMD* , *MPP1* and *MAL1* promoters by sugars in *Hansenula polymorpha*: the outcome depends on cell's ability to phosphorylate sugar. *FEMS Yeast Res.* 13, 219–232. doi: 10.1111/1567-1364.12023

Talebkhan, Y., Samadi, T., Samie, A., Barkhordari, F., Azizi, M., Khalaj, V., Mirabzadeh, E. (2016). Expression of granulocyte colony stimulating factor (GCSF) in *Hansenula polymorpha*. *Iran J. Microbiol.* 8, 21–28.

Tikhomirova, L. P., Ikonomova, R. N., Kuznetsova, E. N., and Fodor, I. I. (1988). Transformation of methylotrophic yeast *Hansenula polymorpha*: cloning and expression of genes. *J. Basic Microbiol.* 28, 343–351. doi: 10.1002/jobm.3620280509

Van Dijk, R., Faber, K. N., Kiel, J. A. K. W., Veenhuis, M., and Van Der Klei, I. (2000). The methylotrophic yeast *Hansenula polymorpha*: a versatile cell factory. *Enzyme Microb. Technol.* 26, 793–800. doi: 10.1016/S0141-0229(00)00173-3

van Zutphen, T., Baerends, R. J. S., Susanna, K. A., de Jong, A., Kuipers, O. P., Veenhuis, M., et al. (2010). Adaptation of *Hansenula polymorpha* to methanol: a transcriptome analysis. *BMC Genomics* 11:1. doi: 10.1186/1471-2164-11-1

Vaquero, M. E., Barriuso, J., Medrano, F. J., Prieto, A., and Martínez, M. J. (2015). Heterologous expression of a fungal sterol esterase/lipase in different hosts: effect on solubility, glycosylation and production. *J. Biosci. Bioeng.* 120, 637–643. doi: 10.1016/j.jbiosc.2015.04.005

Vogl, T., and Glieder, A. (2013). Regulation of *Pichia pastoris* promoters and its consequences for protein production. *N. Biotechnol.* 30, 385–404. doi: 10.1016/j.nbt.2012.11.010

Voronovsky, A. Y., Rohulya, O. V., Abbas, C. A., and Sibirny, A. A. (2009). Development of strains of the thermotolerant yeast Hansenula polymorpha capable of alcoholic fermentation of starch and xylan. *Metab. Eng.* 11, 234–242. doi: 10.1016/j.ymben.2009.04.001

Wang, L., Deng, A., Zhang, Y., Liu, S., Liang, Y., Bai, H., et al. (2018). Efficient CRISPR – Cas9 mediated multiplex genome editing in yeasts. *Biotechnol. Biofuels* 11:277. doi: 10.1186/s13068-018-1271-0

Wang, N., Wang, Y. J., Li, G. Q., Sun, N., and Liu, D. H. (2011). Expression, characterization, and antimicrobial ability of T4 lysozyme from methylotrophic yeast *Hansenula polymorpha* A16. *Sci. China Life Sci.* 54, 520–526. doi: 10.1007/s11427-011-4174-x

Wetzel, D., Müller, J. M., Flaschel, E., Friehs, K., and Risse, J. M. (2016). Fed-batch production and secretion of streptavidin by *Hansenula polymorpha*: evaluation of genetic factors and bioprocess development. *J. Biotechnol.* 225, 3–9. doi: 10.1016/j.jbiotec.2016.03.017

Wetzel, D., Rolf, T., Suckow, M., Merz, J., Kranz, A., Barbian, A., et al. (2018). Establishment of a yeast-based VLP platform for antigen presentation. *Microb. Cell Fact.* 17, 1–17. doi: 10.1186/s12934-018-0868-0

Wickerham, L. J. (1951). *Taxonomy of Yeasts*. Technical Bulletin No. 1029. Washington, DC: US Department Agriculture, 1–56.

Xu, X., Ren, S., Chen, X., Ge, J., Xu, Z., Huang, H., et al. (2014). Generation of hepatitis B virus PreS2-S antigen in *Hansenula polymorpha*. *Virol. Sin.* 29, 403–409. doi: 10.1007/s12250-014-3508-9

Youn, J. K., Shang, L., Kim, M. I, Jeong, C. M., Chang, H. N., Hahm, M. S., et al. (2010). Enhanced production of human serum albumin by fed-batch culture of *Hansenula polymorpha* with high-purity oxygen. *J. Microbiol. Biotechnol.* 20, 1534–1538. doi: 10.4014/jmb.0909.09046

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

microorganisms

MDPI

Article

Evaluation of *Ogataea* (*Hansenula*) *polymorpha* for Hyaluronic Acid Production

João Heitor Colombelli Manfrão-Netto [1], Enzo Bento Queiroz [1], Kelly Assis Rodrigues [1], Cintia M. Coelho [2], Hugo Costa Paes [3], Elibio Leopoldo Rech [4] and Nádia Skorupa Parachin [1,5,*]

1 Grupo Engenharia de Biocatalisadores, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília 70910-900, Brazil; jhnettobio@gmail.com (J.H.C.M.-N.); enzo.beno@gmail.com (E.B.Q.); rodrigues.bio23@gmail.com (K.A.R.)
2 Department of Genetics and Morphology, Institute of Biological Science, University of Brasília, Brasília 70910-900, Brazil; cintia.coelhom@gmail.com
3 Clinical Medicine Division, University of Brasília Medical School, University of Brasília, Brasília 70910-900, Brazil; sorumbatico@gmail.com
4 Brazilian Agriculture Research Corporation—Embrapa—Genetic Resources and Biotechnology—CENARGEN, Brasília 70770-917, Brazil; elibio.rech@embrapa.br
5 Ginkgo Bioworks, Boston, MA 02210, USA
* Correspondence: nadiasp@gmail.com

Abstract: Hyaluronic acid (HA) is a biopolymer formed by UDP-glucuronic acid and UDP-N-acetyl-glucosamine disaccharide units linked by β-1,4 and β-1,3 glycosidic bonds. It is widely employed in medical and cosmetic procedures. HA is synthesized by hyaluronan synthase (HAS), which catalyzes the precursors' ligation in the cytosol, elongates the polymer chain, and exports it to the extracellular space. Here, we engineer *Ogataea* (*Hansenula*) *polymorpha* for HA production by inserting the genes encoding UDP-glucose 6-dehydrogenase, for UDP-glucuronic acid production, and HAS. Two microbial HAS, from *Streptococcus zooepidemicus* (*hasAs*) and *Pasteurella multocida* (*hasAp*), were evaluated separately. Additionally, we assessed a genetic switch using integrases in *O. polymorpha* to uncouple HA production from growth. Four strains were constructed containing both *has* genes under the control of different promoters. In the strain containing the genetic switch, HA production was verified by a capsule-like layer around the cells by scanning electron microscopy in the first 24 h of cultivation. For the other strains, the HA was quantified only after 48 h and in an optimized medium, indicating that HA production in *O. polymorpha* is limited by cultivation conditions. Nevertheless, these results provide a proof-of-principle that *O. polymorpha* is a suitable host for HA production.

Keywords: *Ogataea polymorpha* 1; hyaluronic acid 2; methylotrophic yeast 3; genomic editing 4; genetic switch 5; promoters 6

Citation: Manfrão-Netto, J.H.C.; Queiroz, E.B.; Rodrigues, K.A.; Coelho, C.M.; Paes, H.C.; Rech, E.L.; Parachin, N.S. Evaluation of *Ogataea* (*Hansenula*) *polymorpha* for Hyaluronic Acid Production. *Microorganisms* **2021**, *9*, 312. https://doi.org/10.3390/microorganisms9020312

Academic Editor: Benedetta Turchetti
Received: 7 December 2020
Accepted: 29 January 2021
Published: 3 February 2021

1. Introduction

Microorganisms have been widely used to obtain pharmaceuticals, biopolymers, vaccines, enzymes, and various other chemicals [1]. Through metabolic engineering and synthetic biology approaches, it is possible to generate recombinant strains to produce desired compounds and bypass the chemical synthesis [2]. One of the value-added products that can be obtained through microorganism-based processes is hyaluronic acid (HA). This biopolymer features high viscosity and elasticity and is abundant in the extracellular matrix of vertebrates' connective tissues. Due to its high biocompatibility, HA has various applications in the medical, cosmetic, and pharmaceutical areas (see details in [3]). According to Grand View Research Inc., the global HA market may reach USD 16.6 billion by 2027 [4].

HA is a glycosaminoglycan formed by disaccharide repeats of UDP-glucuronic acid (UDP-GlcUA) and UDP-N-acetylglucosamine (UDP-GlcNAc) linked by β-1,4 and β-1,3

glycosidic bonds [5]. Its synthesis is catalyzed by the enzyme hyaluronan synthase (HAS), encoded by the gene *hasA*, and responsible for assembling the two precursors in the cytosol, elongating the polymer chain and releasing it into the extracellular matrix. Gram-positive bacteria and yeasts are the most used microorganisms for producing heterologous HA, especially those with Generally Recognized as Safe (GRAS) status. Many such strains have been developed in recent years, using different approaches to improve HA titers [6]. *Bacillus subtilis* [7,8], *Corynebacterium glutamicum* [9], and *Lactococcus lactis* [10] are the main bacteria used for the production of this polymer. Bacteria possess the complete pathway for the synthesis of both precursors, whereas yeasts are only naturally capable of synthesizing UDP-GlcNAc. Thus, in addition to *hasA*, HA production in yeast cells requires the expression of the *hasB* gene, which encodes the UDP-glucose 6-dehydrogenase needed for UDP-GlcUA synthesis (Figure 1). To the best of our knowledge, only two yeasts have been engineered for HA production so far: *Kluyveromyces lactis* [11] and *Komagataella phaffii* (previously known as *Pichia pastoris*) [12]. Here, we describe the heterologous production of HA by *Ogataea* (*Hansenula*) *polymorpha*.

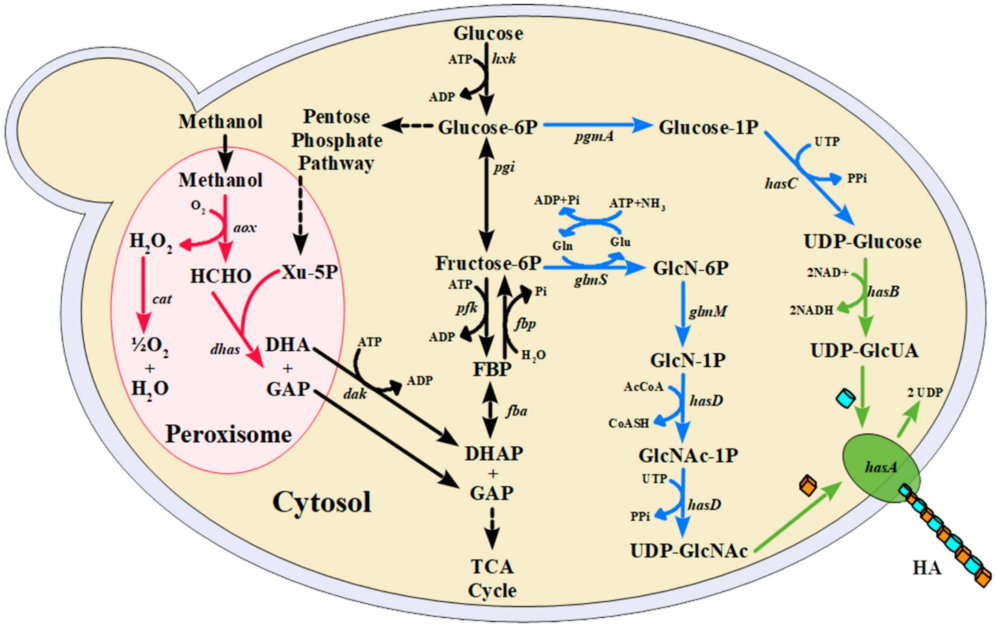


Figure 1. Biosynthetic pathway for hyaluronic acid (HA) production in *Ogataea* (*Hansenula*) *polymorpha*. The black and blue arrows indicate the endogenous glycolysis and HA precursor pathways, respectively. The green arrows indicate the exogenous pathway inserted in *O. polymorpha* for HA production. The red arrows represent methanol metabolism. Genes encoding enzymes of the pathways are shown: *aox*: alcohol oxidase; *cat*: catalase; *dhas*: dihydroxyacetone synthase; *dak*: dihydroxyacetone kinase; *hxk*: hexokinase; *pgi*: glucose-6-phosphate isomerase; *pfk*: phosphofructokinase; *fbp*: fructose-1, 6-bisphosphatase; *fba*: fructose-bisphosphate aldolase; *pgmA*: phosphoglucomutase; *has*C: UDP-glucose pyrophosphorylase; *hasB*: UDP-glucose dehydrogenase; *glmS*: amidotransferase; *glmM*: phosphoglucosamine mutase; *hasD*: acetyltransferase and UDP-N-acetyl-glucosamine pyrophosphorylase; *hasA*: hyaluronic acid synthase. Molecules: DHA: dihydroxyacetone; DHAP: dihydroxyacetone phosphate; GAP: glyceraldehyde-3-phosphate; Xu-5P: xylulose-5-phosphate; FBP: fructose-1, 6-bisphosphate; Glucose-6P: glucose-6-phosphate; Fructose-6P: fructose-6-phosphate; Glucose-1P: glucose-1-phosphate; UDP-glcUA: UDP-glucuronic acid; GlcN-6P: glucosamine-6-phosphate; GlcN-1P: glucosamine-1-phosphate; GlcNAc-1P: N-acetyl-glucosamine-1-phosphate; UDP-GlcNAc: UDP-N-acetyl-glucosamine; HA: hyaluronic acid.The methylotrophic yeast *O. polymorpha* (formerly known as *Hansenula polymorpha* or *Pichia angusta*) is used as an expression platform because of its features, which include fermentation at high cell densities, low-cost substrate utilization, defined synthetic medium, secretion of heterologous proteins in high levels, established strategies for bioreactor cultivation, and availability of synthetic biology tools for genetic manipulation [13]. *O. polymorpha* is widely used to produce biopharmaceuticals, including Insulin (Wosulin®), α-Interferon (Reiferon®), and others [14], and therefore it is an attractive host platform to produce HA. Furthermore, this non-conventional yeast is thermotolerant, and this can be explored for HA production, as it was recently reported that temperature changes affect the chain length of HA as well as the titers [15]. In *B. subtilis*, an increase in the cultivation temperature led to an increase in HA's molecular weight but a decrease in its concentration in the culture broth.

In terms of available genetic tools, a set of endogenous promoters was previously described for *O. polymorpha* with different regulatory mechanisms that allow for developing strategies to regulate gene expression [16]. As a methylotrophic yeast, many of these characterized promoters are related to methanol metabolism, controlling the expression of genes encoding enzymes required to metabolize methanol. Additionally, promoters regulating genes related to nitrate metabolism and inducible by this nitrogen source (*YNR1* [17] and *YNI1* [18]) as well as related to temperature increase (*TPS1* and *HSA1* [19]) were described in *O. polymorpha*. Although a set of promoters is available for this non-conventional yeast, the methanol-inducible one is preferred to control heterologous gene expression. Cultivation strategies using these promoters usually are based on two phases, although other approaches based on derepression of these promoters were also reported [20]. The first step is the growth phase, which focuses on biomass production using glucose or glycerol, followed by a methanol induction phase to produce the heterologous protein. Therefore, the utilization of methanol-inducible promoters allows developing strategies that decouple growth and product synthesis, because in the absence of methanol no significant levels of heterologous protein are achieved. Such decoupled strategies are advantageous for synthesis of compounds which metabolic pathways compete with biomass production, such as HA.

In *O. polymorpha*, the promoter which controls the methanol oxidase gene (pMOX, also referred as pAOX, since the methanol oxidase is the primary alcohol oxidase described in this organism [17]) is preferred for heterologous gene expression [21], and its regulatory mechanisms are widely studied [22]. Although it is mainly regulated by methanol and repressed by glucose, it is active under limiting-glucose conditions [23], leading to promoter leakage, which might affect strategies that require gene regulation. Additionally, pMOX is derepressed in the presence of other carbon sources such as glycerol, xylose, ribose, and sorbitol [24]. Other promoters upregulated by methanol and repressed by glucose include pDHAS (dihydroxyacetone synthase) and pFMD (formaldehyde dehydrogenase) [25]. The latter is an alternative to pMOX since it is considered a strong promoter in the presence of methanol. However, a high level of enzyme production (13.5 g/L) has been achieved using glucose as the carbon source [26]. For HA production, the utilization of methanol-inducible promoters is a feasible strategy to avoid competition between its synthesis and biomass production, which is the main limitation for the heterologous production of HA (Figure 1) [27,28].

Although methanol-inducible promoters are widely used in heterologous protein expression by *O. polymorpha*, their use may be discouraged due to its leakage under some carbon sources and methanol toxicity and flammability, especially at an industrial scale. Thus, promoters regulated by other inducers or constitutive promoters are available alternatives. The strong constitutive pGAP promoter is commonly used for gene regulation in *O. polymorpha* [29]. Other endogenous constitutive promoters already described for this yeast include pTEF1 [13] and pADH1 [30]. However, this type of promoter does not allow one to tune gene regulation, and the heterologous protein is produced similarly regardless of cultivation conditions. Therefore, an alternative genetic tool must be considered for tunning gene regulation. For example, the utilization of genetic switches such as serine integrases enables the building of genetic circuits through targeted DNA rearrangement. If two recognizing sites (*attB* and *attP*) are inserted flanking the desired DNA sequence, the integrase is capable of identifying these sites and flip the sequence at 180° (see details in [31]). Thus, this strategy works as a genetic tool for gene regulation once any genetic element (promoter, coding sequence, and/or terminator) can be constructed flanked by these *att* sites. The rotation of the desired DNA sequence can occur in a specific condition. Recently, a system using the sites *attB* and *attP* and different serine integrases was applied to build a genetic switch in other eukaryotic cells [32].

In this work, the genetic modifications necessary for the heterologous production of HA using *O. polymorpha* as a chassis organism are described (Figure 1). For this, the *hasA* and *hasB* genes were integrated into the *O. polymorpha* genome under different promoters'

regulation. Various combinations were tested to evaluate their influence on HA titers. Two *hasA* genes were used, referred to as *hasAs* (from *Streptococcus zooepidemicus*) and *hasAp* (from *Pasteurella multocida*). The *hasAp* gene was evaluated because, as with the *hasB* (from *Xenopus laevis*) used, we previously demonstrated that the enzymes encoded by these genes were active in *K. lactis*, another non-conventional yeast [11]. The *hasAs* was selected since it is widely used for heterologous production of HA [7,33,34].

Moreover, to better control HA production, we have also employed a genetic switch using a serine integrase to control both *hasAp* and *hasB* gene expression. The integrase-13 (Int13) was selected due to its versatility, successfully applied to design genetic switches in human, bovine, and plant cells [32]. Besides, no point mutations or changes in cell viability caused by the Int13 were detected in cells analyzed in the previous study. Thus, we evaluated a genetic switch using the Int13 as a proof-of-concept in *O. polymorpha* to control the expression of both *hasB* and *hasAp* genes aiming at HA production by this yeast. A capsule-like layer could be seen around cells of the strain containing the genetic switch controlling *hasB* and *hasAp* expression named here as EMB103, using a scanning electron microscopy analysis. In the other strains expressing only *hasB* and *hasA*, the cell surface was similar to the wild-type strain. However, after 48 h of cultivation in an optimized medium, HA was quantified from the culture broth of all three constructed strains tested. This is the first report of an *O. polymorpha* strain developed for HA production to the best of our knowledge.

2. Materials and Methods

2.1. Chemicals and Molecular Biology Procedures

Ampicillin, hygromycin, zeocin, and analytical grade HA for quantification standards were purchased from Sigma-Aldrich. T4 DNA Ligase, Taq DNA polymerase, and restriction endonucleases were from Thermo Fisher. The *hasB* (GenBank accession number MH728986), *hasAp* (GenBank accession number MH728990), and *hasAs* (GenBank AF414053.1) genes were each synthesized and delivered in pBSK plasmids by Síntese Biotecnologia, which also supplied DNA oligonucleotides. The genes *hasB* and *hasAp* were selected from a previous study [11]. The nucleotide sequence for all genes is shown in Supplementary Materials, Sequence S1–S3). The pGEM-T Easy vector, used in cloning procedures, was purchased from Promega (Madison, WI, USA).

General molecular biology procedures were conducted according to Sambrook and Russel [35]. All reagents and kits were supplied by Thermo Fisher (Waltham, MA, USA) and used according to their recommendations. The primers used in this study are shown in Table S1. PCR products were purified using the GeneJet PCR Purification Kit from Thermofisher, and fragments treated with restriction enzymes were recovered from agarose gels 0.8% using the GeneJET Gel Extraction Kit. Plasmid cloning was carried out using T4 DNA Ligase after treating inserts and vectors with appropriate restriction endonucleases. Transformation of *Escherichia coli* through heat shock was performed as described previously [36]. Plasmids were extracted using the GeneJet Plasmid Miniprep Kit from Thermofisher.

2.2. Construction of Plasmids and Strains

The plasmids and strains used in this study are listed in Tables S2 and S3, respectively. Sequences of constructed plasmids are shown in Supplementary Materials, Sequences S5–S11. The *O. polymorpha* strain NCYC495 *yku80* was used and is referred to as Wild Type (WT). This strain, in which the gene encoding KU80 protein was deleted to diminish recombination by non-homologous end joining (NHEJ), and the plasmids of the pHIP series, used as backbone vectors, were cordially provided by Dr. Ida J van der Klei [37]. The *E. coli* DH10B strain (Thermo Fisher) was used for plasmid cloning and expansion. All plasmids were confirmed by restriction enzyme analysis and PCR. In total, four strains of *O. polymorpha* were constructed from the NCYC495 *yku80* strain. The overall strategy for strain construction and the final strain genotypes are summarized in Figure 2 and described in the Supplementary Materials. Stability of construct integration was confirmed thoroughly

throughout the study through PCR of colonies resulting from consecutive passaging on selective media plates.

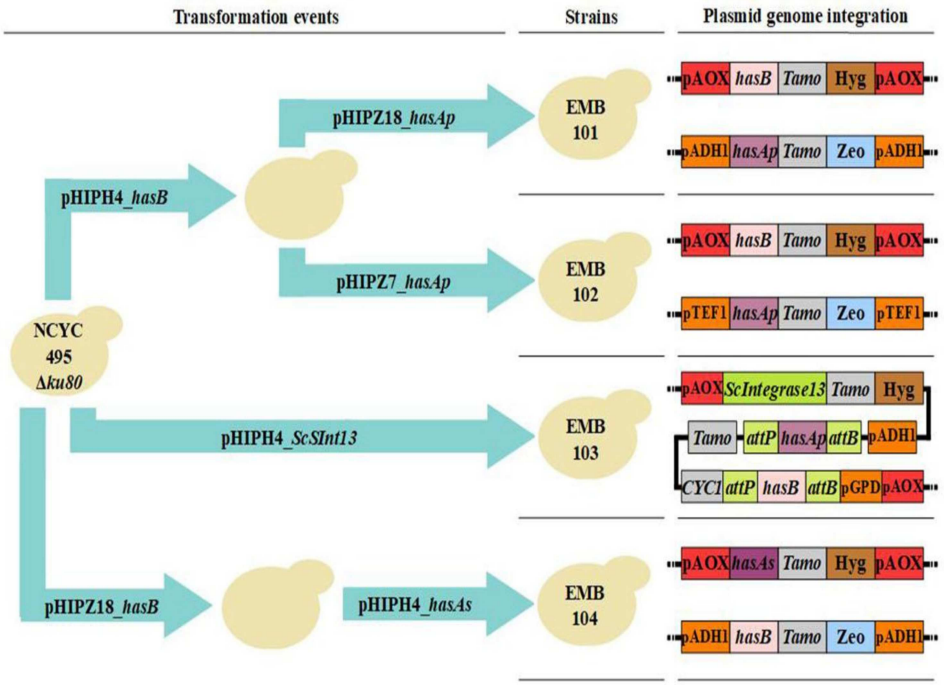


Figure 2. Strains constructed in this study to evaluate HA production in *O. polymorpha*. Light blue arrows represent transformation events with the plasmids specified. Integration of plasmids by single crossover is depicted. Hyg: hygromycin resistance cassette. Zeo: zeocin resistance cassette.

To construct the pHIPH4_*hasB* vector, the *hasB* gene was amplified from pBSK_*hasB* using the primers described in Table S1. Amplification added sites for the restriction enzymes HindIII and SalI to the *hasB* amplicon, which was then ligated into pGEM-T Easy. The resulting pGEM_*hasB* vector and pHIPH4 were treated with HindIII and SalI. The fragment containing *hasB* was purified from an agarose gel for ligation into pHIPH4 between pAOX and the Tamo terminator, resulting in the pHIPH4_*hasB* vector. Similarly, the *hasAp* gene was extracted from the pBSK_*hasAp* vector using HindIII and XhoI, which were also used to remove eGFP-SKL from the pHIPZ18_eGFP-SKL vector. The *hasAp* fragment and the pHIPZ18 backbone were purified from an agarose gel. The former was then ligated between the pADHI promoter and the Tamo terminator of the latter, resulting in the pHIPZ18_*hasAp* vector. The pHIPH4_*hasB* and pHIPZ18_*hasAp* plasmids were inserted into strain NCYC495 *yku80*, resulting in the strain EMB101 (Figure 2).

To construct pHIPZ7_*hasAp*, the pHIPZ18_*hasAp* and pHIPZ7 vectors were treated with HindIII and XhoI and cloned as described above pHIPZ7_*hasAp*, which has the *hasAp* gene under the control of pTEF1 and Tamo terminator. The strain containing the *hasB* gene was transformed with pHIPZ7_*hasAp*, resulting in EMB102 (Figure 2).

The whole pHIPH4_*ScSInt13* plasmid was designed to contain a genetic switch controlling the expression of *hasAp* and *hasB*. The pHIPH4 plasmid was used as a backbone to construct a vector bearing a gene encoding a serine-type phage integrase-13, codon-optimized for *S. cerevisiae*, and the same *hasAp* and *hasB* expression cassettes as in the pHIPZ18_*hasAp* and pHIPH4_*hasB* plasmids. The Integrase13 target sites *attB* and its reverse complement sequence *attP* [38,39] were placed flanking both with genes synthesized in reverse complement orientation [25,26]. The nucleotide sequence for the integrase-13 and the sequences of *attB* and *attP* was based on a previous study [40]. The whole plasmid was synthesized by Epoch Life Science Inc (Missouri City, TX, USA) and confirmed by

sequencing and restriction enzyme analysis. It was linearized in the pAOX promoter and transformed into NCYC495 *yku80* to generate EMB103 (Figure 2).

The pHIPZ18_*hasB* plasmid was constructed using HindIII and XbaI to excise *hasB* from pHIPH4_*hasB* and insert it into the pHIPZ18 vector treated with the same enzymes. Fragment and backbone were gel-purified before ligation. The same restriction enzymes were used to generate the pHIPH4_*hasAs* plasmid by cloning *hasAs* from pBSK_*hasAs* into the pHIPH4 backbone. Strain NCYC495 *yku80* was transformed with linearized pHIPZ18_*hasB* and, after confirmation of *hasB* integration in transformants by colony PCR, the pHIPH4_*hasAs* was inserted, resulting in strain EMB104 (Figure 2).

2.3. Media and Culture Conditions

Escherichia coli cells were cultured in lysogeny broth (LB) at 37 °C and 200 rpm. To select cells harboring plasmids, the medium was supplemented with ampicillin (100 µg/mL).

O. polymorpha strains were cultured and maintained in YPD (1% yeast extract, 2% peptone, and 2% glucose) at 37 °C and 200 rpm. Transformant strains were selected and maintained in media supplemented with 100 µg/mL zeocin or 300 µg/mL hygromycin, according to the resistance markers.

2.4. Electrocompetent Cell Preparation and Transformation

The *O. polymorpha* electrocompetent cells were prepared as described previously [41], using 175 mM β-mercaptoethanol in a substitution of dithiothreitol TED buffer. Sixty microliters of the freshly prepared electrocompetent cells and 5–10 µg of the plasmid linearized at the *O. polymorpha* promoter sequence with the appropriate restriction enzyme were mixed and transferred into an ice-cold 2-m-gap electroporation cuvette. The pulse was performed using a Bio-Rad Xcell MicroPulser cellular electroporator set to 50 µF, 129 Ω, and 1.5 kV (7.5 kV/cm). The recovered cells were plated onto selective media and incubated at 37 °C.0020

2.5. Determination of the Growth Profile

Each constructed strain and the WT were grown overnight in YPD supplemented with 2% glucose and the appropriate antibiotics. On the following day, the cultures were used to inoculate 500 mL baffled shake flasks containing 50 mL of YPD such that initial optical density at 600 nm (OD600) was 0.3. Flasks were then incubated at 37 °C and 200 rpm for biomass production. The cultures were monitored, and samples were taken every two hours during the first 12 h for OD600 measurement using a SpectraMax M2 (Molecular Devices®, San Jose, CA, USA). The cultures were started with glucose as the carbon source as a step focused on producing biomass, thus decoupling growth from the heterologous expression of the genes regulated by the pAOX promoter. Consequently, methanol had to be added (1% at 7 and 21 h) to induce the heterologous expression of *has* genes.

At 25 h, cells were transferred to a fresh medium, since glucose is required to synthesize both HA's precursors (Figure 1). High-density cell cultures were harvested by centrifugation (4000 rpm for 10 min) and transferred to 1-L baffled shake flasks containing 100 mL of medium and kept at 37 °C under agitation (200–250 rpm). The culture ended after 24 h and samples were prepared for Scanning Electron Microscopy (SEM) analysis. All cultures were performed twice.

The maximum growth rate (µMAX) was determined as the slope of the log-transformed linear regression equation. The period of growth used was from 0 h to 6 h, which comprises the exponential growth phase on glucose in the applied conditions.

2.6. Statistical Analysis and Data Presentation

The parameters µMAX (h^{-1}), Final OD, and Doubling time (h), calculated from the growth curves, and HA titer were submitted to analysis of variance (ANOVA) followed by Tukey's post-test using the Graphpad Prism software (version 6.0). The parameters evaluated follow the criteria to choose the one-way ANOVA test. The growth curve in

Figure 3 displays the average values obtained from two biological replicates, with the bars representing the standard deviation (SD). All graphs were prepared using the Graphpad Prism software (version 6.0). All genetic constructions represented in the figures presented in the Supplementary Materials were obtained from the software SnapGene® version 5.1.5.

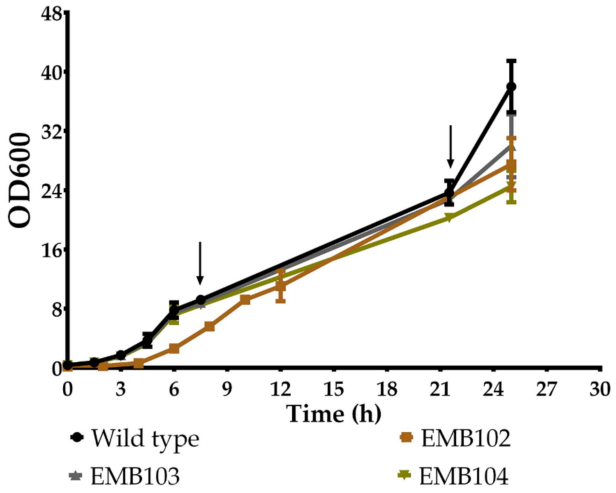


Figure 3. Growth profile of EMB102, EMB103, and EMB104 and *O. polymorpha* NCYC495 *yku80* strain. The experiment was performed in YPD, using shake flasks. The arrows indicate the approximate times of supplementation with methanol to a final concentration of 1%. Means and standard deviations were calculated from two biological replicates.

2.7. Analysis through Scanning Electron Microscopy (SEM)

Cells from 1 mL aliquots of cultures conducted as described above (Section 2.5) were prepared for SEM as described in [11,42]. Images of the surfaces of the cells were made on a Jeol JSM-7000F field emission scanning electron microscope with the assistance of an Emitech K550 automated sputter coater and an Emitech K850 critical point dryer.

2.8. HA Quantification

The protocol applied for HA production and extraction was adapted from a previous study [12]. Cells from each constructed strain and the WT were pre-grown overnight in YPD medium supplemented with appropriate antibiotics, if needed. These cultures were transferred to 1-L baffled shake flasks containing YPD and grown for 24 h aiming a high-cell density culture. Additionally, methanol was added to 1% at 7 and 21 h for induction of *has* genes, as described in Section 2.5. After the 24 h, cells were harvested and washed using a saline solution (9 g/L of NaCl), then transferred to new 1-L baffled shake flasks containing an optimized medium for HA production [43]: 4% glucose, 0.75% yeast extract, 1% peptone, 0.25% K_2HPO_4, 0.05% $MgSO_4$, 0.5% NaCl, 0.04% glutamine, 0.06% glutamic acid, and 0.02% oxalic acid. Glucose was added to 2% during the cultivation to avoid glucose-starvation as indicated by the OD_{600} measured. The cultivation ended after 48 h and cultures were prepared for HA quantification.

For HA extraction, 100 mL samples from each culture were purified as described by [12]. Briefly, an equal volume of 0.1% (w/v) SDS solution was added, then the samples were kept under orbital agitation for 10 min, so that the HA capsule was separated from the cells after being centrifuged for 5 min at 4500× *g*. 1.5 volumes of ethanol were added to the supernatant, which was incubated at 4 °C overnight, for at least 10 h. Then, samples were centrifuged at 4500× *g* and 4 °C for 30 min. The pellet was washed with 25 mL of a 75% ethanol and 25% 0.15 M NaCl solution, then centrifuged once more. The supernatant was discarded, and the pellet was incubated at room temperature and resuspended in MilliQ water after the ethanol evaporated. The WT strain culture was used as a nega-

tive control. The quantifications were performed in three technical replicates from each biological replicate.

The carbazole method was used to quantify HA as previously described [44]. Firstly, 200 µL of water (blank), HA standard and purified samples were added to ice-cold glass tubes with 500 µL of a 25 mM solution of sodium tetraborate in sulfuric acid. After homogenizing, the tubes were incubated at 100 °C under agitation for 10 min and kept at room temperature for cooling. Twenty µL of a 0.125% carbazole solution on methanol was added to each tube. After a second incubation (100 °C for 15 min under agitation), the samples were cooled to room temperature prior reading. The 200 µL volume of each tube was applied to a 96-well microplate, which was read at 530 nm by the Biotek EON microplate reader. The samples, as well as the points of the standard curve, were analyzed in technical triplicate (Supplementary Materials, Figures S16 and S17). The ten-point standard curve used comprised the concentrations of 10 µg/mL, 20 µg/mL, 50 µg/mL, 100 µg/mL, 200 µg/mL, 300 µg/mL, 400 µg/mL, 500 µg/mL, 600 µg/mL, and 700 µg/mL.

3. Results and Discussion

3.1. Strain Construction

Table 1 lists the strains developed in this study, including the promoters utilized for each gene as well as the origin of the heterologous genes. Integration of synthetic constructs into their genomes and their genetic stability were assessed through consecutive cultivation in plates containing antibiotics, followed by PCR. Figures of the 1% agarose gels are shown in Supplementary Materials (Figures S1–S5).

Table 1. Summary of all EMB strains constructed in this study.

EMB Strains	Promoter Driving *hasA*	Source of *hasA*	Promoter Driving *hasB*	Source of *hasB*
EMB102	pTEF1	*P. multocida*	pAOX	*Xenopus laevis*
EMB103	pADH1	*P. multocida*	pGPD (*S. cerevisiae*)	*X. laevis*
EMB104	pAOX	*S. zooepidemicus*	pADH1	*X. laevis*

The presence of *hasB* in strain EMB101 was confirmed by colony PCR. However, even though EMB101 was capable of growing in the presence of the pHIPZ18 plasmid selection marker, zeocin, the presence of *hasAp* was not detected by colony PCR after the third passage in the selective medium (Supplementary Materials, Figure S2). For this reason, EMB101 was not used in further experiments. Assessment of *hasAp* instability when integrated into the *ADH1* locus has not been investigated further. Both *hasA* and *hasB* could be verified in strains EMB102, EMB103, and EMB104 (Supplementary Materials, Figures S3–S5). For the strain EMB103, in which Int13 was used as a genetic switch to control the expression of *hasA* and *hasB*, PCR analysis confirmed the rotation of both genes (Supplementary Materials, Figure S4). Thus, these strains were used in subsequent experiments.

3.2. Growth Profile of Recombinant Strains

Decoupling cell growth and HA production is crucial to its synthesis. Thus, the cultivation was divided into two steps: biomass and enzymes production, and HA synthesis. For the first step, the constructed strains EMB102, EMB103, and EMB104 as well the WT strain were grown on shake flasks containing YPD aiming at achieving high-cell density cultures (Figure 3). Since in the presence of glucose the pAOX promoter is repressed [23], no HA synthesis is expected in this step. At the end of the exponential phase estimated by the OD600 measured (7 h), methanol was added to the cultures to induce the expression of *has* genes required for HA production. Methanol was also added at 21 h of cultivation to ensure high levels of enzymes for HA production (Figure 3). Therefore, this first step of cultivation comprised biomass production and accumulation of the enzymes required for HA synthesis.

While the strains had similar specific growth rate during the exponential phase, they reached different final ODs (Table 2). Nevertheless, the highest mean of final OD, from the WT strain, was statistically different only for the EMB104 strain ($p < 0.05$). This negative impact on cell growth for the engineered strains could be related to (i) a high-level of heterologous protein production after induction using methanol or (ii) production of HA, which impacts cell growth. Indeed, the competition between HA production and cell growth was previously described [6,28].

Table 2. The final OD_{600} and maximum specific growth rates (μMAX) from wild type and EMB strains as well as the HA titer obtained by EMB strains when cultivated in shake flasks.

Strains	μMAX (h^{-1}) *	Final OD *	Doubling Time (h) *	Hyaluronic Acid (HA) Titer (μg/mL) **	Final OD **
WT	0.50 ± 0.02^{a}	38.00 ± 3.46^{a}	1.35 ± 0.06^{a}	NA	21.00 ± 1.41^{a}
EMB102	0.45 ± 0.04^{a}	$27.50 \pm 3.54^{a,b}$	1.55 ± 0.15^{a}	$151.20 \pm 13.04^{a,b}$	27.00 ± 0.00^{a}
EMB103	0.52 ± 0.01^{a}	$30.00 \pm 4.24^{a,b}$	1.34 ± 0.01^{a}	197.76 ± 5.66^{b}	27.50 ± 0.71^{a}
EMB104	0.49 ± 0.02^{a}	24.50 ± 2.12^{b}	1.41 ± 0.07^{a}	123.20 ± 26.56^{a}	21.50 ± 3.54^{a}

NA: Not applied. Equal letters indicate no statistical difference between the strains ($p < 0.05$). * Parameters were calculated based on the data obtained by the growth curves presented in Figure 3 and described in Material and Methods (Section 2.5). ** Values were obtained after 48 h of cultivation in medium optimized for HA production [12,43].

For the *K. lactis* HA-producing yeast, HA production had a negative impact on final cell density [11]. This might be linked to the fact that the HA precursor UDP-GlcNAc is a cell wall component for yeasts, and thus, HA synthesis competes with biomass formation. As methanol is necessary for the expression of the HA synthesis enzymes on the constructed strains, the competition between biomass and HA production only becomes noticeable after the addition of methanol.

After the first phase of growth for biomass generation and induction of *has* genes (Figure 3), the entire culture was transferred to fresh YPD medium, aiming at HA production. In this case, the production was assessed by SEM.

3.3. HA Production Assessed by Scanning Electron Microscopy Analysis

Given that HA is produced and secreted by the cells, a capsular structure can be visualized on their surface. This method was reported previously to confirm the production of HA by *L. lactis* [45] and by the yeast *K. lactis* [11]. The formation of a capsule-like layer was also observed in HA-producing strains of *C. glutamicum* [46], although a phase-contrast microscopy analysis was applied instead of SEM. In this case, the capsule-like layer affected nutrient uptake and cell metabolism, which impaired HA production [46].

As seen on SEM images from 24 h cultures, WT cells are clearly discrete, with separate cell walls and moderate aggregation. Cells from the EMB102 and EMB104 strains showed no significant superficial differences compared to the WT strain, although a discrete capsule-like structure was observed (Figure 4B). Strain EMB103 is clearly distinct from the others. Firstly, discrete cell walls between cells are not distinguishable at the same magnification as in the other strains (Figure 4B). Likewise, seemingly intensified aggregation is observed, and free individual cells are less common. Lastly, EMB103 cells seem to be covered in a substance that glues them together and generates an irregular cell surface indicating the production of HA (Figure 5).

The EMB103 strain was constructed containing a genetic switch for the expression of *hasAp* and *hasB* which allows expression of both genes after the induction of a gene encoding a serine-like integrase-13. Thus, the formation of a capsule-like layer by the EMB103 cells indicates the genetic switch constructed for *O. polymorpha* works properly. Recently, the utilization of integrases to design genetic circuits for gene regulation is emerging as a powerful synthetic biology tool [31]. Serine integrases are also used as a genetic tool to mediate in vivo site-specific recombination between an exogenous DNA fragment, such as plasmids or gene expression cassettes, into a target genome [47]. This has become a versatile

system since its application was demonstrated in prokaryotes [48] and higher eukaryotes such as human, bovine, and plant cells [32]. For yeast cells, several integrases showed activity when expressed by *S. cerevisiae*, with the highest activity reached by the ϕBT1 integrase [49]. Additionally, the authors verified no site damage by integrase-mediated recombination, which is important to reduce point mutations during the site-specific recombination events. In another study using baker's yeast, the PhiC31 integrase was used to successfully recycle the selective markers HIS5 and LEU2 [50]. However, it was verified that the system was efficient only when a low-copy vector bearing the integrase-encoding gene was used, probably due to the toxicity of PhiC31 integrase. The Int13 utilized in this study showed no point mutation or impact on cell viability when tested in different eukaryotic cells [23], representing a reliable genetic tool for yeast cells. Nevertheless, there is no report of a genetic switch using serine integrases for gene regulation in non-conventional yeasts. Here, as a proof of concept, the system was applied to construct an *O. polymorpha* strain that produces HA and showed to be the best genetic approach for constructing HA producing *O. polymorpha* strains.

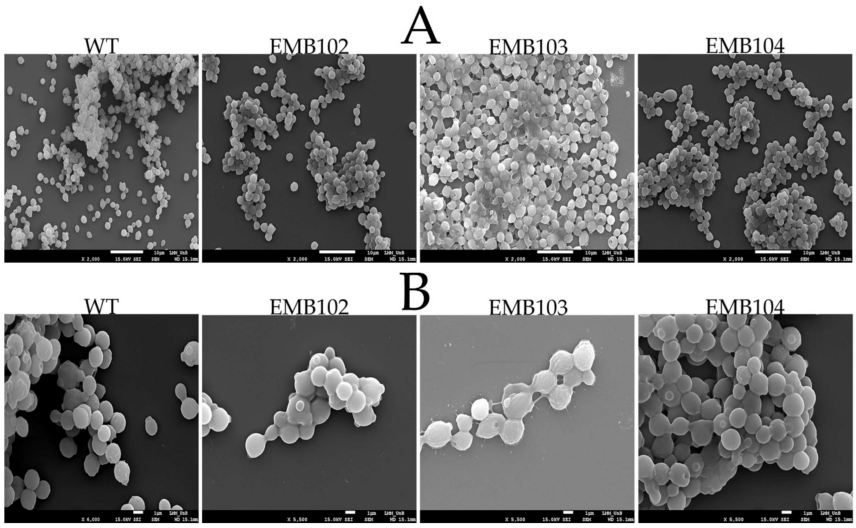


Figure 4. Scanning Electron Microscopy of different *O. polymorpha* strains. The names above the image correspond to the respective strain. Images were acquired at two magnifications: (**A**) 2000× and (**B**) 5500×.

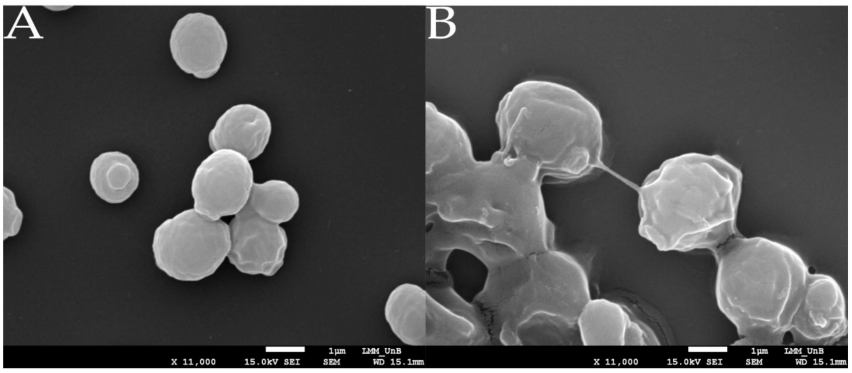


Figure 5. Scanning Electron Microscopy of Wild Type (WT) (**A**) and EMB103 (**B**) in the magnitude of 11,000×.

3.4. HA Quantification

Since a discrete difference between the cell surface of EMB102 and EMB104 compared to the WT was observed, it was hypothesized that HA production could be limited by growth conditions as well as the HA purification protocol applied. Therefore, an adaptation

from the previous protocol applied for *K. phaffi* was performed [12], which is an optimized medium for HA. By this approach, it was possible to quantify the HA in the supernatant from all constructed strains (Table 2). After 48 h of culture, HA titers were assessed through the carbazole method, which was able to accurately quantify the production of HA (Table 2 and Supplementary Materials, Figure S17).

The highest titer of HA (197.76 µg/mL) was achieved by strain EMB103, although its result is not statistically different from that of the EMB102 strain (151.20 µg/mL). The EMB104 strain achieved the lowest HA concentration in the study (123.20 µg/mL). Furthermore, both HAS encoding genes (*hasAp* and *hasAs*) seem to work properly in *O. polymorpha*, indicating that this yeast is a suitable host platform for the application of different strategies for HA production.

Apparently, HA production in *O. polymorpha* in the conditions applied is limited leading to a low synthesis by all strains. Nevertheless, these results demonstrate the potential of *O. polymorpha* as a platform host for HA production, although the process requires optimization regarding cell cultivation. For instance, for the *K.phaffi* HA-producing strain, cultivation was performed in fermenters using a fed-batch mode with controlled dissolved oxygen levels and only 200 mg/L of HA was detected after 48 h of culture when both *hasA* and *hasB* were expressed [12]. Therefore, the conditions used here to produce, and isolate HA are sub-optimal. Furthermore, the purification of HA was conducted solely on the supernatant of cultures. Yet, it has been determined that the HAS enzyme from *P. multocida* serogroup A:1, differently from other HAS, is actually cytosolic and does not possess an HA exporting function [51]. Indeed, there is published evidence of up to 80% of heterologously produced HA inside of algae host cells [52]. While it is established that HA is secreted to the extracellular medium simultaneously with its synthesis by streptococcal HAS enzymes, strains EMB 102 and EMB 103, harboring the *hasAp* gene, might still produce intracellular HA. This evidence suggests that lysates of cultured cells might also be sources for HA quantification and their use in quantification in addition to the supernatant should be considered, depending on the experimental setup and enzymes used.

Additionally, the balance between the two precursors plays a key role in HA production [53,54]. Intracellular concentrations of precursors present a challenge for HA production, and equimolar concentration of UDP-GlcNAc and UDP-GlcUA is required for efficient synthesis of HA. In this regard, the availability of UDP-GlcNAc presents a hurdle, for it is diverted to cell wall formation in *O. polymorpha* [55]. Lastly, the Pentose Phosphate Pathway (PPP) deviates Glucose-6-phosphate which is required for HA precursors biosynthesis. Downregulation of the *zwf* gene, which encodes for the first enzyme of PPP, is a well-described strategy for improving HA production [8,9,56]. The *O. polymorpha* NCYC495 strain, when grown at temperatures above 30 °C, showed an increased flux towards PPP higher than the other thermotolerant yeast tested, *Kluyveromyces marxianus* [57]. It is worth mentioning that PPP and methanol assimilation are intrinsically related in *Ogataea* sp. [58] as well as in other methylotrophic yeasts [59]. Xylulose-5-phosphate, an intermediate from PPP, is required for methanol assimilation, specifically for formaldehyde fixation by the enzyme dihydroxyacetone synthase (Figure 1). In fact, in *O. parapolymorpha* (previously known as *Hansenula polymorpha* DL-1), genes encoding enzymes from the non-oxidative phase of PPP are upregulated in the presence of methanol [58]. Therefore, PPP flux is an important bottleneck to be considered for the heterologous production of HA by *O. polymorpha*. Strategies aiming at funneling the metabolic framework of *O. polymorpha* are required to ensure a proper flux through HA precursors and its efficient synthesis.

Nevertheless, the production of up to 197.76 µg/mL of HA by *O. polymorpha* has been reported, demonstrating the potential of this non-conventional yeast for production of HA, although several aspects of the process require further optimization, as discussed above.

4. Conclusions

O. polymorpha was engineered for HA production. Both *hasB* (*X. laevis*) and two different *hasA* genes, from *P. multocida* (*hasAp*) and *S. zooepidemicus* (*hasAs*), were successfully

integrated into the *O. polymorpha* genome under the control of different promoters. Additionally, a genetic switch using a serine-type integrase-13 to control the expression of both *hasAp* and *hasB* was demonstrated. Both the qualitative and quantitative methods were able to show HA synthesis by the yeast. The capsule-like structure observed in EMB103 cells by SEM analysis indicates HA production by this strain. Indeed, the EMB103 was able to produce 197.76 µg/mL, which is considered a low HA titer. Additionally, it was demonstrated that both HAS encoding genes from *P. multocida* and *S. zooepidemicus* can be utilized in *O. polymorpha*, which enables the design of different strategies for HA production in this industrial yeast. These results demonstrate not only the potential of *O. polymorpha* as a host platform for HA production but also of integrases for gene regulation in this yeast.

Supplementary Materials: The following are available online at https://www.mdpi.com/2076-2607/9/2/312/s1. Table S1: List of primers used in this work; Table S2: List of plasmids used in this work; Table S3: List of strains used in this work; Sequences S1–S4: Gene sequences; Sequences S5–S11: Sequences of constructed vectors; Figures S1–S5: Confirmation of constructed strains; Figures S6–S15: Plasmid maps; Figure S16: Carbazole standard curve; Figure S17: Carbazole assay plate picture.

Author Contributions: J.H.C.M.-N. constructed the strains, performed the growth curves, prepared the samples for scanning electron microscopy analysis, analyzed the data, and drafted the manuscript. E.B.Q. helped to perform the experiments, prepared Figures 1 and 2, and helped to write the manuscript. K.A.R. helped to construct the strains and write the paper. C.M.C. designed the plasmid for the genetic switch, and revised and commented on the manuscript. H.C.P. revised and commented on the manuscript. E.L.R. revised and commented on the manuscript and conceived the study. N.S.P. conceived and coordinated the study, got funding to support the experimental research, helped with the data interpretation, and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the INCT Biologia Sintética (INCT-BioSyn), the Brazilian National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES), and Fundação de Apoio à Pesquisa do Distrito Federal (FAP-DF).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the text and Supplementary Materials.

Acknowledgments: The authors would like to thank Ida J van der Klei (Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen) for kindly providing the *O. polymorpha* NCYC495 yku80 strain and the plasmids pHIP utilized in this work. The authors also thank Sônia Nair Bao and the Laboratório de Microscopia e Microanálises (LMM) from the University of Brasilia for assistance in the electron microscopy analysis performed in this study.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. De Carvalho, C.C.C.R. Whole cell biocatalysts: Essential workers from nature to the industry. *Microb. Biotechnol.* **2017**, *10*, 250–263. [CrossRef] [PubMed]
2. Lin, B.; Tao, Y. Whole-cell biocatalysts by design. *Microb. Cell Fact.* **2017**, *16*, 106. [CrossRef] [PubMed]
3. Fallacara, A.; Baldini, E.; Manfredini, S.; Vertuani, S. Hyaluronic acid in the third millennium. *Polymers* **2018**, *10*, 701. [CrossRef] [PubMed]
4. *Hyaluronic Acid Market Size, Share & Trends Analysis Report by Application (Dermal Fillers, Osteoarthritis (Single Injection, Three Injection, Five Injection), Ophthalmic, Vesicoureteral Reflux), by Region, and Segment Forecasts, 2020–2027*; Grand View Research: San Francisco, CA, USA, 2020; p. 150.
5. Sze, J.H.; Brownlie, J.C.; Love, C.A. Biotechnological production of hyaluronic acid: A mini review. *3 Biotech* **2016**, *6*, 67. [CrossRef] [PubMed]
6. de Oliveira, J.D.; Carvalho, L.S.; Gomes, A.M.V.; Queiroz, L.R.; Magalhães, B.S.; Parachin, N.S. Genetic basis for hyper production of hyaluronic acid in natural and engineered microorganisms. *Microb. Cell Factories* **2016**, *15*, 119. [CrossRef]

7. Jin, P.; Kang, Z.; Yuan, P.; Du, G.; Chen, J. Production of specific-molecular-weight hyaluronan by metabolically engineered *Bacillus subtilis* 168. *Metab. Eng.* **2016**, *35*, 21–30. [CrossRef]
8. Westbrook, A.W.; Ren, X.; Oh, J.; Moo-Young, M.; Chou, C.P. Metabolic Engineering to Enhance Heterologous Production of Hyaluronic Acid in *Bacillus subtilis*. *Metab. Eng.* **2018**, *47*, 401–413. [CrossRef]
9. Cheng, F.; Yu, H.; Stephanopoulos, G. Engineering *Corynebacterium glutamicum* for high-titer biosynthesis of hyaluronic acid. *Metab. Eng.* **2019**, *55*, 276–289. [CrossRef]
10. Jeeva, P.; Shanmuga Doss, S.; Sundaram, V.; Jayaraman, G. Production of controlled molecular weight hyaluronic acid by glucostat strategy using recombinant *Lactococcus lactis* cultures. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 4363–4375. [CrossRef]
11. Gomes, A.M.V.; Netto, J.H.C.M.; Carvalho, L.S.; Parachin, N.S. Heterologous hyaluronic acid production in *Kluyveromyces Lactis*. *Microorganisms* **2019**, *7*, 294. [CrossRef] [PubMed]
12. Jeong, E.; Shim, W.Y.; Kim, J.H. Metabolic engineering of *Pichia Pastoris* for production of hyaluronic acid with high molecular weight. *J. Biotechnol.* **2014**, *185*, 28–36. [CrossRef] [PubMed]
13. Saraya, R.; Krikken, A.M.; Kiel, J.A.K.W.; Baerends, R.J.S.; Veenhuis, M.; Klei, I.J. Novel genetic tools for *Hansenula polymorpha*. *FEMS Yeast Res.* **2012**, *12*, 271–278. [CrossRef] [PubMed]
14. Gellissen, G.; Kunze, G.; Gaillardin, C.; Cregg, J.; Berardi, E.; Veenhuis, M.; Vanderklei, I. New yeast expression platforms based on methylotrophic *Hansenula polymorpha* and *Pichia pastoris* and on dimorphic *Arxula adenivorans* and *Yarrowia lipolytica*–A Comparison. *FEMS Yeast Res.* **2005**, *5*, 1079–1096. [CrossRef] [PubMed]
15. Li, Y.; Li, G.; Zhao, X.; Shao, Y.; Wu, M.; Ma, T. Regulation of hyaluronic acid molecular weight and titer by temperature in engineered *Bacillus subtilis*. *3 Biotech* **2019**, *9*, 225. [CrossRef]
16. Gündüz Ergün, B.; Hüccetoğulları, D.; Öztürk, S.; Çelik, E.; Çalık, P. Established and upcoming yeast expression systems. In *Recombinant Protein Production in Yeast*; Gasser, B., Mattanovich, D., Eds.; Methods in Molecular Biology; Springer: New York, NY, USA, 2019; Volume 1923, pp. 1–74. ISBN 978-1-4939-9023-8.
17. Avila, J.; Pérez, M.D.; Brito, N.; González, C.; Siverio, J.M. Cloning and disruption of the *YNR1* gene encoding the nitrate reductase apoenzyme of the yeast *Hansenula polymorpha*. *FEBS Lett.* **1995**, *366*, 137–142. [CrossRef]
18. Brito, N.; Avila, J.; Pérez, M.D.; González, C.; Siverio, J.M. The genes YNI1 and YNR1, encoding nitrite reductase and nitrate reductase respectively in the yeast *Hansenula polymorpha*, are clustered and co-ordinately regulated. *Biochem. J.* **1996**, *317*, 89–95. [CrossRef]
19. Amuel, C.; Gellissen, G.; Hollenberg, C.P.; Suckow, M. Analysis of heat shock promoters in *Hansenula Polymorpha*: The TPS1 promoter, a novel element for heterologous gene expression. *Biotechnol. Bioprocess Eng.* **2000**, *5*, 247–252. [CrossRef]
20. Stöckmann, C.; Scheidle, M.; Dittrich, B.; Merckelbach, A.; Hehmann, G.; Melmer, G.; Klee, D.; Büchs, J.; Kang, H.A.; Gellissen, G.; et al. Process development in *Hansenula polymorpha* and *Arxula adeninivorans*, a Re-Assessment. *Microb. Cell Factories* **2009**, *8*, 10. [CrossRef]
21. Manfrão-Netto, J.H.C.; Gomes, A.M.V.; Parachin, N.S. Advances in using *Hansenula Polymorpha* as chassis for recombinant protein production. *Front. Bioeng. Biotechnol.* **2019**, *7*, 94. [CrossRef]
22. Suppi, S.; Michelson, T.; Viigand, K.; Alamäe, T. Repression vs. Activation of *MOX*, *FMD*, *MPP1* and *MAL1* promoters by sugars in *Hansenula polymorpha*: The outcome depends on cell's ability to phosphorylate sugar. *FEMS Yeast Res.* **2013**, *13*, 219–232. [CrossRef]
23. Dusny, C.; Schmid, A. The *MOX* promoter in *Hansenula polymorpha* is ultrasensitive to glucose-mediated carbon catabolite repression. *FEMS Yeast Res.* **2016**, *16*, fow067. [CrossRef] [PubMed]
24. Hartner, F.S.; Glieder, A. Regulation of methanol utilisation pathway genes in yeasts. *Microb. Cell Fact* **2006**, *5*, 39. [CrossRef] [PubMed]
25. Vogl, T.; Fischer, J.E.; Hyden, P.; Wasmayer, R.; Sturmberger, L.; Glieder, A. Orthologous promoters from related methylotrophic yeasts surpass expression of endogenous promoters of pichia pastoris. *AMB Expr.* **2020**, *10*, 38. [CrossRef] [PubMed]
26. Mayer, A.F.; Hellmuth, K.; Schlieker, H.; Lopez-Ulibarri, R.; Oertel, S.; Dahlems, U.; Strasser, A.W.M.; van Loon, A.P.G.M. An expression system matures: A highly efficient and cost-effective process for phytase production by recombinant strains of *Hansenula polymorpha*. *Biotechnol. Bioeng.* **1999**, *63*, 373–381. [CrossRef]
27. Chong, B.F.; Blank, L.M.; Mclaughlin, R.; Nielsen, L.K. Microbial hyaluronic acid production. *Appl. Microbiol. Biotechnol.* **2005**, *66*, 341–351. [CrossRef]
28. Liu, L.; Liu, Y.; Li, J.; Du, G.; Chen, J. Microbial production of hyaluronic acid: Current state, challenges, and perspectives. *Microb. Cell Fact* **2011**, *10*, 99. [CrossRef]
29. Kurylenko, O.O.; Ruchala, J.; Vasylyshyn, R.V.; Stasyk, O.V.; Dmytruk, O.V.; Dmytruk, K.V.; Sibirny, A.A. Peroxisomes and peroxisomal transketolase and transaldolase enzymes are essential for xylose alcoholic fermentation by the methylotrophic thermotolerant yeast, *Ogataea* (*Hansenula*) *polymorpha*. *Biotechnol. Biofuels* **2018**, *11*, 197. [CrossRef]
30. Singh, R.; Manivannan, S.; Krikken, A.M.; Boer, R.; Bordin, N.; Devos, D.P.; Klei, I.J. *Hansenula polymorpha* Pex37 is a peroxisomal membrane protein required for organelle fission and segregation. *FEBS J.* **2020**, *287*, 1742–1757. [CrossRef]
31. Merrick, C.A.; Zhao, J.; Rosser, S.J. Serine integrases: Advancing synthetic biology. *ACS Synth. Biol.* **2018**, *7*, 299–310. [CrossRef]
32. Gomide, M.S.; Sales, T.T.; Barros, L.R.C.; Limia, C.G.; de Oliveira, M.A.; Florentino, L.H.; Barros, L.M.G.; Robledo, M.L.; José, G.P.C.; Almeida, M.S.M.; et al. Genetic switches designed for eukaryotic cells and controlled by serine integrases. *Commun. Biol.* **2020**, *3*, 255. [CrossRef]

33. Prasad, S.B.; Jayaraman, G.; Ramachandran, K.B. Hyaluronic acid production is enhanced by the additional co-expression of UDP-glucose pyrophosphorylase in *Lactococcus lactis*. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 273–283. [CrossRef] [PubMed]
34. Jongsareejit, B.; Bhumiratana, A.; Morikawa, M.; Kanaya, S. Cloning of hyaluronan synthase (Sz-Has) gene from *Streptococcus zooepidemicus* in *Escherichia coli*. *ScienceAsia* **2007**, *33*, 389. [CrossRef]
35. Sambrook, J.; Russell, D.W. *Molecular Cloning-A Laboratory Manual*, 3rd ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor: New York, NY, USA, 2001.
36. Green, M.R.; Sambrook, J. The hanahan method for preparation and transformation of competent, *E. Coli*: High-efficiency transformation. *Cold Spring Harbor Protoc.* **2018**, *2018*. pdb.prot101188. [CrossRef] [PubMed]
37. Van der Klei, I.J. The *Hansenula polymorpha* Expression System. Available online: https://www.rug.nl/research/molecular-cell-biology/research/the-hansenula-polymorpha-expression-system (accessed on 19 May 2020).
38. Nash, H.A. Integration and excision of *Bacteriophage* λ: The Mechanism of conservative site specific recombination. *Annu. Rev. Genet.* **1981**, *15*, 143–167. [CrossRef] [PubMed]
39. Rutherford, K.; Yuan, P.; Perry, K.; Sharp, R.; Van Duyne, G.D. Attachment site recognition and regulation of directionality by the serine integrases. *Nucleic Acids Res.* **2013**, *41*, 8341–8356. [CrossRef] [PubMed]
40. Yang, L.; Nielsen, A.A.K.; Fernandez-Rodriguez, J.; McClune, C.J.; Laub, M.T.; Lu, T.K.; Voigt, C.A. Permanent genetic memory with >1-byte capacity. *Nat. Methods* **2014**, *9*, 1261–1266. [CrossRef]
41. Saraya, R.; Gidijala, L.; Veenhuis, M.; van der Klei, I.J. Tools for genetic engineering of the yeast *Hansenula polymorpha*. In *Yeast Metabolic Engineering*; Mapelli, V., Ed.; Methods in Molecular Biology; Springer: New York, NY, USA, 2014; Volume 1152, pp. 43–62. ISBN 978-1-4939-0562-1.
42. Murtey, M.D.; Ramasamy, P. Sample preparations for scanning electron microscopy–life sciences. In *Modern Electron Microscopy in Physical and Life Sciences*; Janecek, M., Kral, R., Eds.; InTech: London, UK, 2016; ISBN 978-953-51-2252-4.
43. Im, J.-H.; Song, J.-M.; Kang, J.-H.; Kang, D.-J. Optimization of medium components for high-molecular-weight hyaluronic acid production by *Streptococcus* Sp. ID9102 via a statistical approach. *J. Ind. Microbiol. Biotechnol.* **2009**, *8*, 1337. [CrossRef]
44. Bitter, T.; Muir, H.M. A modified uronic acid carbazole reaction. *Analytical Biochem.* **1962**, *4*, 330–334. [CrossRef]
45. Hmar, R.V.; Prasad, S.B.; Jayaraman, G.; Ramachandran, K.B. Chromosomal integration of hyaluronic acid synthesis (*Has*) genes enhances the molecular weight of hyaluronan produced in *Lactococcus lactis*. *Biotechnol. J.* **2014**, *9*, 1554–1564. [CrossRef]
46. Wang, Y.; Hu, L.; Huang, H.; Wang, H.; Zhang, T.; Chen, J.; Du, G.; Kang, Z. Eliminating the capsule-like layer to promote glucose uptake for hyaluronan production by engineered *Corynebacterium glutamicum*. *Nat. Commun.* **2020**, *11*, 3120. [CrossRef]
47. Yang, Z.; Blenner, M. Genome editing systems across yeast species. *Curr. Opin. Biotechnol.* **2020**, *66*, 255–266. [CrossRef] [PubMed]
48. Gao, H.; Taylor, G.; Evans, S.K.; Fogg, P.C.M.; Smith, M.C.M. Application of serine integrases for secondary metabolite pathway assembly in *Streptomyces*. *Synth. Syst. Biotechnol.* **2020**, *5*, 111–119. [CrossRef] [PubMed]
49. Xu, Z.; Brown, W.R.A. Comparison and optimization of ten phage encoded serine integrases for genome engineering in *Saccharomyces Cerevisiae*. *BMC Biotechnol.* **2016**, *16*, 13. [CrossRef] [PubMed]
50. Snoeck, N.; De Mol, M.L.; Van Herpe, D.; Goormans, A.; Maryns, I.; Coussement, P.; Peters, G.; Beauprez, J.; De Maeseneire, S.L.; Soetaert, W. Serine Integrase Recombinational Engineering (SIRE): A versatile toolbox for genome editing. *Biotechnol. Bioeng.* **2019**, *116*, 364–374. [CrossRef]
51. Chung, J.Y.; Zhang, Y.; Adler, B. The capsule biosynthetic locus of *Pasteurella multocida* A:1. *FEMS Microbiol. Lett.* **1998**, *166*, 289–296. [CrossRef]
52. Zhang, L.; Toscano Selão, T.; Nixon, P.J.; Norling, B. Photosynthetic conversion of CO_2 to hyaluronic acid by engineered strains of the cyanobacterium *Synechococcus* sp. PCC 7002. *Algal Res.* **2019**, *44*, 101702. [CrossRef]
53. Badle, S.S.; Jayaraman, G.; Ramachandran, K.B. Ratio of intracellular precursors concentration and their flux influences hyaluronic acid molecular weight in *Streptococcus zooepidemicus* and recombinant *Lactococcus lactis*. *Bioresour. Technol.* **2014**, *163*, 222–227. [CrossRef]
54. Kaur, M.; Jayaraman, G. Hyaluronan production and molecular weight is enhanced in pathway-engineered strains of lactate dehydrogenase-deficient *Lactococcus lactis*. *Metab. Eng. Commun.* **2016**, *3*, 15–23. [CrossRef]
55. Park, J.-N.; Choo, J.; Kang, H.A. Functional analysis of a *Hansenula polymorpha* MNN2-2 homologue encoding a putative UDP-N-acetylglucosamine transporter localized in the endoplasmic reticulum. *J. Microbiol.* **2011**, *49*, 1012–1017. [CrossRef]
56. Woo, J.E.; Seong, H.J.; Lee, S.Y.; Jang, Y.-S. Metabolic engineering of *Escherichia coli* for the production of hyaluronic acid from glucose and galactose. *Front. Bioeng. Biotechnol.* **2019**, *7*, 351. [CrossRef]
57. Lehnen, M.; Ebert, B.E.; Blank, L.M. Elevated temperatures do not trigger a conserved metabolic network response among thermotolerant yeasts. *BMC Microbiol.* **2019**, *19*, 100. [CrossRef] [PubMed]
58. Ravin, N.V.; Eldarov, M.A.; Kadnikov, V.V.; Beletsky, A.V.; Schneider, J.; Mardanova, E.S.; Smekalova, E.M.; Zvereva, M.I.; Dontsova, O.A.; Mardanov, A.V.; et al. Genome sequence and analysis of methylotrophic yeast *Hansenula polymorpha* DL1. *BMC Genom.* **2013**, *14*, 837. [CrossRef] [PubMed]
59. Rußmayer, H.; Buchetics, M.; Gruber, C.; Valli, M.; Grillitsch, K.; Modarres, G.; Guerrasio, R.; Klavins, K.; Neubauer, S.; Drexler, H.; et al. Systems-level organization of yeast methylotrophic lifestyle. *BMC Biol.* **2015**, *13*, 80. [CrossRef] [PubMed]

Supplementary Material

Evaluation of *Ogataea* (*Hansenula*) *polymorpha* for hyaluronic acid production

João Heitor Colombelli Manfrão-Netto[1], Enzo Bento Queiroz[1], Kelly Assis Rodrigues[1], Cintia M. Coelho[2], Hugo Costa Paes[3], Elíbio Leopoldo Rech Filho[4], Nádia Skorupa Parachin*[1,5]

[1] Grupo Engenharia de Biocatalisadores, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília-DF, Brazil

[2] Department of Genetics and Morphology, Institute of Biological Science, University of Brasília, Brasília, 70910900, DF, Brazil

[3] Clinical Medicine Division, University of Brasília Medical School, University of Brasília, Brasília, Federal District 70910-900, Brazil

[4] Brazilian Agriculture Research Corporation – Embrapa – Genetic Resources and Biotechnology – CENARGEN, Brasília 70770917 DF, Brazil.

[5] Present address: Ginkgo Bioworks 27th Drydock Ave. 8th floor, Boston, MA, 02210.

*Correspondence: nadiasp@unb.br; nadiasp@gmail.com

Summary of Supporting information

Table S1. List of primers used in this work

Primer name	Sequence to amplify	Primer orientation	Restriction enzyme 5′	[2]Primer sequence 5′→3′
attB_hasB_F	*attL* (*hasB*)	Forward	NA	CCAGTTCCCTGAAATTATTCCCCT
attP_hasB_R	*attL* (*hasB*)	Reverse	NA	CAGCTTCTTGAATTGCACCATCAAT
attB_hasAp_F	*attL* (*hasAp*)	Forward	NA	CTCTTCTTAGGCATCCTTCTATC
attP_hasAp_R	*attL* (*hasAp*)	Reverse	NA	GCAGCACTAGCTTGGAAACC
hasB_F	*hasB*	Forward	HindIII	AAA**AAGCTT**AATGTTTCAGATCAAGAAGATTTGTTGTA
hasB_R	*hasB*	Reverse	SalI	AAA**GTCGAC**TTACACACGTTGCTTC
hasAp_F	*hasAp*	Forward	HindIII	AAGCTT**AAGCTT**ATGAATACCTTATCTC
hasAp_R	*hasAp*	Reverse	XhoI	AAAGTCGACGCCTTTACAATGTGATTGA
AOX-Integration_F[1]	*AOX* promoter	Forward	NA	CAGTTTTTGCCCTACTTGATC
AOX-Integration_R[1]	*AMO*[3] terminator	Reverse	NA	GTAGGAAGGCTGGATGTC

NA: Not applied

[1]These primers were designed to confirm the insert integration when the plasmid pHIPH4 is used for cloning.

[2]In bold, the recognizing sequence for the respective enzyme.

[3] *AMO*: Amino oxidase

Table S2. List of plasmids used in this work

Name	Description	Reference
pBSK_*hasB*	Cloning vector carrying the synthetic *hasB* gene encoding the UDP-glucose dehydrogenase from *Xenopus laevis*	This work
pBSK_*hasAp*	Cloning vector carrying the synthetic *hasAp* gene encoding the hyaluronic acid synthase from *Pasteurella multocida*	This work
pBSK_*hasAs*	Cloning vector carrying the synthetic *hasAs* gene encoding the hyaluronic acid synthase from *Streptococcus zooepidemicus*	This work
pGEM-T Easy	Commercial plasmid used in the cloning steps	Promega
pGEM_*hasB*	pGEM-T Easy plasmid carrying the *hasB* gene	This work
pHIPH4	*O. polymorpha* integrative plasmid containing the native promoter pAOX and the terminator Tamo	[1]
pHIPZ7	*O. polymorpha* integrative plasmid containing the native promoter pTEF1 and the terminator Tamo	[1]
pHIPZ18_eGFP_SKL	*O. polymorpha* integrative plasmid containing the native promoter pAHD1 and the terminator Tamo and carrying the gene encoding eGFP	[1]
pHIPH4_*hasB*	pHIPH4-derived plasmid carrying the *hasB* gene under control of the pAOX promoter	This work
pHIPH4_*hasAs*	pHIPH4-derived plasmid carrying the *hasAs* gene under control of the pAOX promoter	This work
pHIPH4_*ScSInt13*	Synthetic plasmid containing a genetic switch controlling the expression of *hasAp* and *hasB* and carrying a gene of a serine-type phage integrase-13, codon-optimized for *S. cerevisiae*	This work
pHIPZ7_*hasAp*	pHIPZ7-derived plasmid carrying the *hasAp* gene under control of the pTEF1 promoter	This work
pHIPZ18_*hasAp*	pHIPZ18-derived plasmid carrying the *hasAp* gene under control of the pADH1 promoter	This work
pHIPZ18_*hasB*	pHIPZ18-derived plasmid carrying the *hasB* gene under control of the pADH1 promoter	This work

Table S3. List of strains used in this work

Name	Description	Reference
E. coli DH10B	Bacteria strain for plasmid cloning	Invitrogen
O. polymorpha NCYC 495 *yku80*	Methylotrophic yeast with a NHEJ-deficient phenotype; referred to as Wild Type (WT) in this work	[1]
EMB100.1	WT strain transformed with the linearized pHIPH4_*hasB* plasmid. The *hasB* gene was integrated into the genome under control of the pAOX promoter. This strain was used as an intermediate for the construction of the EMB101 and EMB102 strains.	This work
EMB101	EMB100.1 transformed with the linearized pHIPZ18_*hasAp* plasmid. The *hasAp* gene was integrated into the genome under control of the pADH1 promoter. The *hasB* gene is regulated by the pAOX promoter.	This work
EMB102	EMB100.1 transformed with the linearized pHIPZ7_*hasAp* plasmid. The *hasAp* gene was integrated into the genome under control of the pTEF1 promoter. The *hasB* gene is regulated by the pAOX promoter.	This work
EMB103	WT strain transformed with the linearized pHIPH4_ScsInt13 containing a genetic switch controlling the expression of *hasAp* and *hasB* and carrying a gene for a serine-type phage integrase-13, codon-optimized for *S. cerevisiae* and under control of the pAOX promoter. The genes *hasB* and *hasAp* are regulated by the pGDP promoter from *S. cerevisiae* and the pADH1 promoter from *O. polymorpha*, respectively.	This work
EMB100.2	WT strain transformed with the linearized pHIPZ18_*hasB* plasmid. The *hasB* gene was integrated into the genome under control of the pAHD1 promoter. This strain was used as an intermediate for the construction of EMB104.	This work
EMB104	EMB100.2 transformed with the linearized pHIPH4_*hasAs* plasmid. The *hasAs* gene was integrated into the genome under control of the pAOX promoter. The *hasB* gene is regulated by the pADH1 promoter.	This work

Sequence S1. Nucleotide sequence of the *hasB* gene from *Xenopus laevis* (Genbank ID: MH728986). The start and stop codons of the ORF are shown in bold.

ATGTTTCAGATCAAGAAGATTTGTTGTATTGGTGCCGGTTACGTCGGTGGTCCAACCTGT
TCTGTCATTGCACAGATGTGCCCTGACATTAAGGTCACTGTTGTGGATGTGAACCAAGC
CAGGATCAATGCTTGGAATAGTGACACTTTGCCTATCTACGAACCAGGTTTGAAGGAAG
TCGTAGAGTCATGCAGGGGAAAGAATTTGTTCTACTCAACTGACATTGATGGTGCAATT
CAAGAAGCTGATTTGGTGTTCATCTCAGTCAACACTCCAACAAAAACTTACGGTATGGG
TAAGGGAAGGGCAGCCGACTTGAAATACATTGAGGCTTGCGCTAGAAGAATAGTACAG
AATAGTAACGGATACAAGATTGTTACAGAGAAATCTACTGTGCCAGTTAGAGCTGCTG
AATCAATAAGACGTATCTTCGATGCAAATACTAAACCAGATTTGAACTTGCAGGTATTG
AGTAACCCAGAGTTTTTGGCAGAGGGTACAGCCATTAAGGATTTGAAGAACCCTGATA
GAGTTTTGATAGGTGGTGACGAAACCCCTGAAGGTCAGAAAGCTGTTAGAGCTTTGTGT
GACGTATACGAACACTGGGTACCATCTGAGAAAATCATAACCACAAACACCTGGTCTT
CTGAGTTGAGTAAGTTAGCAGCCAACGCATTCTTAGCTCAAAGAATTTCTTCAATCAAC
TCAATTAGTGCCTTATGTGAAGCCACAGGAGCTGACGTTGAAGAGGTTGCCAGAGCTAT
TGGTATGGATCAAAGAATTGGTAACAAGTTTTTGAAGGCTTCAGTGGGATTTGGAGGTT
CATGTTTTCAGAAGGACGTTTTGAACTTGGTTTACTTGTGTGAGGTGTTAAACTTGCACG
AAGTGGCCAAGTACTGGCAACAAGTGATTGATATGAATGATTATCAAAGGAGACGTTT
TACAACTAGGATAATCGATTGTTTGTTTAACACCGTGACCGATAAGAAAATCGCATTGT
TAGGTTTCGCTTTCAAGAAGGATACAGGTGATACTAGAGAGAGTAGTTCAATCTATATC
TCTAAGTATTTGATGGATGAAGGTGCTAAGTTACATATCTACGATCCAAAGGTCCCACG
TGAGCAGATCATCACTGACTTGAGTCAACCTGGTGTTGCAGCTGACGACAGGGTTTCTC
AATTGGTCCACATAAGTACAGATTTGTACGAAGCCTGTGAGAATGCACACGCTATGGTC
ATTTGTACTGAATGGGATATGTTCAAGGAATTAGATTTCAATAGAATCCATAGGATGAT
GTTAAAGCCTGCTTTCATATTCGATGGTAGACGTGTTTTAGATGAATTGCATGGAGAATT
GCAAAACATTGGATTTCAGGTGGAAACCATCGGAAAGAAGGTAGCTTCAAAAAGAAT
ACCATTCACTCCAACTGCTGATATCCCTAAGTTCGGTTTACAGGACTTGCCACACAAGA
AGCAACGTGTG**TAA**

Sequence S2. Nucleotide sequence of the *hasAp* gene from *Pasteurella multocida* (Genbank ID: MH728990). The start and stop codons of the ORF are shown in bold.

ATGAATACCTTATCTCAAGCCATCAAGGCATACAATTCAAATGACTATCAATTGGCTTT
GAAATTGTTCGAAAAGTCAGCAGAAATCTACGGACGTAAGATAGTAGAGTTTCAGATT
ACTAAGTGCAAGGAGAAATTGTCTGCCCATCCAAGTGTCAATTCAGCTCATCCTTCAGT
CAACTCAGCACATTTGTCAGTAAACAAAGAGGAAAAGGTTAATGTGTGCGATTCTCCAT
TGGATATTGCTACCCAATTGTTGTTGTCAAATGTCAAGAAATTAGTGTTATCTGACAGTG
AGAAGAACACATTGAAGAATAAGTGGAAGTTGTTAACTGAAAAGAAAAGTGAGAATG
CTGAAGTTCGTGCTGTAGCATTGGTACCAAAGGATTTCCCAAAGGATTTGGTGTTGGCA
CCATTACCTGACCATGTGAATGACTTTACTTGGTACAAGAAGAGAAAGAAACGTTTAG
GTATCAAACCAGAACATCAACACGTGGGTTTGTCAATAATTGTCACCACTTTCAATCGT
CCTGCAATCTTAAGTATAACTTTGGCATGCTTAGTTAATCAAAAGACTCACTATCCATTC
GAGGTGATTGTCACAGATGATGGATCACAAGAAGATTTGTCTCCAATCATAAGACAAT
ATGAAAACAAATTGGATATCCGTTATGTCAGACAAAAGGACAATGGTTTCCAAGCTAG
TGCTGCTAGGAATATGGGTTTGAGATTAGCAAAGTATGATTTCATTGGTTTGTTGGATTG

CGATATGGCACCTAACCCATTATGGGTGCATTCATATGTCGCTGAATTGTTAGAAGATG
ATGATTTGACAATCATTGGACCAAGAAAGTACATTGATACACAACATATCGACCCAAA
GGACTTCTTAAACAATGCATCTTTGTTGGAATCATTGCCAGAAGTTAAGACCAATAACT
CAGTGGCCGCAAAAGGTGAAGGTACCGTTTCATTGGATTGGAGGTTGGAGCAATTCGA
AAAGACTGAAAACTTAAGATTGTCAGACTCTCCTTTTAGATTCTTCGCAGCTGGTAATGT
TGCTTTCGCCAAGAAGTGGTTGAACAAATCTGGATTCTTTGATGAAGAGTTCAACCATT
GGGGTGGTGAAGATGTTGAGTTTGGATATAGATTGTTTAGGTATGGTTCATTCTTCAAGA
CTATTGACGGTATCATGGCCTACCATCAAGAGCCACCTGGTAAGGAAAACGAAACAGA
TAGGGAAGCTGGAAAGAACATCACATTGGATATTATGAGGGAGAAGGTACCATATATT
TACAGGAAGTTGTTGCCTATCGAAGATTCACACATCAATAGAGTCCCTTTGGTTTCTATC
TATATCCCAGCTTACAACTGTGCCAATTATATTCAACGTTGTGTTGATTCTGCCTTGAAC
CAGACAGTTGTAGATTTGGAAGTCTGTATTTGCAATGATGGTTCTACAGATAATACTTTG
GAAGTTATCAACAAGTTGTACGGTAACAATCCAAGAGTCAGAATCATGAGTAAACCAA
ATGGTGGTATTGCTAGTGCTTCTAATGCAGCAGTGAGTTTTGCCAAAGGATATTACATA
GGTCAATTAGATTCAGATGACTATTTGGAGCCAGATGCCGTAGAGTTATGTTTGAAAGA
GTTCTTGAAAGACAAAACTTTGGCTTGTGTATATACAACAAACAGAAATGTCAATCCTG
ATGGTTCTTTGATAGCAAATGGTTACAACTGGCCAGAGTTTAGTAGGGAGAAGTTGACT
ACTGCAATGATTGCTCATCACTTCCGTATGTTCACTATCAGGGCATGGCATTTGACCGAT
GGTTTTAATGAGAAGATTGAGAATGCTGTGGACTACGATATGTTCTTGAAGTTGAGTGA
AGTTGGTAAGTTCAAGCACTTAAACAAAATCTGCTATAACAGGGTATTGCATGGTGATA
ATACAAGTATTAAGAAGTTGGGTATCCAAAAGAAGAACCATTTCGTGGTCGTCAACCA
GAGTTTGAACAGGCAAGGAATCACTTACTACAATTACGACGAGTTCGATGACTTAGAT
GAGTCTAGGAAATACATCTTTAACAAAACAGCTGAGTACCAGGAAGAAATTGACATCT
TAAAGGACATTAAGATCATACAAAACAAGGACGCTAAAATAGCAGTATCTATCTTCTA
CCCAAATACTTTGAATGGTTTGGTCAAGAAATTGAATAACATCATCGAGTACAACAAG
AACATATTCGTTATTGTCTTGCATGTGGACAAGAACCATTTGACCCCAGATATCAAGAA
AGAGATATTGGCTTTCTACCACAAGCATCAAGTGAATATTTTGTTGAATAACGATATCT
CATACTACACATCAAACCGTTTAATCAAGACCGAGGCACATTTATCAAACATTAATAA
GTTGTCACAGTTGAACTTGAATTGTGAATATATCATATTCGACAATCATGACTCTTTGTT
CGTGAAGAATGATTCTTATGCCTATATGAAGAAGTACGATGTTGGTATGAATTTCTCAG
CCTTAACTCATGATTGGATTGAAAAGATTAACGCACATCCACCATTCAAGAAGTTGATT
AAGACATACTTTAACGATAATGACTTGAAATCTATGAACGTTAAAGGAGCTAGTCAAG
GAATGTTTATGACATATGCATTGGCTCACGAATTGTTGACTATTATCAAAGAGGTTATCA
CTTCTTGCCAATCTATCGATTCTGTACCAGAATACAACACTGAGGACATATGGTTTCAAT
TTGCATTGTTGATCTTGGAAAAGAAAACTGGTCATGTCTTTAACAAGACAAGTACCTTG
ACATACATGCCTTGGGAGAGGAAGTTGCAATGGACCAATGAACAAATTGAATCAGCTA
AACGTGGAGAAAACATTCCAGTGAACAAGTTCATAATCAATTCAATCACATTG**TAA**

Sequence S3. Nucleotide sequence of the *hasAs* gene from *Streptococcus zooepidemicus* (Genbank ID: AF414053.1). The start and stop codons of the ORF are shown in bold.

ATGAGAACTTTGAAAAATTTAATTACAGTTGTTGCATTCTCAATTTTCTGGGTTTTATTA
ATTTATGTTAATGTTTATTTGTTCGGTGCTAAAGGTTCATTGTCAATTTATGGTTTCTTATT
AATTGCATATTTGTTAGTTAAAATGTCTTTGTCATTCTTCTATAAACCATTCAAAGGTAG
AGCAGGTCAATATAAAGTTGCTGCAATTATTCCATCATATAATGAAGATGCTGAATCTT

TATTGGAAACTTTGAAATCAGTTCAACAACAAACATATCCATTGGCAGAAATTTATGTT
GTTGATGATGGTTCAGCCGATGAAACAGGTATTAAAAGAATTGAAGATTATGTTAGAG
ATACTGGTGATTTATCATCTAATGTTATTGTTCATCGTAGCGAAAAAAATCAAGGTAAA
AGACATGCACAAGCATGGGCTTTCGAAAGAAGCGATGCAGATGTTTTCTTGACAGTTGA
TTCAGATACATATATTTATCCAGATGCTTTGGAAGAATTATTGAAAACTTTCAATGATCC
TACAGTTTTCGCAGCTACAGGTCATTTGAATGTTAGAAATAGACAAACTAATTTGTTAA
CAAGATTGACTGATATTAGATATGATAATGCATTCGGTGTTGAAAGAGCAGCACAATCA
GTTACTGGTAATATTTTAGTTTGTTCTGGTCCTTTGTCAGTTTATAGAAGAGAAGTTGTTG
TTCCTAATATTGATAGATATATTAATCAAACTTTCTTAGGTATTCCAGTTTCAATTGGTGA
TGATAGATGTTTAACTAATTATGCTACTGATTTGGGTAAAACAGTTTATCAATCTACAGC
TAAATGTATTACTGATGTTCCAGATAAAATGTCTACATATTTGAAACAACAAAATAGAT
GGAATAAATCATTCTTCAGAGAATCTATTATTTCAGTTAAAAAAATTATGAATAATCCA
TTCGTTGCTTTGTGGACTATTTTGGAAGTTTCTATGTTCATGATGTTAGTTTATTCAGTTGT
TGATTTCTTCGTTGATAATGTTAGAGAGTTCGATTGGTTAAGAGTTTTGGCATTCTTGGTT
ATTATTTTCATTGTTGCTTTGTGTAGAAATATTCATTATATGTTGAAACATCCTTTGTCTTT
CTTGTTATCTCCTTTCTATGGTGTTTTACATTTATTCGTTTTGCAACCATTGAAATTGTATT
CATTGTTCACTATTAGAAATGCTGATTGGGGTACTAGAAAAAAATTGTTG**TAA**

Sequence S4. Nucleotide sequence of the gene of serine integrase-13 (Int13), codon-optimized for *S. cerevisiae*. The start and stop codons of the ORF are shown in bold.

ATGGCGGTTGGGATTTACATCAGAGTCTCAACCCAAGAGCAGGCGAGTGAAGGGCACA
GTATTGAAAGCCAAAAAAAGAAACTGGCGTCTTATTGTGAAATCCAAGGATGGGATGA
CTACAGGTTCTACATCGAAGAGGGCATATCCGGGAAAAACACAAATAGACCGAAGCTT
AAGCTATTAATGGAACATATCGAAAAGGGAAAAATTAACATTTTATTGGTCTACAGGCT
GGATAGGTTGACTAGGTCTGTGATCGATTTACATAAGCTATTAAACTTTTTACAGGAAC
ATGGGTGCGCGTTTAAATCTGCTACAGAAACTTACGACACAACTACTGCAAACGGAAG
GATGAGTATGGGTATAGTGAGTCTTCTAGCCCAATGGGAGACAGAAAATATGAGTGAG
CGTATTAAACTAAACCTGGAACACAAAGTCTTGGTTGAGGGGGAAAGAGTAGGGGCGA
TTCCCTATGGATTCGACTTGTCAGATGATGAAAAGCTTGTGAAGAATGAAAAGTCTGCA
ATTTTATTGGACATGGTCGAAAGGGTGGAGAACGGCTGGTCCGTCAATAGGATCGTCAA
CTATCTTAATTTAACTAACAATGATCGTAACTGGTCACCTAATGGGGTGCTACGTTTGTT
AAGGAACCCTGCACTATATGGCGCTACAAGGTGGAATGATAAAATCGCAGAGAACAC
ACACGAGGGTATAATTAGCAAGGAACGTTTCAACCGTCTGCAGCAAATACTTGCAGAC
CGTAGCATCCATCACAGACGTGATGTGAAAGGAACATACATATTTCAAGGAGTTTTGA
GATGTCCGGTTTGTGATCAGACGCTGTCCGTTAATAGGTTTATTAAGAAGCGTAAGGAT
GGAACAGAGTACTGTGGTGTCCTTTATAGGTGTCAGCCATGTATTAAGCAAAACAAGTA
CAATTTAGCTATCGGCGAAGCTAGGTTCCTGAAGGCCCTTAACGAGTACATGTCTACGG
TGGAATTCCAGACAGTTGAAGACGAGGTGATACCCAAGAAAAGTGAGAGAGAAATGT
TGGAATCTCAGCTGCAACAGATCGCAAGAAAGAGGGAGAAATACCAAAAGGCATGGG
CGAGCGATTTAATGTCCGATGATGAATTTGAGAAACTTATGGTCGAGACCCGTGAAACT
TATGACGAATGCAAGCAAAAACTGGAGAGTTGCGAGGACCCTATTAAGATCGACGAG
ACATATTTGAAGGAAATAGTTTACATGTTTCATCAAACATTCAATGATTTAGAGTCCGA
GAAGCAAAAGGAGTTTATATCAAAATTTATAAGGACTATCCGTTACACCGTCAAAGAG
CAGCAACCTATCAGACCTGATAAGTCTAAGACAGGTAAGGGTAAACAGAAAGTGATA
ATTACGGAAGTGGAGTTTTACCAG**TAA**

Sequence S5. The nucleotide sequence of the pGEM_*hasB* plasmid. The sequence of the *hasB* gene is shown in yellow.

TATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAA
CCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTA
ATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGA
ATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCG
TGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCT
CGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCG
ATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTA
GTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTA
ATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTG
ATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAA
AAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTCCTGATGCGGTATTTTC
TCCTTACGCATCTGTGCGGTATTTCACACCGCATCAGGTGGCACTTTTCGGGGAAATGTG
CGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGA
CAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAAC
ATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCA
GAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTACA
TCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTT
CCAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCC
GGGCAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTC
ACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCT
GCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGAC
CGAAGGAGCTAACCGCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGT
TGGGAACCGGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCT
GTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGGCGAACTACTTACTCTAGCTTC
CCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGACCACTTCTGCGC
TCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGGAGCCGGTGAGCGTGGGTCT
CGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTA
CACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACAGATCGCTGAGATAGG
TGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTCATATATACTTTAGAT
TGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCT
CATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAA
AGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAA
AAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTT
CCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTTCTTCTAGTGTAGCC
GTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAA

TCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCA
AGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACAC
AGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATG
AGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAG
GGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTAT
AGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGG
GGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTG
CTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATT
ACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGT
CAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTG
GCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAG
CGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCCCAGGCTTTACACTTTAT
GCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCACACAGGAAACA
GCTATGACCATGATTACGCCAAGCTATTTAGGTGACACTATAGAATACTCAAGCTATGC
ATCCAACGCGTTGGGAGCTCTCCCATATGGTCGACCTGCAGGCGGCCGCGAATTCACTA
GTGATTAAAAAGCTTAATGTTTCAGATCAAGAAGATTTGTTGTATTGGTGCCGGTTACGT
CGGTGGTCCAACCTGTTCTGTCATTGCACAGATGTGCCCTGACATTAAGGTCACTGTTGT
GGATGTGAACCAAGCCAGGATCAATGCTTGGAATAGTGACACTTTGCCTATCTACGAAC
CAGGTTTGAAGGAAGTCGTAGAGTCATGCAGGGGAAAGAATTTGTTCTACTCAACTGA
CATTGATGGTGCAATTCAAGAAGCTGATTTGGTGTTCATCTCAGTCAACACTCCAACAA
AAACTTACGGTATGGGTAAGGGAAGGGCAGCCGACTTGAAATACATTGAGGCTTGCGC
TAGAAGAATAGTACAGAATAGTAACGGATACAAGATTGTTACAGAGAAATCTACTGTG
CCAGTTAGAGCTGCTGAATCAATAAGACGTATCTTCGATGCAAATACTAAACCAGATTT
GAACTTGCAGGTATTGAGTAACCCAGAGTTTTTGGCAGAGGGTACAGCCATTAAGGATT
TGAAGAACCCTGATAGAGTTTTGATAGGTGGTGACGAAACCCCTGAAGGTCAGAAAGC
TGTTAGAGCTTTGTGTGACGTATACGAACACTGGGTACCATCTGAGAAAATCATAACCA
CAAACACCTGGTCTTCTGAGTTGAGTAAGTTAGCAGCCAACGCATTCTTAGCTCAAAGA
ATTTCTTCAATCAACTCAATTAGTGCCTTATGTGAAGCCACAGGAGCTGACGTTGAAGA
GGTTGCCAGAGCTATTGGTATGGATCAAAGAATTGGTAACAAGTTTTTGAAGGCTTCAG
TGGGATTTGGAGGTTCATGTTTTCAGAAGGACGTTTTGAACTTGGTTTACTTGTGTGAGG
TGTTAAACTTGCACGAAGTGGCCAAGTACTGGCAACAAGTGATTGATATGAATGATTAT
CAAAGGAGACGTTTTACAACTAGGATAATCGATTGTTTGTTTAACACCGTGACCGATAA
GAAAATCGCATTGTTAGGTTTCGCTTTCAAGAAGGATACAGGTGATACTAGAGAGAGT
AGTTCAATCTATATCTCTAAGTATTTGATGGATGAAGGTGCTAAGTTACATATCTACGAT
CCAAAGGTCCCACGTGAGCAGATCATCACTGACTTGAGTCAACCTGGTGTTGCAGCTGA
CGACAGGGTTTCTCAATTGGTCCACATAAGTACAGATTTGTACGAAGCCTGTGAGAATG
CACACGCTATGGTCATTTGTACTGAATGGGATATGTTCAAGGAATTAGATTTCAATAGA

ATCCATAGGATGATGTTAAAGCCTGCTTTCATATTCGATGGTAGACGTGTTTTAGATGAA
TTGCATGGAGAATTGCAAAACATTGGATTTCAGGTGGAAACCATCGGAAAGAAGGTAG
CTTCAAAAAGAATACCATTCACTCCAACTGCTGATATCCCTAAGTTCGGTTTACAGGAC
TTGCCACACAAGAAGCAACGTGTGTAAGTCGACTTTAATCGAATTCCCGCGGCCGCCAT
GGCGGCCGGGAGCATGCGACGTCGGGCCCAATTCGCCC

Sequence S6. The nucleotide sequence of the pHIPZ7_*hasAp* plasmid. The sequence of the *hasAp* gene is shown in green.

AGGGGATATCCTCGAGACTTGCCTTTGAAGGCTCTTGTTGCGGTAAATAAGTATATAGG
ACACGACAATCTAGTAATCTCCACTATTGACGAGCTCGTCGAACTGCGAAAATAGGTTT
TCCATCTGGTCTGTAGGCATCAGCCCGGCGTCATCCTCCTGCGCAGGAGCAGCGGGCTC
AGGGCCGGCCTGGGCGGGCTGATCCAGAAAGTCGAGGTTCAGATCCCCCACACACCAT
AGCTTCAAAATGTTTCTACTCCTTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCG
CCGTACCACTTCAAAACACCCAAGCACAGCATACTAAATTTTCCCTCTTTCTTCCTCTAG
GGTGTCGTTAATTACCCGTACTAAAGGTTTGGAAAAGAAAAAAGAGACCGCCTCGTTTC
TTTTTCTTCGTCGAAAAAGGCAATAAAAATTTTTATCACGTTTCTTTTTCTTGAAATTTTTT
TTTTTAGTTTTTTTCTCTTTCAGTGACCTCCATTGATATTTAAGTTAATAAACGGTCTTCAA
TTTCTCAAGTTTCAGTTTCATTTTTCTTGTTCTATTACAACTTTTTTTACTTCTTGTTCATTA
GAAAGAAAGCATAGCAATCTAATCTAAGGGGCGGTGTTGACAATTAATCATCGGCATA
GTATATCGGCATAGTATAATACGACAAGGTGAGGAACTAAACCATGGCCAAGTTGACC
AGTGCCGTTCCGGTGCTCACCGCGCGCGACGTCGCCGGAGCGGTCGAGTTCTGGACCG
ACCGGCTCGGGTTCTCCCGGGACTTCGTGGAGGACGACTTCGCCGGTGTGGTCCGGGAC
GACGTGACCCTGTTCATCAGCGCGGTCCAGGACCAGGTGGTGCCGGACAACACCCTGG
CCTGGGTGTGGGTGCGCGGCCTGGACGAGCTGTACGCCGAGTGGTCGGAGGTCGTGTCC
ACGAACTTCCGGGACGCCTCCGGGCCGGCCATGACCGAGATCGGCGAGCAGCCGTGGG
GGCGGGAGTTCGCCCTGCGCGACCCGGCCGGCAACTGCGTGCACTTCGTGGCCGAGGA
GCAGGACTGACACGTCCGACGGCGGCCCACGGGTCCCAGGCCTCGGAGATCCGTCCCC
CTTTTCCTTTGTCGATATCATGTAATTAGTTATGTCACGCTTACATTCACGCCCTCCCCCC
ACATCCGCTCTAACCGAAAAGGAAGGAGTTAGACAACCTGAAGTCTAGGTCCCTATTT
ATTTTTTTATAGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTTCTTTTTTTTC
TGTACAGACGCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTCCGAGCTTGGCG
TAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAAC
ATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCA
CATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGC
ATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCT
TCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCA
CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACAT

GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTT
TTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGG
TGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCG
TGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGG
AAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCG
CTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCG
GTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCC
ACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT
GGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTG
GTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAA
GAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTA
AGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAA
AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAA
TGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCC
TGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGC
TGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAG
CCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGT
CTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAAC
GTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTC
AGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAG
CGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCAC
TCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTT
CTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGT
TGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGT
GCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGA
GATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCA
CCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAA
GGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTT
ATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAA
ATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGGGAAATTGTAAACGTTAAT
ATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCG
AAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGT
TCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGA
AAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTT
GGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGA
GCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGG

AGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCC
GCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCATTCGCCATTCAGGCTGCGCAA
CTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGG
GGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG
TAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGAGCTCCAC
CGCGGTGGCGGCCGCTCTAGCATGGAACCAAGACCCATGACGTTGTTTCTTGATGATCT
CTTTTCGTTTCTGCTTTAATGTTTTTATTTGCTGATCGATTGCCTCTACATCGTTTTTCTGC
ACATCGCTTGAATCGCAGTCGTTATTTGTTTCCACTTTTTCTTTCGGGTTTGACAACCTCC
TCGATTCGTTGACAGTATCAAACACATCGGTAAAAGAACTGCTACGTGAAGTCTTGTCA
GGCTCAGGTATCATTTCCTCTGACTTTTCTTCCTCCTTTTGGCTTTCATGAGAGAACTGCT
TCTCGGAATCACACACACTATCTAGTGTTCGGTCTGCATCTTCTGGCGTCGAATGCTCGG
GCTCAAAGCATAGAATGTCCAAGTTTGAGGGCTCTGTTTGGAACTGCTCCAATTCTTTGC
TCGTTGGTCTTTGGAATAAATTGTTGTTGAGTAGTTTTGATCGAACTGGAGAGTCCCTAT
AATTGCGATTTGTATGTGCTGCAAGAATATGTGTAGAAGCAGTATTATGGATACTTCTCA
AAGAGCTACTTCTCACCGATACCTCGTCCTTGAACGAGACCGTTTTTTTACCCAATTGCT
CCTCGTCCACTAGTTTGTTTGGTTTCATGATGGACTTCAGCATTGAGCTATGTAGATCAC
CTTCTGGGCCTCCTGGGTTGTCTGGCTCAACTTGCTCTTTGTCATCATGTACAAAGGATA
ATGGGATGGTGGTGACTTTTGGAAACTGTACAATGGGGTGAGTATTGAGTTGTAATGTC
CCCCAGACATTCAGAAGGTCAGGCTTGGGAGCAGGCTTTTTACTCAGCCTCAAACTACC
GTTTCTCCGCCTTCGTCTTCTTGACACAGTTGATACGCTGTTGGAAATGGAATGGTCGTC
AGTCTCGTTTAGCCCTAAAGACACCGTTGCCATAATTAGGGATAAATCTATTTAGGCTC
CTTATTTTTCACTCGGCGCGAAAAAAAATGGGCACACGGCGATCGACGAGAGGTCGAT
CGTCTCTACGTTCCAGCTACATGGCAGTTCTAAGACGGGAAGTAAGATGACACTAGTAG
ATGTTTTGCAAATTAGGATCACAGGCCTGCTCCAGAAAACTTTTTATCCTTCAGCGGAA
GGTCTGTCCAGGGGCACAAATCTCAAACAGAGCAAAGGCACCCTTACCCTCTTATCTCC
TTGATGAATTTTTTTTTTCTTTTAAATTCTTTTTTAGAAGGCCGGGTAACAACTTTAAGAA
CCTGTTTCAGCTGTCCCCAGGAAGCTCCATTGGCAGTTCCAGTACGTTGTAGATGTGTAT
AGTGTGCTGAAAAGCGTTGTCCCTGGTTTTTTCAAGCAAAATCTTCGTCTCGGAGCTGGA
TAGCCCACCAAGGTATTGTTCCTGTGCGTAATTTTTGGCACGCAGACGACTCGAATAAG
TTTGGCAATAAAAAAATTTTTTTCACTATATAAAGAGGAGACATTCCCACATGAGATTT
TTTCTGATCTTTTAATTAGTACATTCGTAGGATCCAAGCTTATGAATACCTTATCTCAAG
CCATCAAGGCATACAATTCAAATGACTATCAATTGGCTTTGAAATTGTTCGAAAAGTCA
GCAGAAATCTACGGACGTAAGATAGTAGAGTTTCAGATTACTAAGTGCAAGGAGAAAT
TGTCTGCCCATCCAAGTGTCAATTCAGCTCATCCTTCAGTCAACTCAGCACATTTGTCAG
TAAACAAAGAGGAAAAGGTTAATGTGTGCGATTCTCCATTGGATATTGCTACCCAATTG
TTGTTGTCAAATGTCAAGAAATTAGTGTTATCTGACAGTGAGAAGAACACATTGAAGAA
TAAGTGGAAGTTGTTAACTGAAAAGAAAAGTGAGAATGCTGAAGTTCGTGCTGTAGCA

TTGGTACCAAAGGATTTCCCAAAGGATTTGGTGTTGGCACCATTACCTGACCATGTGAA
TGACTTTACTTGGTACAAGAAGAGAAAGAAACGTTTAGGTATCAAACCAGAACATCAA
CACGTGGGTTTGTCAATAATTGTCACCACTTTCAATCGTCCTGCAATCTTAAGTATAACT
TTGGCATGCTTAGTTAATCAAAAGACTCACTATCCATTCGAGGTGATTGTCACAGATGA
TGGATCACAAGAAGATTTGTCTCCAATCATAAGACAATATGAAAACAAATTGGATATC
CGTTATGTCAGACAAAAGGACAATGGTTTCCAAGCTAGTGCTGCTAGGAATATGGGTTT
GAGATTAGCAAAGTATGATTTCATTGGTTTGTTGGATTGCGATATGGCACCTAACCCATT
ATGGGTGCATTCATATGTCGCTGAATTGTTAGAAGATGATGATTTGACAATCATTGGAC
CAAGAAAGTACATTGATACACAACATATCGACCCAAAGGACTTCTTAAACAATGCATC
TTTGTTGGAATCATTGCCAGAAGTTAAGACCAATAACTCAGTGGCCGCAAAAGGTGAA
GGTACCGTTTCATTGGATTGGAGGTTGGAGCAATTCGAAAAGACTGAAAACTTAAGATT
GTCAGACTCTCCTTTTAGATTCTTCGCAGCTGGTAATGTTGCTTTCGCCAAGAAGTGGTT
GAACAAATCTGGATTCTTTGATGAAGAGTTCAACCATTGGGGTGGTGAAGATGTTGAGT
TTGGATATAGATTGTTTAGGTATGGTTCATTCTTCAAGACTATTGACGGTATCATGGCCT
ACCATCAAGAGCCACCTGGTAAGGAAAACGAAACAGATAGGGAAGCTGGAAAGAAC
ATCACATTGGATATTATGAGGGAGAAGGTACCATATATTTACAGGAAGTTGTTGCCTAT
CGAAGATTCACACATCAATAGAGTCCCTTTGGTTTCTATCTATATCCCAGCTTACAACTG
TGCCAATTATATTCAACGTTGTGTTGATTCTGCCTTGAACCAGACAGTTGTAGATTTGGA
AGTCTGTATTTGCAATGATGGTTCTACAGATAATACTTTGGAAGTTATCAACAAGTTGTA
CGGTAACAATCCAAGAGTCAGAATCATGAGTAAACCAAATGGTGGTATTGCTAGTGCT
TCTAATGCAGCAGTGAGTTTTGCCAAAGGATATTACATAGGTCAATTAGATTCAGATGA
CTATTTGGAGCCAGATGCCGTAGAGTTATGTTTGAAAGAGTTCTTGAAAGACAAAACTT
TGGCTTGTGTATATACAACAAACAGAAATGTCAATCCTGATGGTTCTTTGATAGCAAAT
GGTTACAACTGGCCAGAGTTTAGTAGGGAGAAGTTGACTACTGCAATGATTGCTCATCA
CTTCCGTATGTTCACTATCAGGGCATGGCATTTGACCGATGGTTTTAATGAGAAGATTGA
GAATGCTGTGGACTACGATATGTTCTTGAAGTTGAGTGAAGTTGGTAAGTTCAAGCACT
TAAACAAAATCTGCTATAACAGGGTATTGCATGGTGATAATACAAGTATTAAGAAGTTG
GGTATCCAAAAGAAGAACCATTTCGTGGTCGTCAACCAGAGTTTGAACAGGCAAGGAA
TCACTTACTACAATTACGACGAGTTCGATGACTTAGATGAGTCTAGGAAATACATCTTT
AACAAAACAGCTGAGTACCAGGAAGAAATTGACATCTTAAAGGACATTAAGATCATAC
AAAACAAGGACGCTAAAATAGCAGTATCTATCTTCTACCCAAATACTTTGAATGGTTTG
GTCAAGAAATTGAATAACATCATCGAGTACAACAAGAACATATTCGTTATTGTCTTGCA
TGTGGACAAGAACCATTTGACCCCAGATATCAAGAAAGAGATATTGGCTTTCTACCAC
AAGCATCAAGTGAATATTTTGTTGAATAACGATATCTCATACTACACATCAAACCGTTT
AATCAAGACCGAGGCACATTTATCAAACATTAATAAGTTGTCACAGTTGAACTTGAATT
GTGAATATATCATATTCGACAATCATGACTCTTTGTTCGTGAAGAATGATTCTTATGCCT
ATATGAAGAAGTACGATGTTGGTATGAATTTCTCAGCCTTAACTCATGATTGGATTGAA

AAGATTAACGCACATCCACCATTCAAGAAGTTGATTAAGACATACTTTAACGATAATG
ACTTGAAATCTATGAACGTTAAAGGAGCTAGTCAAGGAATGTTTATGACATATGCATTG
GCTCACGAATTGTTGACTATTATCAAAGAGGTTATCACTTCTTGCCAATCTATCGATTCT
GTACCAGAATACAACACTGAGGACATATGGTTTCAATTTGCATTGTTGATCTTGGAAAA
GAAAACTGGTCATGTCTTTAACAAGACAAGTACCTTGACATACATGCCTTGGGAGAGG
AAGTTGCAATGGACCAATGAACAAATTGAATCAGCTAAACGTGGAGAAAACATTCCAG
TGAACAAGTTCATAATCAATTCAATCACATTGTAAAGGCCTAA

Sequence S7. The nucleotide sequence of the pHIPZ18_*hasAp* plasmid. The sequence of the *hasAp* gene is shown in green.

TCGAGACTTGCCTTTGAAGGCTCTTGTTGCGGTAAATAAGTATATAGGACACGACAATC
TAGTAATCTCCACTATTGACGAGCTCGTCGAACTGCGAAAATAGGTTTTCCATCTGGTCT
GTAGGCATCAGCCCGGCGTCATCCTCCTGCGCAGGAGCAGCGGGCTCAGGGCCGGCCT
GGGCGGGCTGATCCAGAAAGTCGAGGTTCAGATCCCCCACACACCATAGCTTCAAAAT
GTTTCTACTCCTTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTT
CAAAACACCCAAGCACAGCATACTAAATTTTCCCTCTTTCTTCCTCTAGGGTGTCGTTAA
TTACCCGTACTAAAGGTTTGGAAAAGAAAAAAGAGACCGCCTCGTTTCTTTTTCTTCGTC
GAAAAAGGCAATAAAAATTTTTATCACGTTTCTTTTTCTTGAAATTTTTTTTTTTAGTTTTT
TTCTCTTTCAGTGACCTCCATTGATATTTAAGTTAATAAACGGTCTTCAATTTCTCAAGTT
TCAGTTTCATTTTTCTTGTTCTATTACAACTTTTTTTACTTCTTGTTCATTAGAAAGAAAGC
ATAGCAATCTAATCTAAGGGGCGGTGTTGACAATTAATCATCGGCATAGTATATCGGCA
TAGTATAATACGACAAGGTGAGGAACTAAACCATGGCCAAGTTGACCAGTGCCGTTCC
GGTGCTCACCGCGCGCGACGTCGCCGGAGCGGTCGAGTTCTGGACCGACCGGCTCGGG
TTCTCCCGGGACTTCGTGGAGGACGACTTCGCCGGTGTGGTCCGGGACGACGTGACCCT
GTTCATCAGCGCGGTCCAGGACCAGGTGGTGCCGGACAACACCCTGGCCTGGGTGTGG
GTGCGCGGCCTGGACGAGCTGTACGCCGAGTGGTCGGAGGTCGTGTCCACGAACTTCC
GGGACGCCTCCGGGCCGGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGT
TCGCCCTGCGCGACCCGGCCGGCAACTGCGTGCACTTCGTGGCCGAGGAGCAGGACTG
ACACGTCCGACGGCGGCCCACGGGTCCCAGGCCTCGGAGATCCGTCCCCCTTTTCCTTT
GTCGATATCATGTAATTAGTTATGTCACGCTTACATTCACGCCCTCCCCCCACATCCGCT
CTAACCGAAAAGGAAGGAGTTAGACAACCTGAAGTCTAGGTCCCTATTTATTTTTTTAT
AGTTATGTTAGTATTAAGAACGTTATTTATATTTCAAATTTTTCTTTTTTTTCTGTACAGAC
GCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGG
TCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCC
GGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTG
CGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGA
ATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCT

CACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGG
CGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAA
AAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAG
GCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAAC
CCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCC
TGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGC
GCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCT
GGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATC
GTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCT
AACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTAC
CTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGT
GGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATC
CTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTT
TGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGT
TTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATC
AGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCC
GTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGAT
ACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGA
AGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTG
TTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCA
TTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTT
CCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTC
CTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTAT
GGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGG
TGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCC
CGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCAT
TGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTT
CGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTT
CTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACA
CGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTT
ATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTT
CCGCGCACATTTCCCCGAAAAGTGCCACCTGGGAAATTGTAAACGTTAATATTTTGTTA
AAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGC
AAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTG
GAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGT
CTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGA

GGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACG
GGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCG
CTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCT
TAATGCGCCGCTACAGGGCGCGTCGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGG
AAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGC
TGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGA
CGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGG
CGGCCGCCCCCTGCATTATTAATCACCACCCCGTCTACGATGACAGGCTCGCGACTGCA
GAGATGGCCTTTGTTACGGGCAATACATTTGTCACACGGCACCTCGAAGTTGCACTTAA
CTTTGCGCTTCTTGCAGCTGAGACATGTCTTGGAAATACGAGTAGTACGCCGTAGCTTG
GTTATGTCGAGCCTAGTAAACTCCGCAGGAGAATTCCTCTTCTCCGAAGGCACTCTCTC
CGCAGAAGCCATAAATTATTTGAAATCAGAAATGTGAGCGATATAAACACCCCTGCAC
GGGGCTGCACCCCGATTTTTTTCGGGTGATCGTATGATGCTACTATGAGCCCGCAGATA
ACATCTGACTTTACATTAAGGCAAAATTCTGGTGTAGGGTGTCCGCAGGCCGAGTAATG
CTGACCGGTACGCACCTTTTTCGTCCGTTGTGGTGTGCTGCTTCATGGTCAATTTTTTTTA
GCTCAGATGACATCGGAAGGTGTTCCTACAAAGCACACCTCCAATTCGCCGTAAAACCT
CGGAGTAAATCATTCTCCCTGTAGTGTAAACTAAGCTGATGAGGCGTTCGGCAGATCCA
ATTGGACACACCCCGTCATGGGGTGTGAAATACCCCGCCACGGCCATCCGGCAGCCCC
ACTCCACCACTTTGGGTTGAATATGGTCCAATTGTTATGTGCAAATTTTCGGACATTGGA
AATCACCCACTCGGAGTAAAGCAACTCGTAGTGACTGAAAAATAAGACGTCATCATTA
ATTAGGCTGCAACAGACATATAAATACGAGACACTTTTCCCTCTTCTTAGGCATCCTTCT
ATCAATCAATCAATCAATTTAAAAAGCTTATGAATACCTTATCTCAAGCCATCAAGGCA
TACAATTCAAATGACTATCAATTGGCTTTGAAATTGTTCGAAAAGTCAGCAGAAATCTA
CGGACGTAAGATAGTAGAGTTTCAGATTACTAAGTGCAAGGAGAAATTGTCTGCCCATC
CAAGTGTCAATTCAGCTCATCCTTCAGTCAACTCAGCACATTTGTCAGTAAACAAAGAG
GAAAAGGTTAATGTGTGCGATTCTCCATTGGATATTGCTACCCAATTGTTGTTGTCAAAT
GTCAAGAAATTAGTGTTATCTGACAGTGAGAAGAACACATTGAAGAATAAGTGGAAGT
TGTTAACTGAAAAGAAAAGTGAGAATGCTGAAGTTCGTGCTGTAGCATTGGTACCAAA
GGATTTCCCAAAGGATTTGGTGTTGGCACCATTACCTGACCATGTGAATGACTTTACTTG
GTACAAGAAGAGAAAGAAACGTTTAGGTATCAAACCAGAACATCAACACGTGGGTTTG
TCAATAATTGTCACCACTTTCAATCGTCCTGCAATCTTAAGTATAACTTTGGCATGCTTA
GTTAATCAAAAGACTCACTATCCATTCGAGGTGATTGTCACAGATGATGGATCACAAGA
AGATTTGTCTCCAATCATAAGACAATATGAAAACAAATTGGATATCCGTTATGTCAGAC
AAAAGGACAATGGTTTCCAAGCTAGTGCTGCTAGGAATATGGGTTTGAGATTAGCAAA
GTATGATTTCATTGGTTTGTTGGATTGCGATATGGCACCTAACCCATTATGGGTGCATTC
ATATGTCGCTGAATTGTTAGAAGATGATGATTTGACAATCATTGGACCAAGAAAGTACA
TTGATACACAACATATCGACCCAAAGGACTTCTTAAACAATGCATCTTTGTTGGAATCA

TTGCCAGAAGTTAAGACCAATAACTCAGTGGCCGCAAAAGGTGAAGGTACCGTTTCAT
TGGATTGGAGGTTGGAGCAATTCGAAAAGACTGAAAACTTAAGATTGTCAGACTCTCCT
TTTAGATTCTTCGCAGCTGGTAATGTTGCTTTCGCCAAGAAGTGGTTGAACAAATCTGGA
TTCTTTGATGAAGAGTTCAACCATTGGGGTGGTGAAGATGTTGAGTTTGGATATAGATTG
TTTAGGTATGGTTCATTCTTCAAGACTATTGACGGTATCATGGCCTACCATCAAGAGCCA
CCTGGTAAGGAAAACGAAACAGATAGGGAAGCTGGAAAGAACATCACATTGGATATT
ATGAGGGAGAAGGTACCATATATTTACAGGAAGTTGTTGCCTATCGAAGATTCACACAT
CAATAGAGTCCCTTTGGTTTCTATCTATATCCCAGCTTACAACTGTGCCAATTATATTCA
ACGTTGTGTTGATTCTGCCTTGAACCAGACAGTTGTAGATTTGGAAGTCTGTATTTGCAA
TGATGGTTCTACAGATAATACTTTGGAAGTTATCAACAAGTTGTACGGTAACAATCCAA
GAGTCAGAATCATGAGTAAACCAAATGGTGGTATTGCTAGTGCTTCTAATGCAGCAGTG
AGTTTTGCCAAAGGATATTACATAGGTCAATTAGATTCAGATGACTATTTGGAGCCAGA
TGCCGTAGAGTTATGTTTGAAAGAGTTCTTGAAAGACAAAACTTTGGCTTGTGTATATAC
AACAAACAGAAATGTCAATCCTGATGGTTCTTTGATAGCAAATGGTTACAACTGGCCAG
AGTTTAGTAGGGAGAAGTTGACTACTGCAATGATTGCTCATCACTTCCGTATGTTCACTA
TCAGGGCATGGCATTTGACCGATGGTTTTAATGAGAAGATTGAGAATGCTGTGGACTAC
GATATGTTCTTGAAGTTGAGTGAAGTTGGTAAGTTCAAGCACTTAAACAAAATCTGCTA
TAACAGGGTATTGCATGGTGATAATACAAGTATTAAGAAGTTGGGTATCCAAAAGAAG
AACCATTTCGTGGTCGTCAACCAGAGTTTGAACAGGCAAGGAATCACTTACTACAATTA
CGACGAGTTCGATGACTTAGATGAGTCTAGGAAATACATCTTTAACAAAACAGCTGAG
TACCAGGAAGAAATTGACATCTTAAAGGACATTAAGATCATACAAAACAAGGACGCTA
AAATAGCAGTATCTATCTTCTACCCAAATACTTTGAATGGTTTGGTCAAGAAATTGAAT
AACATCATCGAGTACAACAAGAACATATTCGTTATTGTCTTGCATGTGGACAAGAACCA
TTTGACCCCAGATATCAAGAAAGAGATATTGGCTTTCTACCACAAGCATCAAGTGAATA
TTTTGTTGAATAACGATATCTCATACTACACATCAAACCGTTTAATCAAGACCGAGGCA
CATTTATCAAACATTAATAAGTTGTCACAGTTGAACTTGAATTGTGAATATATCATATTC
GACAATCATGACTCTTTGTTCGTGAAGAATGATTCTTATGCCTATATGAAGAAGTACGA
TGTTGGTATGAATTTCTCAGCCTTAACTCATGATTGGATTGAAAAGATTAACGCACATCC
ACCATTCAAGAAGTTGATTAAGACATACTTTAACGATAATGACTTGAAATCTATGAACG
TTAAAGGAGCTAGTCAAGGAATGTTTATGACATATGCATTGGCTCACGAATTGTTGACT
ATTATCAAAGAGGTTATCACTTCTTGCCAATCTATCGATTCTGTACCAGAATACAACACT
GAGGACATATGGTTTCAATTTGCATTGTTGATCTTGGAAAAGAAAACTGGTCATGTCTTT
AACAAGACAAGTACCTTGACATACATGCCTTGGGAGAGGAAGTTGCAATGGACCAATG
AACAAATTGAATCAGCTAAACGTGGAGAAAACATTCCAGTGAACAAGTTCATAATCAA
TTCAATCACATTGTAAAGGCCTAAAGGGGATATCC

Sequence S8. The nucleotide sequence of the pHIPZ18_*hasB* plasmid. The sequence of the *hasB* gene is shown in yellow.

CTAGAGGATCGATCCCCGGGCCTGGACATCCAGCCTTCCTACGCCATGACCACCTCCGA
GGCTAAGAGGGCCGTGCACAAGGAGACCAAGGACAAAACCTCGAGACTTGCCTTTGA
AGGCTCTTGTTGCGGTAAATAAGTATATAGGACACGACAATCTAGTAATCTCCACTATT
GACGAGCTCGTCGAACTGCGAAAATAGGTTTTCCATCTGGTCTGTAGGCATCAGCCCGG
CGTCATCCTCCTGCGCAGGAGCAGCGGGCTCAGGGCCGGCCTGGGCGGGCTGATCCAG
AAAGTCGAGGTTCAGATCCCCCACACACCATAGCTTCAAAATGTTTCTACTCCTTTTTTA
CTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTACCACTTCAAAACACCCAAGCAC
AGCATACTAAATTTTCCCTCTTTCTTCCTCTAGGGTGTCGTTAATTACCCGTACTAAAGGT
TTGGAAAAGAAAAAAGAGACCGCCTCGTTTCTTTTTCTTCGTCGAAAAAGGCAATAAA
AATTTTTATCACGTTTCTTTTTCTTGAAATTTTTTTTTTTAGTTTTTTTCTCTTTCAGTGACCT
CCATTGATATTTAAGTTAATAAACGGTCTTCAATTTCTCAAGTTTCAGTTTCATTTTTCTT
GTTCTATTACAACTTTTTTTACTTCTTGTTCATTAGAAAGAAAGCATAGCAATCTAATCT
AAGGGGCGGTGTTGACAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGAC
AAGGTGAGGAACTAAACCATGGCCAAGTTGACCAGTGCCGTTCCGGTGCTCACCGCGC
GCGACGTCGCCGGAGCGGTCGAGTTCTGGACCGACCGGCTCGGGTTCTCCCGGGACTTC
GTGGAGGACGACTTCGCCGGTGTGGTCCGGGACGACGTGACCCTGTTCATCAGCGCGGT
CCAGGACCAGGTGGTGCCGGACAACACCCTGGCCTGGGTGTGGGTGCGCGGCCTGGAC
GAGCTGTACGCCGAGTGGTCGGAGGTCGTGTCCACGAACTTCCGGGACGCCTCCGGGC
CGGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGACCC
GGCCGGCAACTGCGTGCACTTCGTGGCCGAGGAGCAGGACTGACACGTCCGACGGCGG
CCCACGGGTCCCAGGCCTCGGAGATCCGTCCCCCTTTTCCTTTGTCGATATCATGTAATT
AGTTATGTCACGCTTACATTCACGCCCTCCCCCCACATCCGCTCTAACCGAAAAGGAAG
GAGTTAGACAACCTGAAGTCTAGGTCCCTATTTATTTTTTTATAGTTATGTTAGTATTAAG
AACGTTATTTATATTTCAAATTTTTCTTTTTTTTCTGTACAGACGCGGTACCCAGCTTTTGT
TCCCTTTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTG
TGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTA
AAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCC
GCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGG
GAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC
GGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA
CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCC
AGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGA
GCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAG
ATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCT
TACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACG

CTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAAC
CCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCG
GTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGA
GGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGA
AGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCA
GCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGT
CTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAA
AAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTA
TATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCA
GCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACG
ATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCT
CACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAA
GTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAG
TAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGG
TGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAG
TTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTG
TCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCT
CTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCA
TTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAA
TACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGC
GAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA
CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGG
AAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCAT
ACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATAC
ATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAA
AGTGCCACCTGGGAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGT
TAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAA
AGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAA
AGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACT
ACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATC
GGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGG
CGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTA
GCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCG
CGTCGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCC
TCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGT
AACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATAC

GACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCCCCCTGCATTATTA
ATCACCACCCCGTCTACGATGACAGGCTCGCGACTGCAGAGATGGCCTTTGTTACGGGC
AATACATTTGTCACACGGCACCTCGAAGTTGCACTTAACTTTGCGCTTCTTGCAGCTGAG
ACATGTCTTGGAAATACGAGTAGTACGCCGTAGCTTGGTTATGTCGAGCCTAGTAAACT
CCGCAGGAGAATTCCTCTTCTCCGAAGGCACTCTCTCCGCAGAAGCCATAAATTATTTG
AAATCAGAAATGTGAGCGATATAAACACCCCTGCACGGGGCTGCACCCCGATTTTTTTC
GGGTGATCGTATGATGCTACTATGAGCCCGCAGATAACATCTGACTTTACATTAAGGCA
AAATTCTGGTGTAGGGTGTCCGCAGGCCGAGTAATGCTGACCGGTACGCACCTTTTTCG
TCCGTTGTGGTGTGCTGCTTCATGGTCAATTTTTTTTAGCTCAGATGACATCGGAAGGTG
TTCCTACAAAGCACACCTCCAATTCGCCGTAAAACCTCGGAGTAAATCATTCTCCCTGT
AGTGTAAACTAAGCTGATGAGGCGTTCGGCAGATCCAATTGGACACACCCCGTCATGG
GGTGTGAAATACCCCGCCACGGCCATCCGGCAGCCCCACTCCACCACTTTGGGTTGAAT
ATGGTCCAATTGTTATGTGCAAATTTTCGGACATTGGAAATCACCCACTCGGAGTAAAG
CAACTCGTAGTGACTGAAAAATAAGACGTCATCATTAATTAGGCTGCAACAGACATAT
AAATACGAGACACTTTTCCCTCTTCTTAGGCATCCTTCTATCAATCAATCAATCAATTTA
AAAAGCTTAATGTTTCAGATCAAGAAGATTTGTTGTATTGGTGCCGGTTACGTCGGTGGT
CCAACCTGTTCTGTCATTGCACAGATGTGCCCTGACATTAAGGTCACTGTTGTGGATGTG
AACCAAGCCAGGATCAATGCTTGGAATAGTGACACTTTGCCTATCTACGAACCAGGTTT
GAAGGAAGTCGTAGAGTCATGCAGGGGAAAGAATTTGTTCTACTCAACTGACATTGAT
GGTGCAATTCAAGAAGCTGATTTGGTGTTCATCTCAGTCAACACTCCAACAAAAACTTA
CGGTATGGGTAAGGGAAGGGCAGCCGACTTGAAATACATTGAGGCTTGCGCTAGAAGA
ATAGTACAGAATAGTAACGGATACAAGATTGTTACAGAGAAATCTACTGTGCCAGTTA
GAGCTGCTGAATCAATAAGACGTATCTTCGATGCAAATACTAAACCAGATTTGAACTTG
CAGGTATTGAGTAACCCAGAGTTTTTGGCAGAGGGTACAGCCATTAAGGATTTGAAGA
ACCCTGATAGAGTTTTGATAGGTGGTGACGAAACCCCTGAAGGTCAGAAAGCTGTTAG
AGCTTTGTGTGACGTATACGAACACTGGGTACCATCTGAGAAAATCATAACCACAAAC
ACCTGGTCTTCTGAGTTGAGTAAGTTAGCAGCCAACGCATTCTTAGCTCAAAGAATTTCT
TCAATCAACTCAATTAGTGCCTTATGTGAAGCCACAGGAGCTGACGTTGAAGAGGTTGC
CAGAGCTATTGGTATGGATCAAAGAATTGGTAACAAGTTTTTGAAGGCTTCAGTGGGAT
TTGGAGGTTCATGTTTTCAGAAGGACGTTTTGAACTTGGTTTACTTGTGTGAGGTGTTAA
ACTTGCACGAAGTGGCCAAGTACTGGCAACAAGTGATTGATATGAATGATTATCAAAG
GAGACGTTTTACAACTAGGATAATCGATTGTTTGTTTAACACCGTGACCGATAAGAAAA
TCGCATTGTTAGGTTTCGCTTTCAAGAAGGATACAGGTGATACTAGAGAGAGTAGTTCA
ATCTATATCTCTAAGTATTTGATGGATGAAGGTGCTAAGTTACATATCTACGATCCAAA
GGTCCCACGTGAGCAGATCATCACTGACTTGAGTCAACCTGGTGTTGCAGCTGACGACA
GGGTTTCTCAATTGGTCCACATAAGTACAGATTTGTACGAAGCCTGTGAGAATGCACAC
GCTATGGTCATTTGTACTGAATGGGATATGTTCAAGGAATTAGATTTCAATAGAATCCA

TAGGATGATGTTAAAGCCTGCTTTCATATTCGATGGTAGACGTGTTTTAGATGAATTGCA
TGGAGAATTGCAAAACATTGGATTTCAGGTGGAAACCATCGGAAAGAAGGTAGCTTCA
AAAAGAATACCATTCACTCCAACTGCTGATATCCCTAAGTTCGGTTTACAGGACTTGCC
ACACAAGAAGCAACGTGTGTAAGTCGACT

Sequence S9. The nucleotide sequence of the pHIPH4_*hasB* plasmid. The sequence of the *hasB* gene is shown in yellow.

AAGCTTAATGTTTCAGATCAAGAAGATTTGTTGTATTGGTGCCGGTTACGTCGGTGGTCC
AACCTGTTCTGTCATTGCACAGATGTGCCCTGACATTAAGGTCACTGTTGTGGATGTGAA
CCAAGCCAGGATCAATGCTTGGAATAGTGACACTTTGCCTATCTACGAACCAGGTTTGA
AGGAAGTCGTAGAGTCATGCAGGGGAAAGAATTTGTTCTACTCAACTGACATTGATGGT
GCAATTCAAGAAGCTGATTTGGTGTTCATCTCAGTCAACACTCCAACAAAAACTTACGG
TATGGGTAAGGGAAGGGCAGCCGACTTGAAATACATTGAGGCTTGCGCTAGAAGAATA
GTACAGAATAGTAACGGATACAAGATTGTTACAGAGAAATCTACTGTGCCAGTTAGAG
CTGCTGAATCAATAAGACGTATCTTCGATGCAAATACTAAACCAGATTTGAACTTGCAG
GTATTGAGTAACCCAGAGTTTTTGGCAGAGGGTACAGCCATTAAGGATTTGAAGAACCC
TGATAGAGTTTTGATAGGTGGTGACGAAACCCCTGAAGGTCAGAAAGCTGTTAGAGCTT
TGTGTGACGTATACGAACACTGGGTACCATCTGAGAAAATCATAACCACAAACACCTG
GTCTTCTGAGTTGAGTAAGTTAGCAGCCAACGCATTCTTAGCTCAAAGAATTTCTTCAAT
CAACTCAATTAGTGCCTTATGTGAAGCCACAGGAGCTGACGTTGAAGAGGTTGCCAGA
GCTATTGGTATGGATCAAAGAATTGGTAACAAGTTTTTGAAGGCTTCAGTGGGATTTGG
AGGTTCATGTTTTCAGAAGGACGTTTTGAACTTGGTTTACTTGTGTGAGGTGTTAAACTT
GCACGAAGTGGCCAAGTACTGGCAACAAGTGATTGATATGAATGATTATCAAAGGAGA
CGTTTTACAACTAGGATAATCGATTGTTTGTTTAACACCGTGACCGATAAGAAAATCGC
ATTGTTAGGTTTCGCTTTCAAGAAGGATACAGGTGATACTAGAGAGAGTAGTTCAATCT
ATATCTCTAAGTATTTGATGGATGAAGGTGCTAAGTTACATATCTACGATCCAAAGGTC
CCACGTGAGCAGATCATCACTGACTTGAGTCAACCTGGTGTTGCAGCTGACGACAGGGT
TTCTCAATTGGTCCACATAAGTACAGATTTGTACGAAGCCTGTGAGAATGCACACGCTA
TGGTCATTTGTACTGAATGGGATATGTTCAAGGAATTAGATTTCAATAGAATCCATAGG
ATGATGTTAAAGCCTGCTTTCATATTCGATGGTAGACGTGTTTTAGATGAATTGCATGGA
GAATTGCAAAACATTGGATTTCAGGTGGAAACCATCGGAAAGAAGGTAGCTTCAAAAA
GAATACCATTCACTCCAACTGCTGATATCCCTAAGTTCGGTTTACAGGACTTGCCACAC
AAGAAGCAACGTGTGTAAGTCGACTCTAGAGGATCGATCCCCGGGCCTGGACATCCAG
CCTTCCTACGCCATGACCACCTCCGAGGCTAAGAGGGCCGTGCACAAGGAGACCAAGG
ACAAAACCTCGAGACTTGCCTTTGAAGGCTCTTGTTGCGGTAAATAAGTATATAGGACA
CGACAATCTAGTAATCTCCACTATTGACGAGCTCGTCGAACTGCGAAAATAGGTTTTCC
ATCTGGTCTGTAGGCATCAGCCCGGCGTCATCCTCCTGCGCAGGAGCAGCGGGCTCAGG

GCCGGCCTGGGCGGGCTGATCCAGAAAGTCGAGGTTCAGATCCCCCACACACCATAGC
TTCAAAATGTTTCTACTCCTTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCG
TACCACTTCAAAACACCCAAGCACAGCATACTAAATTTTCCCTCTTTCTTCCTCTAGGGT
GTCGTTAATTACCCGTACTAAAGGTTTGGAAAAGAAAAAAGAGACCGCCTCGTTTCTTT
TTCTTCGTCGAAAAAGGCAATAAAAATTTTTATCACGTTTCTTTTTCTTGAAATTTTTTTTT
TTAGTTTTTTTCTCTTTCAGTGACCTCCATTGATATTTAAGTTAATAAACGGTCTTCAATTT
CTCAAGTTTCAGTTTCATTTTTCTTGTTCTATTACAACTTTTTTTACTTCTTGTTCATTAGA
AAGAAAGCATAGCAATCTAATCTAAGGGGCGGTGTTGACAATTAATCATCGGCATAGT
ATATCGGCATAGTATAATACGACAAGGTGAGGAACTAAACCATGGGTAAAAAGCCTGA
ACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACC
TGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGT
GGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTAT
CGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAG
CGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGC
CTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCT
GCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTC
AATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGC
AAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGAT
GCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCA
ACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGAT
GTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTT
GTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCC
GCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGA
CGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCG
GAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGA
TGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGG
GCAAAGGAATAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCATTT
GTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGATTTATATTTT
TTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGTAATATCAT
GCGTCAATCGTATGTGAATGCTGGTCGCTATACTGCTGTCGATTCGATACTAACGCCGC
CATCCAGTGTCGAAAACGAGCTCGAATTCATCGATGATGTACCCAGCTTTTGTTCCCTTT
AGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT
GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTG
GGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCA
GTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGC
GGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTC
GGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATC

AGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC
GTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCAC
AAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG
GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGA
TACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG
TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACA
CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTA
GGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAG
TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTT
GATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATT
ACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGC
TCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCT
TCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGT
AAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT
CTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAG
GGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCC
AGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCA
ACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCG
CCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCG
TCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCC
CCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCAT
GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAAT
AGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCA
CATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTC
AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGAT
CTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAT
GCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTT
TTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAAT
GTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACC
TGGGAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGC
TCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGAC
CGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGG
ACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACC
ATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTA
AAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGG

AAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGC
TGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCA
TTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTAT
TACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGG
GTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTAT
AGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCCACGGCGATATCGGATCTCGACG
CGGAGAACGATCTCCTCGAGCTGCTCGCGGATCAGCTTGTGGCCCGGTAATGGAACCA
GGCCGACGCGACGCTCCTTGCGGACCACGGTGGCTGGCGAGCCCAGTTTGTGAACGAG
GTCGTTTAGAACGTCCTGCGCAAAGTCCAGTGTCAGATGAATGTCCTCCTCGGACCAAT
TCAGCATGTTCTCGAGCAGCCATCTGTCTTTGGAGTAGAAGCGTAATCTCTGCTCCTCGT
TACTGTACCGGAAGAGGTAGTTTGCCTCGCCGCCCATAATGAACAGGTTCTCTTTCTGGT
GGCCTGTGAGCAGCGGGGACGTCTGGACGGCGTCGATGAGGCCCTTGAGGCGCTCGTA
GTACTTGTTCCGTCGCTGTAGCCGGCCGCGGTGACGATACCCACATAGAGGTCCTTGGC
CATTAGTTTGATGAGGTGGGGCAGGATGGGCGACTCGGCATCGAAATTTTTGCCGTCGT
CGTACAGTGTGATGTCACCATCGAATGTAATGAGCTGCAGCTTGCGATCTCGGATGGTT
TTGGAATGGAAGAACCGCGACATCTCCAACAGCTGGGCCGTGTTGAGAATGAGCCGGA
CGTCGTTGAACGAGGGGGCCACAAGCCGGCGTTTGCTGATGGCGCGGCGCTCGTCCTCG
ATGTAGAAGGCCTTTTCCAGAGGCAGTCTCGTGAAGAAGCTGCCAACGCTCGGAACCA
GCTGCACGAGCCGAGACAATTCGGGGGTGCCGGCTTTGGTCATTTCAATGTTGTCGTCG
ATGAGGAGTTCGAGGTCGTGGAAGATTTCCGCGTAGCGGCGTTTTGCCTCAGAGTTTAC
CATGAGGTCGTCCACTGCAGAGATGCCGTTGCTCTTCACCGCGTACAGGACGAACGGC
GTGGCCAGCAGGCCCTTGATCCATTCTATGAGGCCATCTCGACGGTGTTCCTTGAGTGC
GTACTCCACTCTGTAGCGACTGGACATCTCGAGACTGGGCTTGCTGTGCTGGATGCACC
AATTAATTGTTGCCGCATGCATCCTTGCACCGCAAGTTTTTAAAACCCACTCGCTTTAGC
CGTCGCGTAAAACTTGTGAATCTGGCAACTGAGGGGGTTCTGCAGCCGCAACCGAACTT
TTCGCTTCGAGGACGCAGCTGGATGGTGTCATGTGAGGCTCTGTTTGCTGGCGTAGCCTA
CAACGTGACCTTGCCTAACCGGACGGCGCTACCCACTGCTGTCTGTGCCTGCTACCAGA
AAATCACCAGAGCAGCAGAGGGCCGATGTGGCAACTGGTGGGGTGTCGGACAGGCTGT
TTCTCCACAGTGCAAATGCGGGTGAACCGGCCAGAAAGTAAATTCTTATGCTACCGTGC
AGTGACTCCGACATCCCCAGTTTTTGCCCTACTTGATCACAGATGGGGTCAGCGCTGCC
GCTAAGTGTACCCAACCGTCCCCACACGGTCCATCTATAAATACTGCTGCCAGTGCACG
GTGGTGACATCAATCTAAAGTACAAAAAC

Sequence S10. The nucleotide sequence of the pHIPH4_*hasAs* plasmid. The sequence of the *hasAs* gene is shown in blue.

AAGCTTAAAAATGAGAACTTTGAAAAATTTAATTACAGTTGTTGCATTCTCAATTTTCTG
GGTTTTATTAATTTATGTTAATGTTTATTTGTTCGGTGCTAAAGGTTCATTGTCAATTTATG

GTTTCTTATTAATTGCATATTTGTTAGTTAAAATGTCTTTGTCATTCTTCTATAAACCATTC
AAAGGTAGAGCAGGTCAATATAAAGTTGCTGCAATTATTCCATCATATAATGAAGATG
CTGAATCTTTATTGGAAACTTTGAAATCAGTTCAACAACAAACATATCCATTGGCAGAA
ATTTATGTTGTTGATGATGGTTCAGCCGATGAAACAGGTATTAAAAGAATTGAAGATTA
TGTTAGAGATACTGGTGATTTATCATCTAATGTTATTGTTCATCGTAGCGAAAAAAATCA
AGGTAAAAGACATGCACAAGCATGGGCTTTCGAAAGAAGCGATGCAGATGTTTTCTTG
ACAGTTGATTCAGATACATATATTTATCCAGATGCTTTGGAAGAATTATTGAAAACTTTC
AATGATCCTACAGTTTTCGCAGCTACAGGTCATTTGAATGTTAGAAATAGACAAACTAA
TTTGTTAACAAGATTGACTGATATTAGATATGATAATGCATTCGGTGTTGAAAGAGCAG
CACAATCAGTTACTGGTAATATTTTAGTTTGTTCTGGTCCTTTGTCAGTTTATAGAAGAG
AAGTTGTTGTTCCTAATATTGATAGATATATTAATCAAACTTTCTTAGGTATTCCAGTTTC
AATTGGTGATGATAGATGTTTAACTAATTATGCTACTGATTTGGGTAAAACAGTTTATCA
ATCTACAGCTAAATGTATTACTGATGTTCCAGATAAAATGTCTACATATTTGAAACAAC
AAAATAGATGGAATAAATCATTCTTCAGAGAATCTATTATTTCAGTTAAAAAAATTATG
AATAATCCATTCGTTGCTTTGTGGACTATTTTGGAAGTTTCTATGTTCATGATGTTAGTTT
ATTCAGTTGTTGATTTCTTCGTTGATAATGTTAGAGAGTTCGATTGGTTAAGAGTTTTGGC
ATTCTTGGTTATTATTTTCATTGTTGCTTTGTGTAGAAATATTCATTATATGTTGAAACAT
CCTTTGTCTTTCTTGTTATCTCCTTTCTATGGTGTTTTACATTTATTCGTTTTGCAACCATTG
AAATTGTATTCATTGTTCACTATTAGAAATGCTGATTGGGGTACTAGAAAAAAATTGTTG
TAAAGGCCTGATATCGAATTCCTGCAGCCCGGGGGATCCACTAGTTCTAGAGGATCGAT
CCCCGGGCCTGGACATCCAGCCTTCCTACGCCATGACCACCTCCGAGGCTAAGAGGGC
CGTGCACAAGGAGACCAAGGACAAAACCTCGAGACTTGCCTTTGAAGGCTCTTGTTGC
GGTAAATAAGTATATAGGACACGACAATCTAGTAATCTCCACTATTGACGAGCTCGTCG
AACTGCGAAAATAGGTTTTCCATCTGGTCTGTAGGCATCAGCCCGGCGTCATCCTCCTG
CGCAGGAGCAGCGGGCTCAGGGCCGGCCTGGGCGGGCTGATCCAGAAAGTCGAGGTTC
AGATCCCCCACACACCATAGCTTCAAAATGTTTCTACTCCTTTTTTACTCTTCCAGATTTT
CTCGGACTCCGCGCATCGCCGTACCACTTCAAAACACCCAAGCACAGCATACTAAATTT
TCCCTCTTTCTTCCTCTAGGGTGTCGTTAATTACCCGTACTAAAGGTTTGGAAAAGAAAA
AAGAGACCGCCTCGTTTCTTTTTCTTCGTCGAAAAAGGCAATAAAAATTTTTATCACGTT
TCTTTTTCTTGAAATTTTTTTTTTTAGTTTTTTTCTCTTTCAGTGACCTCCATTGATATTTAA
GTTAATAAACGGTCTTCAATTTCTCAAGTTTCAGTTTCATTTTTCTTGTTCTATTACAACTT
TTTTTACTTCTTGTTCATTAGAAAGAAAGCATAGCAATCTAATCTAAGGGGCGGTGTTGA
CAATTAATCATCGGCATAGTATATCGGCATAGTATAATACGACAAGGTGAGGAACTAA
ACCATGGGTAAAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAA
AGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTC
AGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTT
CTACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGT

GCTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGG
GTGTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTCGCG
GAGGCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCAT
TCGGACCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCT
GATCCCCATGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGC
GCAGGCTCTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCG
TGCACGCGGATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATAACAGCGGT
CATTGACTGGAGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCT
TCTGGAGGCCGTGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCAT
CCGGAGCTTGCAGGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCA
ACTCTATCAGAGCTTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGAT
GCGACGCAATCGTCCGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAG
AAGCGCGGCCGTCTGGACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGA
CGCCCCAGCACTCGTCCGAGGGCAAAGGAATAATCAGTACTGACAATAAAAAGATTCT
TGTTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAA
ATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAA
GTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACTGCTG
TCGATTCGATACTAACGCCGCCATCCAGTGTCGAAAACGAGCTCGAATTCATCGATGAT
GTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCA
TAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGG
AAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGT
TGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATC
GGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCAC
TGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGG
TAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAG
GCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCT
CCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCG
ACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGT
TCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCT
TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGG
CTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCT
TGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGG
ATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTA
CGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG
GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT
TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGA
TCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTC

ATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAA
ATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTG
AGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCG
TGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCG
CGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGG
CCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCC
GGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCT
ACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAA
CGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGG
TCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGC
ACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTA
CTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGT
CAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAA
ACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGT
AACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGT
GAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAA
TGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTC
TCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGC
ACATTTCCCCGAAAAGTGCCACCTGGGAAATTGTAAACGTTAATATTTTGTTAAAATTC
GCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATC
CCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAA
GAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAG
GGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCG
TAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAG
CCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGC
GCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCG
CCGCTACAGGGCGCGTCGCGCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGC
GATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAG
GCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCC
AGTGAATTGTAATACGACTCACTATAGGGCGAATTGGAGCTCCACCGCGGTGGCGGCC
GCCACGGCGATATCGGATCTCGACGCGGAGAACGATCTCCTCGAGCTGCTCGCGGATC
AGCTTGTGGCCCGGTAATGGAACCAGGCCGACGCGACGCTCCTTGCGGACCACGGTGG
CTGGCGAGCCCAGTTTGTGAACGAGGTCGTTTAGAACGTCCTGCGCAAAGTCCAGTGTC
AGATGAATGTCCTCCTCGGACCAATTCAGCATGTTCTCGAGCAGCCATCTGTCTTTGGA
GTAGAAGCGTAATCTCTGCTCCTCGTTACTGTACCGGAAGAGGTAGTTTGCCTCGCCGC
CCATAATGAACAGGTTCTCTTTCTGGTGGCCTGTGAGCAGCGGGGACGTCTGGACGGCG
TCGATGAGGCCCTTGAGGCGCTCGTAGTACTTGTTCCGTCGCTGTAGCCGGCCGCGGTG

ACGATACCCACATAGAGGTCCTTGGCCATTAGTTTGATGAGGTGGGGCAGGATGGGCG
ACTCGGCATCGAAATTTTTGCCGTCGTCGTACAGTGTGATGTCACCATCGAATGTAATG
AGCTGCAGCTTGCGATCTCGGATGGTTTTGGAATGGAAGAACCGCGACATCTCCAACA
GCTGGGCCGTGTTGAGAATGAGCCGGACGTCGTTGAACGAGGGGGCCACAAGCCGGCG
TTTGCTGATGGCGCGGCGCTCGTCCTCGATGTAGAAGGCCTTTTCCAGAGGCAGTCTCGT
GAAGAAGCTGCCAACGCTCGGAACCAGCTGCACGAGCCGAGACAATTCGGGGGTGCC
GGCTTTGGTCATTTCAATGTTGTCGTCGATGAGGAGTTCGAGGTCGTGGAAGATTTCCGC
GTAGCGGCGTTTTGCCTCAGAGTTTACCATGAGGTCGTCCACTGCAGAGATGCCGTTGC
TCTTCACCGCGTACAGGACGAACGGCGTGGCCAGCAGGCCCTTGATCCATTCTATGAGG
CCATCTCGACGGTGTTCCTTGAGTGCGTACTCCACTCTGTAGCGACTGGACATCTCGAG
ACTGGGCTTGCTGTGCTGGATGCACCAATTAATTGTTGCCGCATGCATCCTTGCACCGCA
AGTTTTTAAAACCCACTCGCTTTAGCCGTCGCGTAAAACTTGTGAATCTGGCAACTGAG
GGGGTTCTGCAGCCGCAACCGAACTTTTCGCTTCGAGGACGCAGCTGGATGGTGTCATG
TGAGGCTCTGTTTGCTGGCGTAGCCTACAACGTGACCTTGCCTAACCGGACGGCGCTAC
CCACTGCTGTCTGTGCCTGCTACCAGAAAATCACCAGAGCAGCAGAGGGCCGATGTGG
CAACTGGTGGGGTGTCGGACAGGCTGTTTCTCCACAGTGCAAATGCGGGTGAACCGGC
CAGAAAGTAAATTCTTATGCTACCGTGCAGTGACTCCGACATCCCCAGTTTTTGCCCTAC
TTGATCACAGATGGGGTCAGCGCTGCCGCTAAGTGTACCCAACCGTCCCCACACGGTCC
ATCTATAAATACTGCTGCCAGTGCACGGTGGTGACATCAATCTAAAGTACAAAAAC

Sequence S11. The nucleotide sequence of the pHIPH4_*ScSInt13* plasmid. The sequence of the *hasB* gene is shown in yellow. The reverse sequence of the *hasAp* gene is shown in green. The sequence of the Int13 encoding gene is shown in red.

GGCCGCCCCCTGCATTATTAATCACCACCCCGTCTACGATGACAGGCTCGCGACTGCAG
AGATGGCCTTTGTTACGGGCAATACATTTGTCACACGGCACCTCGAAGTTGCACTTAAC
TTTGCGCTTCTTGCAGCTGAGACATGTCTTGGAAATACGAGTAGTACGCCGTAGCTTGGT
TATGTCGAGCCTAGTAAACTCCGCAGGAGAATTCCTCTTCTCCGAAGGCACTCTCTCCG
CAGAAGCCATAAATTATTTGAAATCAGAAATGTGAGCGATATAAACACCCCTGCACGG
GGCTGCACCCCGATTTTTTTCGGGTGATCGTATGATGCTACTATGAGCCCGCAGATAAC
ATCTGACTTTACATTAAGGCAAAATTCTGGTGTAGGGTGTCCGCAGGCCGAGTAATGCT
GACCGGTACGCACCTTTTTCGTCCGTTGTGGTGTGCTGCTTCATGGTCAATTTTTTTTAGC
TCAGATGACATCGGAAGGTGTTCCTACAAAGCACACCTCCAATTCGCCGTAAAACCTCG
GAGTAAATCATTCTCCCTGTAGTGTAAACTAAGCTGATGAGGCGTTCGGCAGATCCAAT
TGGACACACCCCGTCATGGGGTGTGAAATACCCCGCCACGGCCATCCGGCAGCCCCAC
TCCACCACTTTGGGTTGAATATGGTCCAATTGTTATGTGCAAATTTTCGGACATTGGAAA
TCACCCACTCGGAGTAAAGCAACTCGTAGTGACTGAAAAATAAGACGTCATCATTAAT
TAGGCTGCAACAGACATATAAATACGAGACACTTTTCCCTCTTCTTAGGCATCCTTCTAT
CAATCAATCAATCAATTTAAAAGCATACATTGTTGTTGTTTTTCCAGATCCAGTTGGTCC
TGTAAATATAAGCAATCCATGTGAGTTTACAATGTGATTGAATTGATTATGAACTTGTTC
ACTGGAATGTTTTCTCCACGTTTAGCTGATTCAATTTGTTCATTGGTCCATTGCAACTTCC
TCTCCCAAGGCATGTATGTCAAGGTACTTGTCTTGTTAAAGACATGACCAGTTTTCTTTT

CCAAGATCAACAATGCAAATTGAAACCATATGTCCTCAGTGTTGTATTCTGGTACAGAA
TCGATAGATTGGCAAGAAGTGATAACCTCTTTGATAATAGTCAACAATTCGTGAGCCAA
TGCATATGTCATAAACATTCCTTGACTAGCTCCTTTAACGTTCATAGATTTCAAGTCATT
ATCGTTAAAGTATGTCTTAATCAACTTCTTGAATGGTGGATGTGCGTTAATCTTTTCAAT
CCAATCATGAGTTAAGGCTGAGAAATTCATACCAACATCGTACTTCTTCATATAGGCAT
AAGAATCATTCTTCACGAACAAAGAGTCATGATTGTCGAATATGATATATTCACAATTC
AAGTTCAACTGTGACAACTTATTAATGTTTGATAAATGTGCCTCGGTCTTGATTAAACGG
TTTGATGTGTAGTATGAGATATCGTTATTCAACAAAATATTCACTTGATGCTTGTGGTAG
AAAGCCAATATCTCTTTCTTGATATCTGGGGTCAAATGGTTCTTGTCCACATGCAAGACA
ATAACGAATATGTTCTTGTTGTACTCGATGATGTTATTCAATTTCTTGACCAAACCATTC
AAAGTATTTGGGTAGAAGATAGATACTGCTATTTTAGCGTCCTTGTTTTGTATGATCTTA
ATGTCCTTTAAGATGTCAATTTCTTCCTGGTACTCAGCTGTTTTGTTAAAGATGTATTTCC
TAGACTCATCTAAGTCATCGAACTCGTCGTAATTGTAGTAAGTGATTCCTTGCCTGTTCA
AACTCTGGTTGACGACCACGAAATGGTTCTTCTTTTGGATACCCAACTTCTTAATACTTG
TATTATCACCATGCAATACCCTGTTATAGCAGATTTTGTTTAAGTGCTTGAACTTACCAA
CTTCACTCAACTTCAAGAACATATCGTAGTCCACAGCATTCTCAATCTTCTCATTAAAAC
CATCGGTCAAATGCCATGCCCTGATAGTGAACATACGGAAGTGATGAGCAATCATTGC
AGTAGTCAACTTCTCCCTACTAAACTCTGGCCAGTTGTAACCATTTGCTATCAAAGAAC
CATCAGGATTGACATTTCTGTTTGTTGTATATACACAAGCCAAAGTTTTGTCTTTCAAGA
ACTCTTTCAAACATAACTCTACGGCATCTGGCTCCAAATAGTCATCTGAATCTAATTGAC
CTATGTAATATCCTTTGGCAAAACTCACTGCTGCATTAGAAGCACTAGCAATACCACCA
TTTGGTTTACTCATGATTCTGACTCTTGGATTGTTACCGTACAACTTGTTGATAACTTCCA
AAGTATTATCTGTAGAACCATCATTGCAAATACAGACTTCCAAATCTACAACTGTCTGG
TTCAAGGCAGAATCAACACAACGTTGAATATAATTGGCACAGTTGTAAGCTGGGATAT
AGATAGAAACCAAAGGGACTCTATTGATGTGTGAATCTTCGATAGGCAACAACTTCCTG
TAAATATATGGTACCTTCTCCCTCATAATATCCAATGTGATGTTCTTTCCAGCTTCCCTAT
CTGTTTCGTTTTCCTTACCAGGTGGCTCTTGATGGTAGGCCATGATACCGTCAATAGTCTT
GAAGAATGAACCATACCTAAACAATCTATATCCAAACTCAACATCTTCACCACCCCAA
TGGTTGAACTCTTCATCAAAGAATCCAGATTTGTTCAACCACTTCTTGGCGAAAGCAAC
ATTACCAGCTGCGAAGAATCTAAAAGGAGAGTCTGACAATCTTAAGTTTTCAGTCTTTT
CGAATTGCTCCAACCTCCAATCCAATGAAACGGTACCTTCACCTTTTGCGGCCACTGAG
TTATTGGTCTTAACTTCTGGCAATGATTCCAACAAAGATGCATTGTTTAAGAAGTCCTTT
GGGTCGATATGTTGTGTATCAATGTACTTTCTTGGTCCAATGATTGTCAAATCATCATCTT
CTAACAATTCAGCGACATATGAATGCACCCATAATGGGTTAGGTGCCATATCGCAATCC
AACAAACCAATGAAATCATACTTTGCTAATCTCAAACCCATATTCCTAGCAGCACTAGC
TTGGAAACCATTGTCCTTTTGTCTGACATAACGGATATCCAATTTGTTTTCATATTGTCTT
ATGATTGGAGACAAATCTTCTTGTGATCCATCATCTGTGACAATCACCTCGAATGGATA
GTGAGTCTTTTGATTAACTAAGCATGCCAAAGTTATACTTAAGATTGCAGGACGATTGA
AAGTGGTGACAATTATTGACAAACCCACGTGTTGATGTTCTGGTTTGATACCTAAACGTT
TCTTTCTCTTCTTGTACCAAGTAAAGTCATTCACATGGTCAGGTAATGGTGCCAACACCA
AATCCTTTGGGAAATCCTTTGGTACCAATGCTACAGCACGAACTTCAGCATTCTCACTTT
TCTTTTCAGTTAACAACTTCCACTTATTCTTCAATGTGTTCTTCTCACTGTCAGATAACAC
TAATTTCTTGACATTTGACAACAACAATTGGGTAGCAATATCCAATGGAGAATCGCACA
CATTAACCTTTTCCTCTTTGTTTACTGACAAATGTGCTGAGTTGACTGAAGGATGAGCTG
AATTGACACTTGGATGGGCAGACAATTTCTCCTTGCACTTAGTAATCTGAAACTCTACTA

TCTTACGTCCGTAGATTTCTGCTGACTTTTCGAACAATTTCAAAGCCAATTGATAGTCAT
TTGAATTGTATGCCTTGATGGCTTGAGATAAGGTATTCATAAGCTTAAATATTCGGAGTT
GTATTTATGTTACTAAAACAACTGGTCAAGTTCTACAAATACAACCGTTATTGAAGCTT
GCATGCCTGCAGGTCGACTCTAGAGGATCGATCCCCGGGCCTGGACATCCAGCCTTCCT
ACGCCATGACCACCTCCGAGGCTAAGAGGGCCGTGCACAAGGAGACCAAGGACAAAA
CCTCGAGACTTGCCTTTGAAGGCTCTTGTTGCGGTAAATAAGTATATAGGACACGACAA
TCTAGTAATCTCCACTATTGACGAGCTCGTCGAACTGCGAAAATAGGTTTTCCATCTGGT
CTGTAGGCATCAGCCCGGCGTCATCCTCCTGCGCAGGAGCAGCGGGCTCAGGGCCGGC
CTGGGCGGGCTGATCCAGAAAGTCGAGGTTCAGCAAATTAAAGCCTTCGAGCGTCCCA
AAACCTTCTCAAGCAAGGTTTTCAGTATAATGTTACATGCGTACACGCGTCTGTACAGA
AAAAAAAGAAAAATTTGAAATATAAATAACGTTCTTAATACTAACATAACTATAAAAA
AATAAATAGGGACCTAGACTTCAGGTTGTCTAACTCCTTCCTTTTCGGTTAGAGCGGATG
TGGGGGGAGGGCGTGAATGTAAGCGTGACATAACTAATTACATGACAATAACGGTTGT
ATTTGTAGAACTTGACCAGTTGTTTTAGTAACATAAATACAACTCCGAATAAAAAATGT
TTCAGATCAAGAAGATTTGTTGTATTGGTGCCGGTTACGTCGGTGGTCCAACCTGTTCTG
TCATTGCACAGATGTGCCCTGACATTAAGGTCACTGTTGTGGATGTGAACCAAGCCAGG
ATCAATGCTTGGAATAGTGACACTTTGCCTATCTACGAACCAGGTTTGAAGGAAGTCGT
AGAGTCATGCAGGGGAAAGAATTTGTTCTACTCAACTGACATTGATGGTGCAATTCAAG
AAGCTGATTTGGTGTTCATCTCAGTCAACACTCCAACAAAAACTTACGGTATGGGTAAG
GGAAGGGCAGCCGACTTGAAATACATTGAGGCTTGCGCTAGAAGAATAGTACAGAATA
GTAACGGATACAAGATTGTTACAGAGAAATCTACTGTGCCAGTTAGAGCTGCTGAATC
AATAAGACGTATCTTCGATGCAAATACTAAACCAGATTTGAACTTGCAGGTATTGAGTA
ACCCAGAGTTTTTGGCAGAGGGTACAGCCATTAAGGATTTGAAGAACCCTGATAGAGTT
TTGATAGGTGGTGACGAAACCCCTGAAGGTCAGAAAGCTGTTAGAGCTTTGTGTGACGT
ATACGAACACTGGGTACCATCTGAGAAAATCATAACCACAAACACCTGGTCTTCTGAG
TTGAGTAAGTTAGCAGCCAACGCATTCTTAGCTCAAAGAATTTCTTCAATCAACTCAAT
TAGTGCCTTATGTGAAGCCACAGGAGCTGACGTTGAAGAGGTTGCCAGAGCTATTGGTA
TGGATCAAAGAATTGGTAACAAGTTTTTGAAGGCTTCAGTGGGATTTGGAGGTTCATGT
TTTCAGAAGGACGTTTTGAACTTGGTTTACTTGTGTGAGGTGTTAAACTTGCACGAAGTG
GCCAAGTACTGGCAACAAGTGATTGATATGAATGATTATCAAAGGAGACGTTTTACAA
CTAGGATAATCGATTGTTTGTTTAACACCGTGACCGATAAGAAAATCGCATTGTTAGGT
TTCGCTTTCAAGAAGGATACAGGTGATACTAGAGAGAGTAGTTCAATCTATATCTCTAA
GTATTTGATGGATGAAGGTGCTAAGTTACATATCTACGATCCAAAGGTCCCACGTGAGC
AGATCATCACTGACTTGAGTCAACCTGGTGTTGCAGCTGACGACAGGGTTTCTCAATTG
GTCCACATAAGTACAGATTTGTACGAAGCCTGTGAGAATGCACACGCTATGGTCATTTG
TACTGAATGGGATATGTTCAAGGAATTAGATTTCAATAGAATCCATAGGATGATGTTAA
AGCCTGCTTTCATATTCGATGGTAGACGTGTTTTAGATGAATTGCATGGAGAATTGCAA
AACATTGGATTTCAGGTGGAAACCATCGGAAAGAAGGTAGCTTCAAAAAGAATACCAT
TCACTCCAACTGCTGATATCCCTAAGTTCGGTTTACAGGACTTGCCACACAAGAAGCAA
CGTGTGTAAACTCACATGGATTGCTTATATTTACAGGACCAACTGGATCTGGAAAAACA
ACAACAATGTATGCGGATCCACTAGTTCTAGAATCCGTCGAAACTAAGTTCTGGTGTTT
TAAAACTAAAAAAAAGACTAACTATAAAAGTAGAATTTAAGAAGTTTAAGAAATAGAT
TTACAGAATTACAATCAATACCTACCGTCTTTATATACTTATTAGTCAAGTAGGGGAAT
AATTTCAGGGAACTGGTTTCAACCTTTTTTTTCAGCTTTTCCAAATCAGAGAGAGCAGAA
GGTAATAGAAGGTGTAAGAAAATGAGATAGATACATGCGTGGGTCAATTGCCTTGTGT

CATCATTTACTCCAGGCAGGTTGCATCACTCCATTGAGGTTGTGCCCGTTTTTTGCCTGTT
TGTGCCCCTGTTCTCTGTAGTTGCGCTAAGAGAATGGACCTATGAACTGATGGTTGGTGA
AGAAAACAATATTTTGGTGCTGGGATTCTTTTTTTTTCTGGATGCCAGCTTAAAAAGCGG
GCTCCATTATATTTAGTGGATGCCAGGAATAAACTGTTCACCCAGACACCTACGATGTT
ATATATTCTGTGTAACCCGCCCCCTATTTTGGGCATGTACGGGTTACAGCAGAATTAAA
AGGCTAATTTTTTGACTAAATAAAGTTAGGAAAATCACTACTATTAATTATTTACGTATT
CTTTGAAATGGCAGTATTGATAATGATAAACTCGAGGATCCCCCACACACCATAGCTTC
AAAATGTTTCTACTCCTTTTTTACTCTTCCAGATTTTCTCGGACTCCGCGCATCGCCGTAC
CACTTCAAAACACCCAAGCACAGCATACTAAATTTTCCCTCTTTCTTCCTCTAGGGTGTC
GTTAATTACCCGTACTAAAGGTTTGGAAAAGAAAAAAGAGACCGCCTCGTTTCTTTTTC
TTCGTCGAAAAAGGCAATAAAAATTTTTATCACGTTTCTTTTTCTTGAAATTTTTTTTTTT
AGTTTTTTTCTCTTTCAGTGACCTCCATTGATATTTAAGTTAATAAACGGTCTTCAATTTC
TCAAGTTTCAGTTTCATTTTTCTTGTTCTATTACAACTTTTTTTACTTCTTGTTCATTAGAA
AGAAAGCATAGCAATCTAATCTAAGGGGCGGTGTTGACAATTAATCATCGGCATAGTA
TATCGGCATAGTATAATACGACAAGGTGAGGAACTAAACCATGGGTAAAAAGCCTGAA
CTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAAAAGTTCGACAGCGTCTCCGACCT
GATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTCAGCTTCGATGTAGGAGGGCGTG
GATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTCTACAAAGATCGTTATGTTTATC
GGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTGCTTGACATTGGGGAATTCAGC
GAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGTGTCACGTTGCAAGACCTGCC
TGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTCGCGGAGGCCATGGATGCGATCGCT
GCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGACCGCAAGGAATCGGTC
AATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCCCATGTGTATCACTGGC
AAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCTCTCGATGAGCTGAT
GCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCGGATTTCGGCTCCA
ACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGGAGCGAGGCGAT
GTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCGTGGTTGGCTT
GTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCAGGATCGCC
GCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGCTTGGTTGA
CGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTCCGATCCG
GAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGGACCGA
TGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCGAGG
GCAAAGGAATAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCATTT
GTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGATTTATATTTT
TTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGTAATATCAT
GCGTCAATCGTATGTGAATGCTGGTCGCTATACTGCTGTCGATTCGATACTAACGCCGC
CATCCAGTGTCGAAAACGAGCTCGAATTCATCGATGATGTACCCAGCTTTTGTTCCCTTT
AGTGAGGGTTAATTCCGAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATT
GTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTG
GGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCA
GTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGC
GGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTC
GGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATC
AGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC
GTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCAC

AAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAG
GCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGA
TACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGG
TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTT
CAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACA
CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTA
GGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAG
TATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTT
GATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATT
ACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGC
TCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCT
TCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGT
AAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGT
CTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAG
GGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCC
AGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCA
ACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCG
CCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCG
TCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCC
CCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCAT
GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAAT
AGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCA
CATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTC
AAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGAT
CTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAAT
GCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTT
TTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAAT
GTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACC
TGGGAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGC
TCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGAC
CGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGG
ACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACC
ATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTA
AAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGG
AAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGC
TGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCGCGCCA
TTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTAT
TACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGG
GTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTAT
AGGGCGAATTGGAGCTCCACCGCGGTGGCGGCCGCCACGGCGATATCGGATCTCGACG
CGGAGAACGATCTCCTCGAGCTGCTCGCGGATCAGCTTGTGGCCCGGTAATGGAACCA
GGCCGACGCGACGCTCCTTGCGGACCACGGTGGCTGGCGAGCCCAGTTTGTGAACGAG
GTCGTTTAGAACGTCCTGCGCAAAGTCCAGTGTCAGATGAATGTCCTCCTCGGACCAAT
TCAGCATGTTCTCGAGCAGCCATCTGTCTTTGGAGTAGAAGCGTAATCTCTGCTCCTCGT

TACTGTACCGGAAGAGGTAGTTTGCCTCGCCGCCCATAATGAACAGGTTCTCTTTCTGGT
GGCCTGTGAGCAGCGGGGACGTCTGGACGGCGTCGATGAGGCCCTTGAGGCGCTCGTA
GTACTTGTTCCGTCGCTGTAGCCGGCCGCGGTGACGATACCCACATAGAGGTCCTTGGC
CATTAGTTTGATGAGGTGGGGCAGGATGGGCGACTCGGCATCGAAATTTTTGCCGTCGT
CGTACAGTGTGATGTCACCATCGAATGTAATGAGCTGCAGCTTGCGATCTCGGATGGTT
TTGGAATGGAAGAACCGCGACATCTCCAACAGCTGGGCCGTGTTGAGAATGAGCCGGA
CGTCGTTGAACGAGGGGGCCACAAGCCGGCGTTTGCTGATGGCGCGGCGCTCGTCCTCG
ATGTAGAAGGCCTTTTCCAGAGGCAGTCTCGTGAAGAAGCTGCCAACGCTCGGAACCA
GCTGCACGAGCCGAGACAATTCGGGGGTGCCGGCTTTGGTCATTTCAATGTTGTCGTCG
ATGAGGAGTTCGAGGTCGTGGAAGATTTCCGCGTAGCGGCGTTTTGCCTCAGAGTTTAC
CATGAGGTCGTCCACTGCAGAGATGCCGTTGCTCTTCACCGCGTACAGGACGAACGGC
GTGGCCAGCAGGCCCTTGATCCATTCTATGAGGCCATCTCGACGGTGTTCCTTGAGTGC
GTACTCCACTCTGTAGCGACTGGACATCTCGAGACTGGGCTTGCTGTGCTGGATGCACC
AATTAATTGTTGCCGCATGCATCCTTGCACCGCAAGTTTTTAAAACCCACTCGCTTTAGC
CGTCGCGTAAAACTTGTGAATCTGGCAACTGAGGGGGTTCTGCAGCCGCAACCGAACTT
TTCGCTTCGAGGACGCAGCTGGATGGTGTCATGTGAGGCTCTGTTTGCTGGCGTAGCCTA
CAACGTGACCTTGCCTAACCGGACGGCGCTACCCACTGCTGTCTGTGCCTGCTACCAGA
AAATCACCAGAGCAGCAGAGGGCCGATGTGGCAACTGGTGGGGTGTCGGACAGGCTGT
TTCTCCACAGTGCAAATGCGGGTGAACCGGCCAGAAAGTAAATTCTTATGCTACCGTGC
AGTGACTCCGACATCCCCAGTTTTTGCCCTACTTGATCACAGATGGGGTCAGCGCTGCC
GCTAAGTGTACCCAACCGTCCCCACACGGTCCATCTATAAATACTGCTGCCAGTGCACG
GTGGTGACATCAATCTAAAGTACAAAAACATGGCGGTTGGGATTTACATCAGAGTCTC
AACCCAAGAGCAGGCGAGTGAAGGGCACAGTATTGAAAGCCAAAAAAAGAAACTGGC
GTCTTATTGTGAAATCCAAGGATGGGATGACTACAGGTTCTACATCGAAGAGGGCATAT
CCGGGAAAAACACAAATAGACCGAAGCTTAAGCTATTAATGGAACATATCGAAAAGG
GAAAAATTAACATTTTATTGGTCTACAGGCTGGATAGGTTGACTAGGTCTGTGATCGATT
TACATAAGCTATTAAACTTTTTACAGGAACATGGGTGCGCGTTTAAATCTGCTACAGAA
ACTTACGACACAACTACTGCAAACGGAAGGATGAGTATGGGTATAGTGAGTCTTCTAG
CCCAATGGGAGACAGAAAATATGAGTGAGCGTATTAAACTAAACCTGGAACACAAAG
TCTTGGTTGAGGGGGAAAGAGTAGGGGCGATTCCCTATGGATTCGACTTGTCAGATGAT
GAAAAGCTTGTGAAGAATGAAAAGTCTGCAATTTTATTGGACATGGTCGAAAGGGTGG
AGAACGGCTGGTCCGTCAATAGGATCGTCAACTATCTTAATTTAACTAACAATGATCGT
AACTGGTCACCTAATGGGGTGCTACGTTTGTTAAGGAACCCTGCACTATATGGCGCTAC
AAGGTGGAATGATAAAATCGCAGAGAACACACACGAGGGTATAATTAGCAAGGAACG
TTTCAACCGTCTGCAGCAAATACTTGCAGACCGTAGCATCCATCACAGACGTGATGTGA
AAGGAACATACATATTTCAAGGAGTTTTGAGATGTCCGGTTTGTGATCAGACGCTGTCC
GTTAATAGGTTTATTAAGAAGCGTAAGGATGGAACAGAGTACTGTGGTGTCCTTTATAG
GTGTCAGCCATGTATTAAGCAAAACAAGTACAATTTAGCTATCGGCGAAGCTAGGTTCC
TGAAGGCCCTTAACGAGTACATGTCTACGGTGGAATTCCAGACAGTTGAAGACGAGGT
GATACCCAAGAAAAGTGAGAGAGAAATGTTGGAATCTCAGCTGCAACAGATCGCAAG
AAAGAGGGAGAAATACCAAAAGGCATGGGCGAGCGATTTAATGTCCGATGATGAATTT
GAGAAACTTATGGTCGAGACCCGTGAAACTTATGACGAATGCAAGCAAAAACTGGAG
AGTTGCGAGGACCCTATTAAGATCGACGAGACATATTTGAAGGAAATAGTTTACATGTT
TCATCAAACATTCAATGATTTAGAGTCCGAGAAGCAAAAGGAGTTTATATCAAAATTTA
TAAGGACTATCCGTTACACCGTCAAAGAGCAGCAACCTATCAGACCTGATAAGTCTAA

GACAGGTAAGGGTAAACAGAAAGTGATAATTACGGAAGTGGAGTTTTACCAGTAAAA
GCTTGCATGCCTGCAGGTCGACTCTAGAGGATCGATCCCCGGGCCTGGACATCCAGCCT
TCCTACGCCATGACCACCTCCGAGGCTAAGAGGGCCGTGCACAAGGAGACCAAGGAC
AAAACCTCGAGACTTGCCTTTGAAGGCTCTTGTTGCGGTAAATAAGTATATAGGACACG
ACAATCTAGTAATCTCCACTATTGACGAGCTCGTCGAACTGCGAAAATAGGTTTTCCAT
CTGGTCTGTAGGCATCAGCCCGGCGTCATCCTCCTGCGCAGGAGCAGCGGGCTCAGGG
CCGGCCTGGGCGGGCTGATCCAGAAAGTCGAGGTTCA

Figure S1. Confirmation of integration of both genes into the EMB101 strain by colony PCR of 12 colonies selected after transformation and plating on selective medium. The upper gel shows amplification of the *hasB* gene (1505 bp) using the primer pair hasB_F and hasB_R. The bottom gel shows amplification of the *hasAp* gene (2966 bp) using the primer pair hasAp_F and hasAp_R. The agarose content was 1% and the ladders used were 1kB Ladder Plus M1191/M1192 (M1; Sinapse Inc) 1 kb Ladder K9 (M2; Kasvi). The *O. polymorpha* NCYC495 *yku80* was used as a negative control (C-) for the PCR reactions. The pHIPH4_*hasB* and pHIPZ18_*hasAp* plasmids were used as positive controls (C+) for the PCR reactions of *hasB* and *hasAp*, respectively.

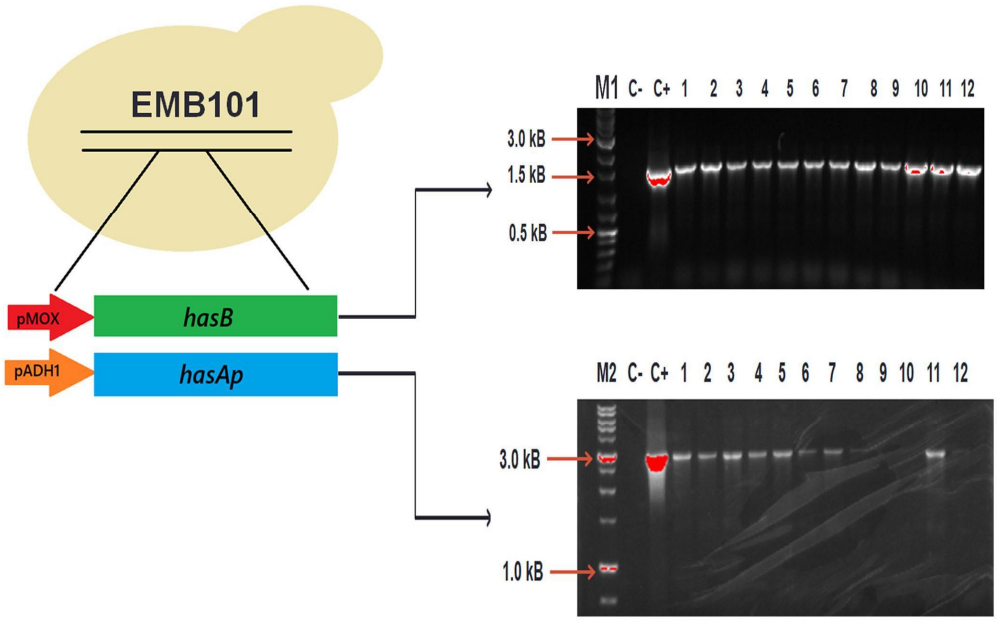


Figure S2. Colony PCRs to verify the *hasAp* (2966 bp) and *hasB* (1505 bp) genes stability in the genome of EMB101 strain after three successive passages on YPD supplemented with zeocin and hygromycin. Five colonies were selected after the transformation and plating on YPD plate containing both antibiotics for the verification of *hasB* and *hasAp* stability on the genome of *O. polymorpha* NCYC495 *yku80*. All PCR reactions for *hasB* gene were performed using the primer pair hasB_F and hasB_R while the PCRs reactions for *hasAp* gene using the primer pair hasAp_F and hasAp_R. In all PCRs reactions performed, the *O. polymorpha* NCYC495 *yku80* was used as a negative control (C-) and the pHIPH4_*hasB* and pHIPZ18_*hasAp* plasmids were used as positive controls (C+) for the of *hasB* and *hasAp*, respectively. The fragments amplified at the height of 1.5 kb in the upper gels correspond to the *hasB* gene and the fragments in the height of 3.0 kb at the bottom gels correspond to the *hasAp* gene. The black arrows indicate a passaging to another YPD plate supplemented with zeocin and hygromycin. The agarose content was 1% and the ladder used was 1 kb Plus DNA Ladder (M1; Invitrogen). **(A)** Colony PCR for the *hasAp* gene of the first passaging on selective medium, **(B)** Colony PCR for the hasB (upper gel) and hasP (bottom gel) genes of the second passaging on YPD plate with both antibiotics. **(C)** Colony PCR for the *hasB* (upper gel) and *hasAp* (bottom gel) genes of the third passaging on a YPD plate with both antibiotics.

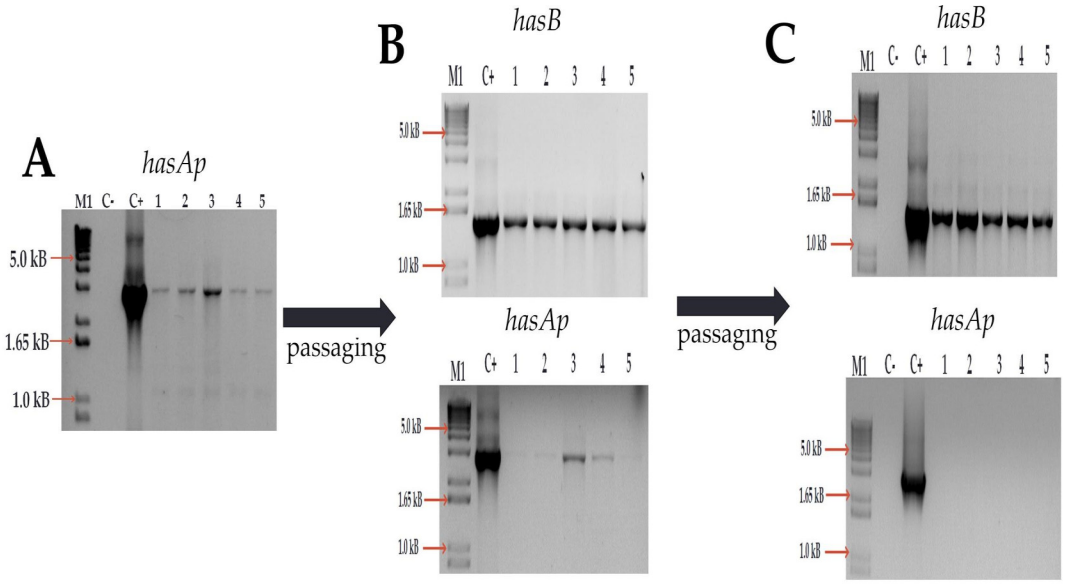


Figure S3. Confirmation of integration of both genes into the EMB102 strain by colony PCR of 12 colonies selected after the transformation and plating on selective medium. The upper gel shows amplification of the *hasB* gene (1.6 kB) using the primer pair hasB_F and hasB_R. The bottom gel shows amplification of the *hasAp* gene (2.9 kB) using the primer pair hasAp_F and hasAp_R. The agarose content was 1% and the ladders used were 1kB Ladder Plus M1191/M1192 (M1; Sinapse Inc) 1 kb Ladder K9 (M2; Kasvi). The *O. polymorpha* NCYC495 *yku80* was used as a negative control (C-) for the PCR reactions. The pHIPH4_*hasB* and pHIPZ7_*hasAp* plasmids were utilized as positive controls (C+) for the PCR reactions of *hasB* and *hasAp*, respectively.

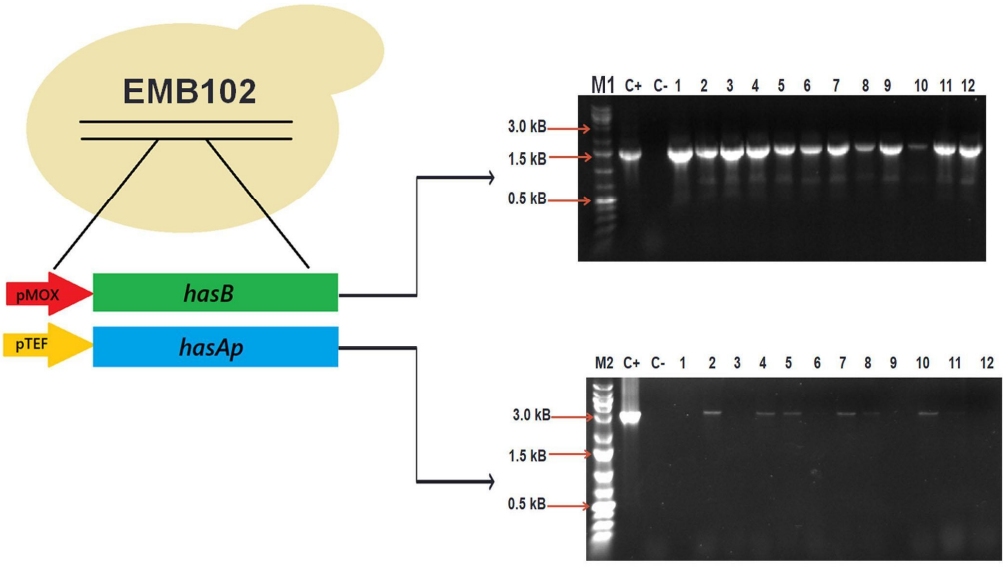


Figure S4. Scheme representing the genetic switch constructed to control the expression of both *hasB* and *hasAp* by a serine integrase. (A) The gene encoding the Int13 codon-optimized for *S. cerevisiae* is regulated by the promoter pAOX which is inducible by methanol (details in Figure S14). Thus, the addition of methanol leads to the production of Int13 that recognizes the sites *attB* and *attP* flanking both *has* genes synthesized in reverse complement orientation. The action of Int13 causes a rotation of 180° in both genes resulting in two different flanking sequences named *attL* and *attR*. In the final, both genes are in ORF with promoter and terminator and can be properly transcripted. The correct gene orientation as well the formation of *attL* sequence could be evaluated by PCR using the pair of primer attB_hasB_F and attP_hasB_R for *hasB* (resulting fragment of 493 bp) and attB_hasAp_F and attP_hasAp_R for *hasAp* (resulting fragment 827 bp). When the genes are in the initial orientation, the primer annealing fails and no amplification occurs. The black arrows represent the primer orientation. (B) Confirmation by PCR of the both genes rotation in the EMB103 strain after methanol induction. For all PCRs reactions, the genomic DNA was utilized as the template and for the extraction the phenol/chloroform method was applied according to [2]. The gel in the left shows the fragment (827 bp) containing the *attL* sequence and the *hasAp* flipped and in the right the fragment for *hasB* (493 bp).The *O. polymorpha* NCYC495 *yku80* was used as a negative control (C-) for the PCR reactions. The pHIPZ18_*hasAp* plasmid which contains the *hasAp* in ORF was utilized as positive controls (C+) for the PCR reactions of *hasAp*. The plasmid PKLAC2-BP constructed previously for our group (REF) was utilized as the control positive for the PCR reaction of *hasB* once all plasmids constructed in this work has the *hasB* gene controlled by endogenous promoters of *O. polymorpha* instead of pGDP promoter from *S. cerevisiae*.

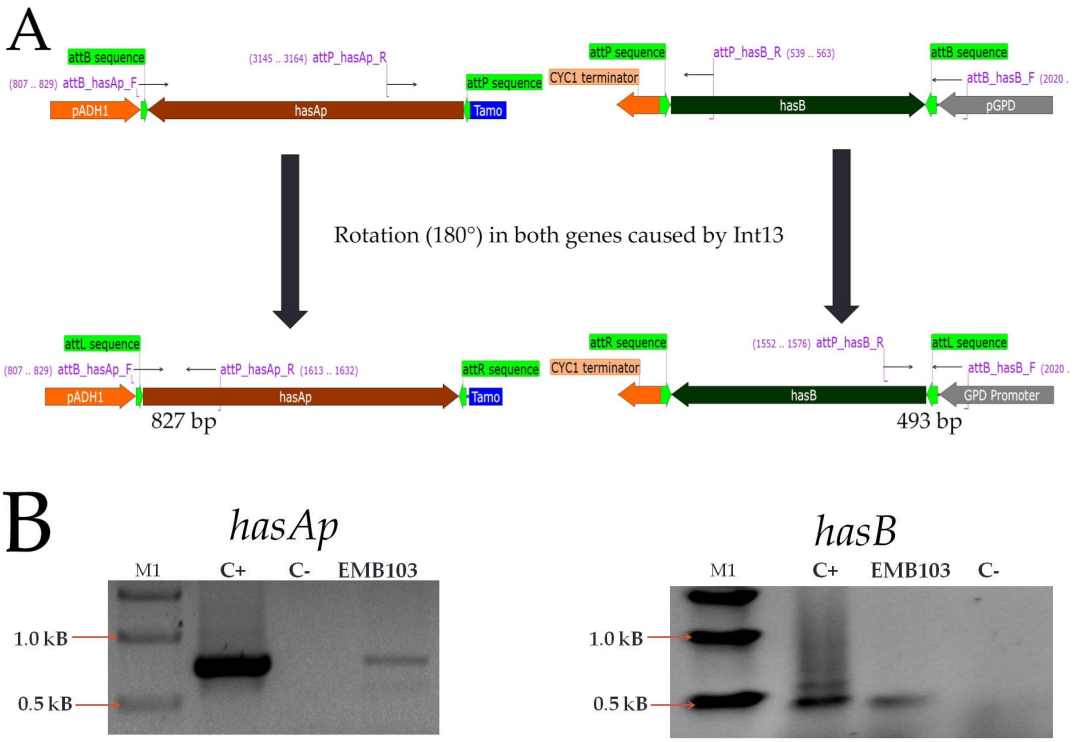


Figure S5. Confirmation of integration of both genes into the EMB104 strain by colony PCR of 12 colonies selected after the transformation and plating on selective medium. The upper gel shows amplification of the *hasB* gene (1.6 kB) using the pair of primers hasB_F and hasB_R. The bottom gel shows amplification of the *hasAs* gene between the pAOX (pMOX) promoter and the *AMO* terminator (1.5 kB) using the pair of primers AOX-Integration_F and AOX-Integration_R. The agarose content was 1% and the ladder used was 1kB Ladder M1181/M1182 (M1; Sinapse Inc). The *O. polymorpha* NCYC495 *yku80* was used as a negative control (C-) for the PCR reactions. The pHIPZ18_*hasB* and pHIPH4_*hasAs* plasmids were used as positive controls (C+) for the PCR reactions of *hasB* and *hasAs*, respectively.

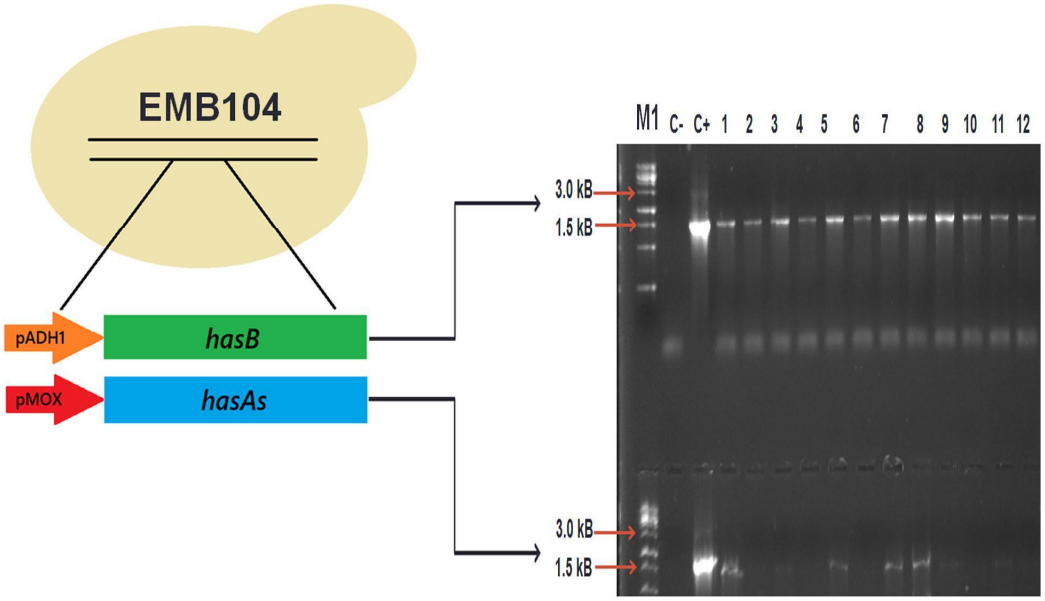


Figure S6. Map of the synthetic plasmid pBSK_*hasB* harbouring the *hasB* gene from *Xenopus laevis*. Oligonucleotide primers for amplification of *hasB* (HasB_F and HasB_R) are shown on their annealing sites.

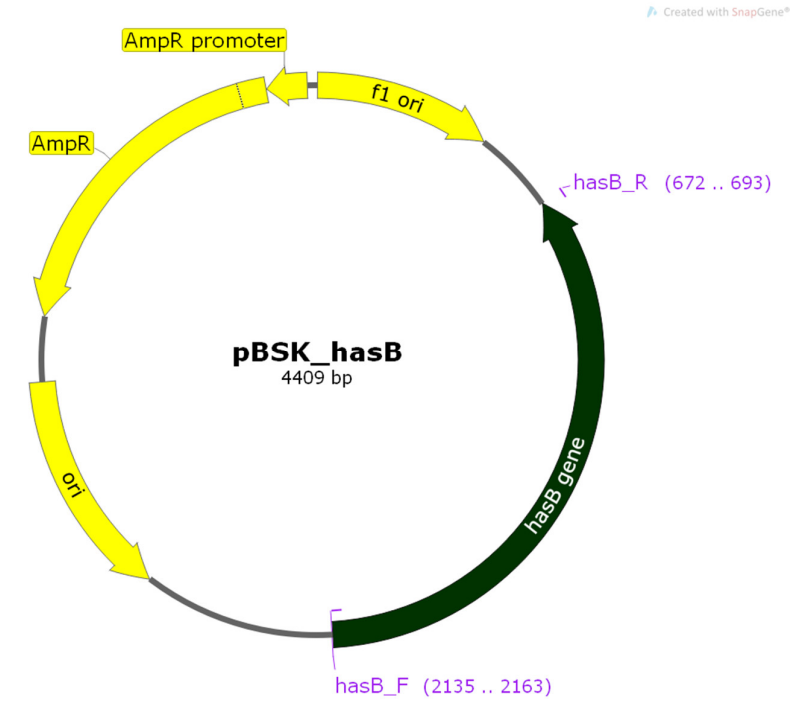


Figure S7. Map of the synthetic plasmid pBSK_*hasAp* harbouring the *hasA* gene from *Pasteurella multocida*. Primers for amplification of *hasAp* (HasAp_F and HasAp_R) are shown on their annealing sites.

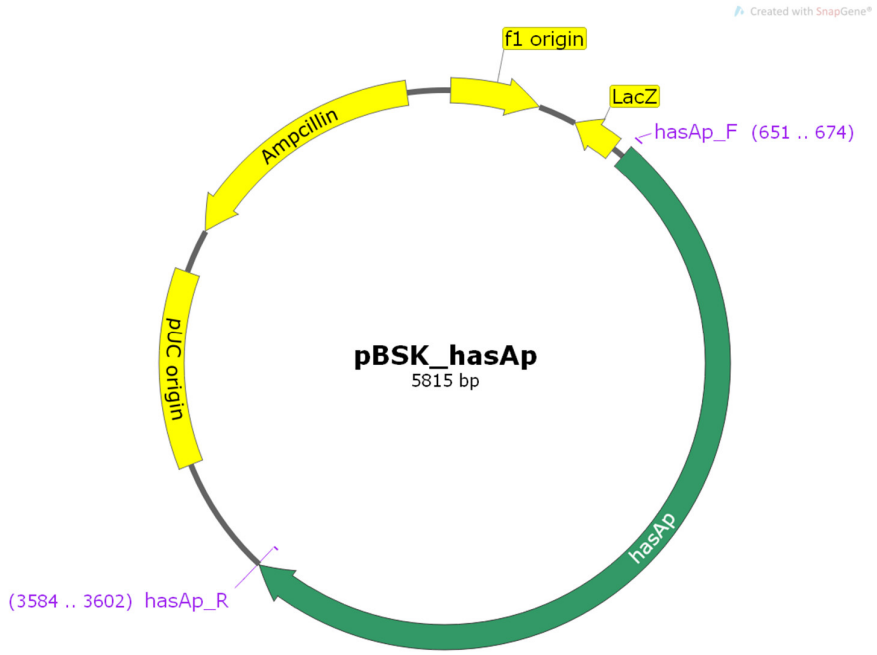


Figure S8. Map of the synthetic plasmid pBSK_*hasAs* harbouring the *hasA* gene from *Streptococcus zooepidemicus*. Restriction sites for HindIII and XbaI are shown.

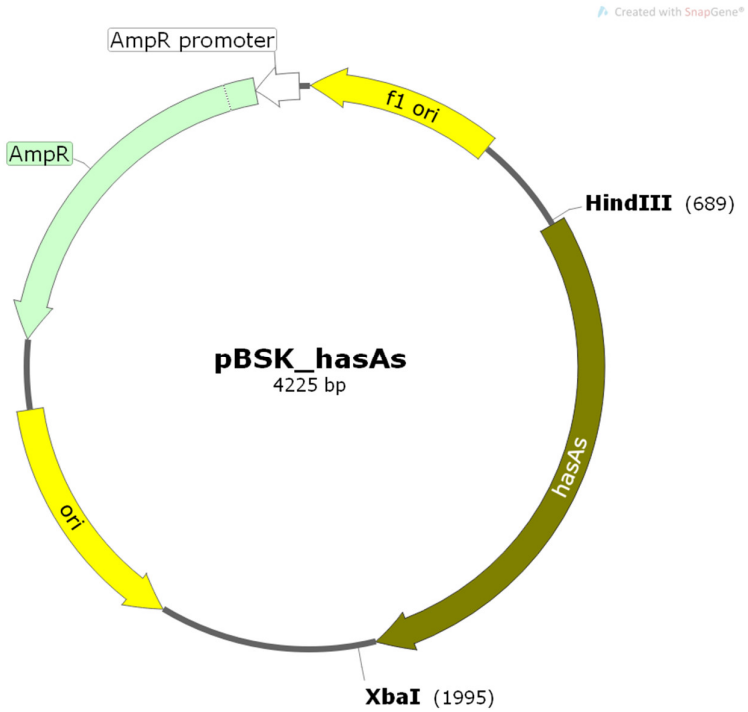


Figure S9. Map of the pGEM_*hasB* plasmid harbouring the *hasB* gene from *X. laevis*. Restriction sites for HindIII and SalI are shown.

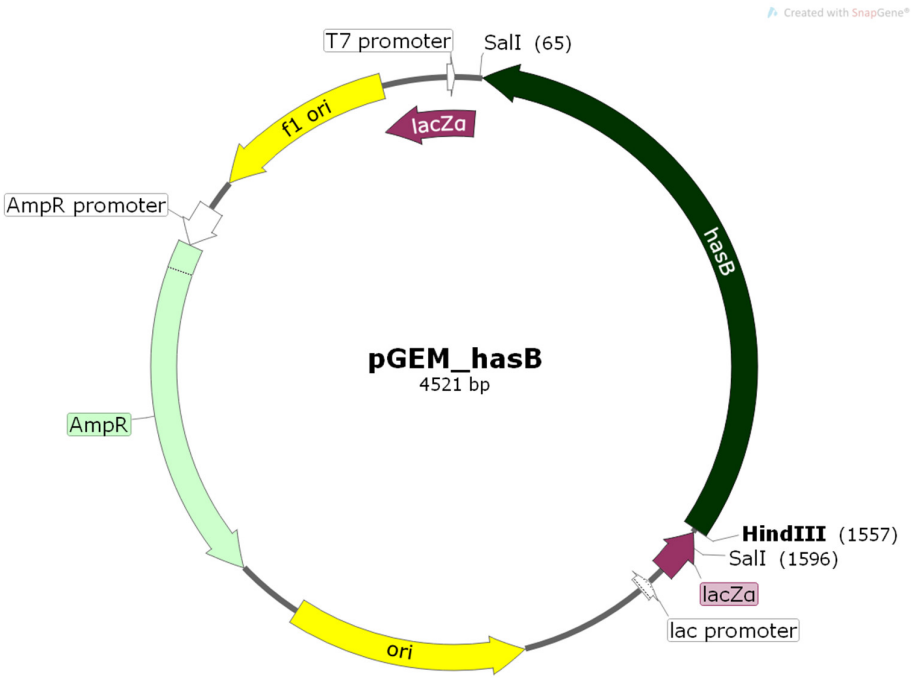


Figure S10. Map of the pHIPZ7_*hasAp* plasmid bearing the *hasAp* gene from *P. multocida*. Primers for amplification of *hasAp* (HasAp_F and HasAp_R) are shown on their annealing sites. Restriction sites for HindIII and XhoI are shown.

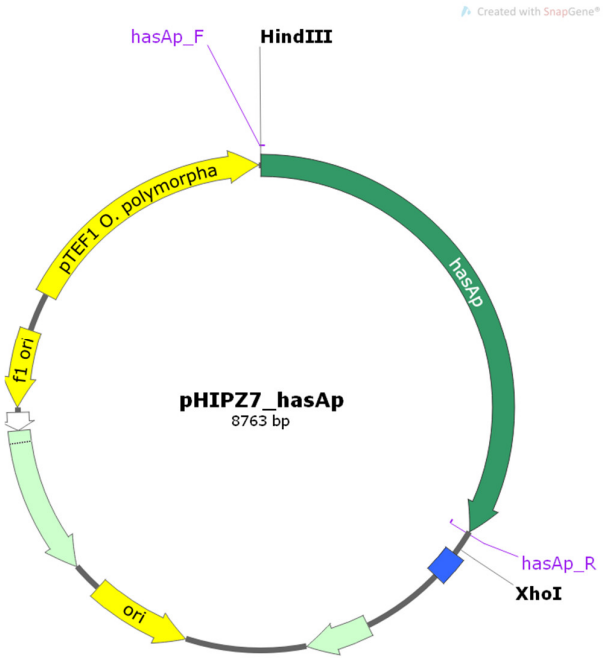


Figure S11. Map of the pHIPZ18_*hasAp* plasmid harbouring the *hasA* gene from *P. multocida*. Primers for amplification of *hasAp* (HasAp_F and HasAp_R) are shown on their annealing sites. Restriction sites for HindIII and XhoI are shown.

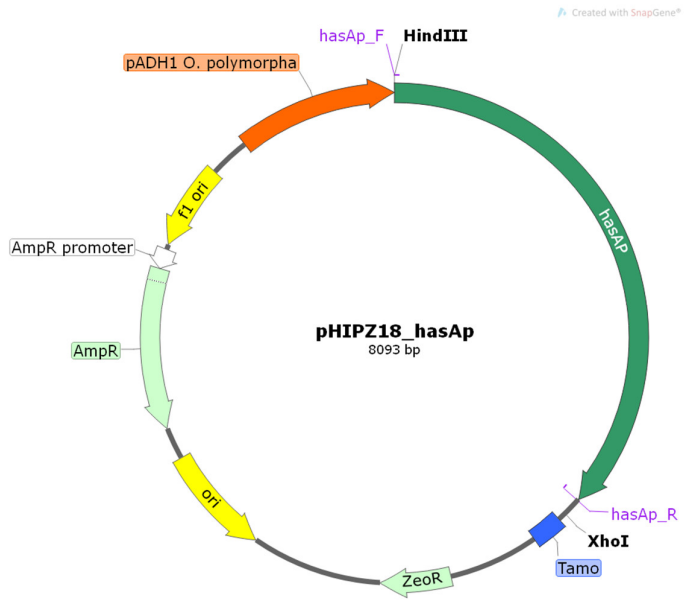


Figure S12. Map of the pHIPZ18_*hasB* plasmid bearing the *hasB* gene from *X. laevis*. Primers for amplification of *hasB* are shown on their annealing sites. Restriction sites for HindIII and SalI are shown.

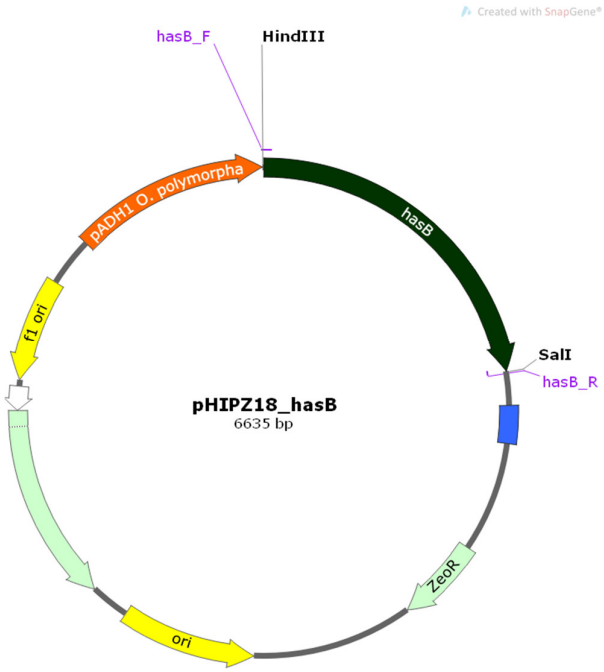


Figure S13. Map of the pHIPH4_*hasB* plasmid harbouring the *hasB* gene from *X. laevis*. Primers for amplification of *hasB* (HasB_F and HasB_R) are shown on their annealing sites. Restriction sites for HindIII and SalI are shown.

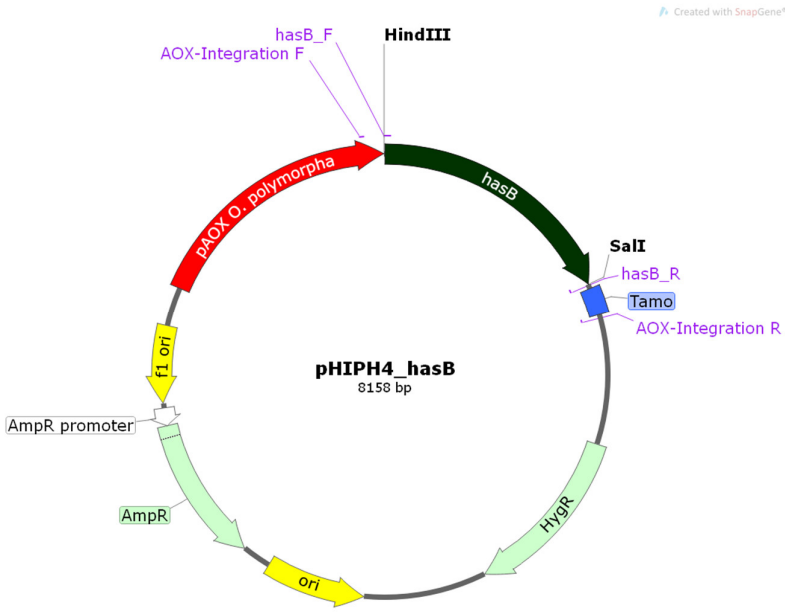


Figure S14. Map of the pHIPH4_*hasAs* plasmid harbouring the *hasA* gene from *S. zooepidemicus*. Primers for amplification of *hasA* (AOX-Integration F and AOX-Integration R) are shown on their annealing sites. Restriction sites for HindIII and XbaI are shown.

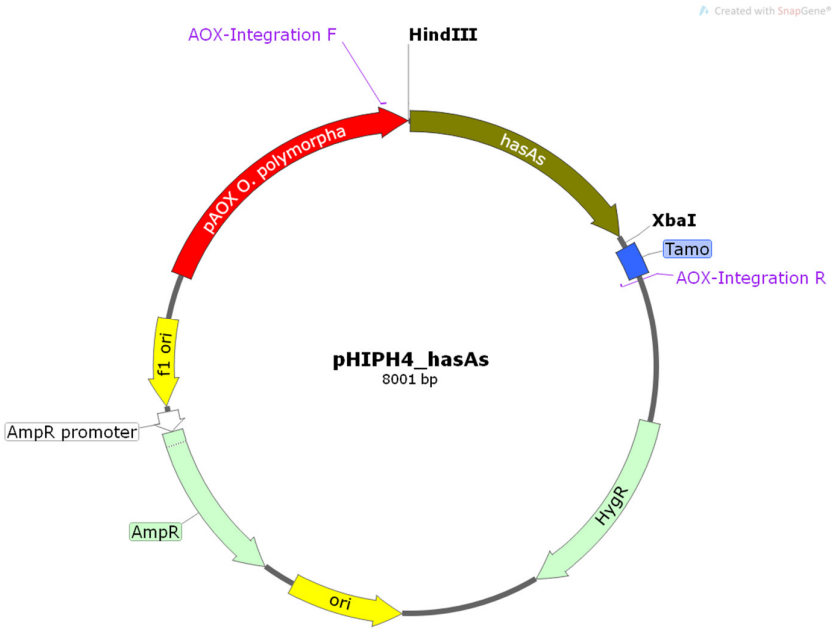


Figure S15. Map of the pHIPH4_*ScSInt13* plasmid harbouring the *hasB* gene from *X. laevis* and the *hasA* gene from *P. multocida*. Primers for amplification of *hasB* (HasB Forward and HasB Reverse) and *hasAp* (HasAp Forward and HasAp Reverse) are shown on their annealing sites.

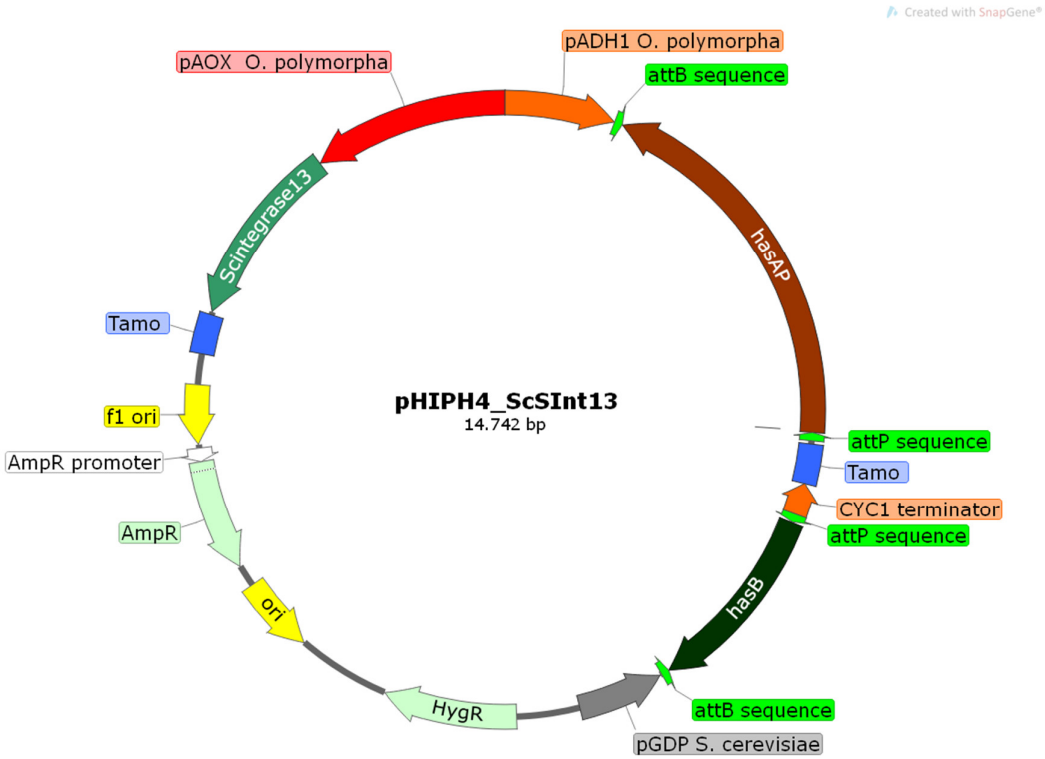


Figure S16. Standard curve obtained by the carbazole method for quantifying HA.

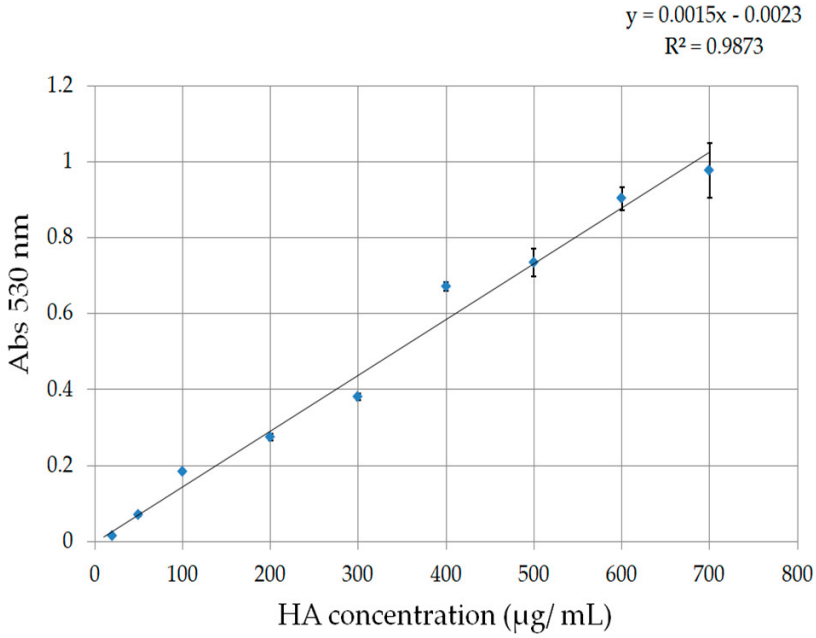
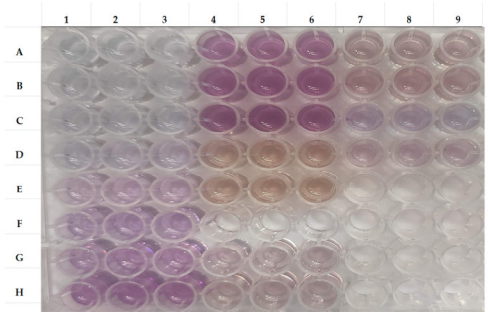


Figure S17. Carbazole assay plate picture. (A) The 96-well plate of the carbazole assay performed to quantify the HA in the supernatants after the 48 h of cultivation. (B) The plate design indicating the position of each sample in (A). I and II indicated the biological replicate. All quantifications were performed in technical replicate. The empty wells are represented with a line."

A



B

	1	2	3	4	5	6	7	8	9
A	Blank	Blank	Blank	500 µg/mL	500 µg/mL	500 µg/mL	EMB103 I	EMB103 I	EMB103 I
B	10 µg/mL	10 µg/mL	10 µg/mL	600 µg/mL	600 µg/mL	600 µg/mL	EMB103 I	EMB103 I	EMB103 I
C	20 µg/mL	20 µg/mL	20 µg/mL	700 µg/mL	700 µg/mL	700 µg/mL	EMB104 I	EMB104 I	EMB104 I
D	50 µg/mL	50 µg/mL	50 µg/mL	WT I	WT I	WT I	EMB104 II	EMB104 II	EMB104 II
E	100 µg/mL	100 µg/mL	100 µg/mL	WT II	WT II	WT II	-	-	-
F	200 µg/mL	200 µg/mL	200 µg/mL	-	-	-	-	-	-
G	300 µg/mL	300 µg/mL	300 µg/mL	EMB102 I	EMB102 I	EMB102 I	-	-	-
H	400 µg/mL	400 µg/mL	400 µg/mL	EMB102 II	EMB102 II	EMB102 II	-	-	-

References

1. van der Klei, I.J. The Hansenula polymorpha expression system Available online: https://www.rug.nl/research/molecular-cell-biology/research/the-hansenula-polymorpha-expression-system (accessed on May 19, 2020).

2. Sambrook, J.; Russell, D.W. *Molecular Cloning - A Laboratory Manual*; 3rd ed.; Cold Spring Harbor Laboratory Press, Cold Spring Harbor: New York, 2001;

3. V. Gomes, A.M.; C. M. Netto, J.H.; Carvalho, L.S.; Parachin, N.S. Heterologous Hyaluronic Acid Production in Kluyveromyces lactis. Microorganisms 2019, 7, 294, doi:10.3390/microorganisms7090294.

microbial biotechnology

Open Access

Metabolic engineering of *Pseudomonas putida* for production of vanillylamine from lignin-derived substrates

João Heitor Colombelli Manfrão-Netto,[1] Fredrik Lund,[2,†] Nina Muratovska,[2,†] Elin M. Larsson,[2,3,†] Nádia Skorupa Parachin[1,‡] and Magnus Carlquist[2,†]

[1]*Grupo Engenharia de Biocatalisadores, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília, Brazil.*

[2]*Division of Applied Microbiology, Department of Chemistry, Faculty of Engineering, Lund University, PO Box 124, Lund, 221 00, Sweden.*

[3]*Department of Bioengineering, California Institute of Technology, 1200 East California Blvd, Pasadena, CA 91125, USA.*

Summary

Whole-cell bioconversion of technical lignins using *Pseudomonas putida strains* overexpressing amine transaminases (ATAs) has the potential to become an eco-efficient route to produce phenolic amines. Here, a novel cell growth-based screening method to evaluate the *in vivo* activity of recombinant ATAs towards vanillylamine in *P. putida* KT2440 was developed. It allowed the identification of the native enzyme Pp-SpuC-II and ATA from *Chromobacterium violaceum* (*Cv-ATA*) as highly active towards vanillylamine *in vivo*. Overexpression of *Pp-SpuC-II* and *Cv-ATA* in the strain GN442ΔPP_2426, previously engineered for reduced vanillin assimilation, resulted in 94- and 92-fold increased specific transaminase activity, respectively. Whole-cell bioconversion of vanillin yielded 0.70 ± 0.20 mM and 0.92 ± 0.30 mM vanillylamine, for *Pp-SpuC-II* and *Cv-ATA*, respectively. Still, amine production was limited by a substantial re-assimilation of the product and formation of the by-products vanillic acid and vanillyl alcohol. Concomitant overexpression of *Cv-ATA* and alanine dehydrogenase from *Bacillus subtilis* increased the production of vanillylamine with ammonium as the only nitrogen source and a reduction in the amount of amine product re-assimilation. Identification and deletion of additional native genes encoding oxidoreductases acting on vanillin are crucial engineering targets for further improvement.

Received 14 October, 2020; accepted 20 January, 2021.
For correspondence. E-mail magnus.carlquist@tmb.lth.se; Tel. (+46) 46 222 98 75; Fax: (+46) 46 222 42 03.
Present address: Applied Microbiology, Lund University, Kemicentrum, Naturvetarvägen 14, Lund, 22100, Sweden.
Present address: Ginkgo Bioworks, 27 Drydock Ave, Boston, MA, 02210, USA.
Microbial Biotechnology (2021) **0**(0), 1–15
doi:10.1111/1751-7915.13764
Funding information
This study was financially supported by the Program for Institutional Internationalization from the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES/PrInt) – Project number: 88887.364008/2019-00, and Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES/MEC) – Grant Number: 88881.155712/2017-01 and STINT (The Swedish Foundation for International Cooperation in Research and Higher Education) Grant number: BR2017-7077.

Introduction

Amines are essential chemical building blocks in the chemical industry used to produce various pharmaceuticals, agrochemicals, cleaning agents, personal care products and polymers (Froidevaux *et al*., 2016; Kelly *et al*., 2018). Most amines produced industrially are derived from platform chemicals of fossil origin. Nevertheless, there is a growing interest in developing bio-based production processes from non-food renewable resources following the green chemistry principles (Masuo *et al*., 2016; Zhou *et al*., 2018; Blondiaux *et al*., 2019). In recent years, technical lignins, such as Kraft lignin, lignosulfonates, soda lignin and organosolv lignin, have been recognized as an abundant potential source of platform chemicals through bio-refining (see previous reviews (Abdelaziz *et al*., 2016; Wang *et al*., 2019)). Microbial bioconversion of depolymerized technical lignin (DTL) to various products, such as polyhydroxyalkanoate (Salvachúa *et al*., 2020), muconic acid (Kohlstedt *et al*., 2018; van Duuren *et al*., 2020), lipids (Bhatia *et al*., 2019) and others (reviewed by Becker and Wittmann (2019)) have been developed. Bioconversion of DTL to amines has, however, been studied little.

Bio-catalytic amine production can be achieved using amine transaminases (ATAs), which is now the preferred method in the pharmaceutical industry (Bryan *et al*., 2018). ATAs are pyridoxal 5'-phosphate (PLP)-dependent enzymes (E.C. 2.6.1.x) that catalyse reversible transamination of carbonyl compounds through a

ping-pong bi-bi mechanism (Patil *et al*., 2018). Unfavourable thermodynamic equilibrium towards the amine product and enzyme inhibition by (co-) products often hinder reaction efficiency. Thus, strategies for surpassing this typically involves using an excess amount of amine donors (such as L-alanine or isopropylamine) and/or coupled reactions to remove the co-product or to recycle the amine donor (Wu *et al*., 2016; Wu *et al*., 2017a; Zhou *et al*., 2018). Whole-cell transamination (Weber *et al*., 2014a; Weber *et al*., 2017; Patil *et al*., 2019; Molnár *et al*., 2019) offers several advantages over *in vitro* because it supplies PLP and amine donor, recycles cofactors, removes co-product, as well as simplifies upstream preparation of the biocatalyst, which altogether result in significant process improvement (Tufvesson, Lima-Ramos, Jensen, *et al*., 2011; Tufvesson, Lima-Ramos, Nordblad, *et al*., 2011). On the other hand, whole-cell systems are often sensitive to adverse process conditions, such as high titres of inhibitory substrates and products. They may harbour native activities resulting in by-product formation (Straathof *et al*., 2019). In the case of ATA-based whole-cell production of amines, identification and knock-out of native enzymes that acts on the carbonyl substrate are typical engineering targets.

Pseudomonas putida is a gram-negative bacterium with high tolerance to aromatic compounds, and it carries several well-characterized funnelling pathways for the assimilation of various aromatics (Brink *et al*., 2019; Xu *et al*., 2019). Furthermore, it is a somewhat industrially domesticated bacterium with a wide range of genetic tools for metabolic engineering (Aparicio *et al*., 2020; Wirth *et al*., 2020; Batianis *et al*., 2020), as well as suitable strategies for bioprocess scale-up (Nikel and Lorenzo, 2018). Altogether this has resulted in *P. putida* being one of the most preferred platform hosts for the valourization of DTL (Becker and Wittmann, 2019). Metabolically engineered *P. putida* has previously been demonstrated to be efficient in converting DTL to various carboxylic acids, for example muconic acid, pyruvic acid, lactic acid and 2,5-furandicarboxylic acid (Johnson *et al*., 2019; Weimer *et al*., 2020). To the best of our knowledge, the use of *P. putida* for bioconversion of DTL to phenolic amines has, however, not been studied previously. Herein, *P. putida* was engineered and evaluated for whole-cell bioconversion of vanillin and ferulic acid to vanillylamine (VA; Fig. 1). The interest in VA comes from its use as intermediate in the synthesis of bioactive compounds, such as capsaicinoids (Sudhakar Johnson *et al*., 1996; Anderson *et al*., 2014; Arce-Rodríguez and Ochoa-Alejo, 2019), and as a precursor for the production of polyepoxides (Fache *et al*., 2015; Mogheiseh *et al*., 2020).

Pseudomonas putida KT2440 has several native enzymes that use vanillin as substrate, including vanillin dehydrogenase and several aldehyde reductases resulting in the formation of the by-products vanillyl alcohol or vanillic acid (Simon *et al*., 2014), with the latter being further assimilated via protocatechuate and the β-ketoadipate pathway (Fig. 1). A more suitable host strain than KT2440 for production of VA may therefore be GN442ΔPP_2426, which was previously engineered to produce vanillin from ferulic acid (Graf and Altenbuchner, 2014; García-Hidalgo *et al*., 2020). In this strain, the native genes coding for vanillin dehydrogenase (*vdh*), two aldehyde dehydrogenases (*aldB-II* (PP_2680) and *aldB-I* (PP_0545)), a benzaldehyde dehydrogenase (PP_1948) and the coniferyl alcohol dehydrogenase *calA* (PP_2426) were previously deleted (García-Hidalgo *et al*., 2020; Fig. 1). Also, genes coding for feruloyl-CoA synthetase (*fcs*) and enoyl-CoA hydratase/aldolase (*ech*) are also overexpressed, resulting in the improved conversion of ferulic acid to vanillin (Graf and Altenbuchner, 2014). Under non-growing conditions, ferulic acid was converted to vanillin with 82% yield (mol/mol) using this strain (García-Hidalgo *et al*., 2020).

In the present work, three native and two non-native ATAs were evaluated for *in vivo* activity towards VA in *P. putida*. For this purpose, a novel growth-based screening method suitable for high-throughput evaluation of candidate genes was developed. Identified genes encoding the most active enzymes, putrescine transaminase Pp-SpuC-II (Galman *et al*., 2018) and the well-known ATA from *Chromobacterium violaceum* (Cv-ATA; Kaulmann *et al*., 2007) were overexpressed in the strain GN442ΔPP_2426 (García-Hidalgo *et al*., 2020). Furthermore, the effect of coexpressing Alanine dehydrogenase (AlaDH) from *Bacillus subtilis* for the regeneration of L-alanine was evaluated. Our findings provide a proof-of-principle that *P. putida* with engineered ATA activity in combination with reduced vanillin assimilation indeed can be used for VA production using ammonium as a nitrogen source and without supplementation of L-alanine in the reaction broth. Still, adverse oxidoreductase background activities converting vanillin to vanillic acid and vanillyl alcohol remain key bioengineering targets to reach increased titre and yield.

Results and discussion

Development of a growth-based method to measure transaminase activity towards vanillylamine in P. putida *KT2440 and identification of ATAs*

Pseudomonas spp. are known to metabolize biogenic amines (see review by Luengo and Olivera (2020)). Also, several natural isolates of *P. putida* were previously found to grow using VA as the sole carbon source (Flagan and Leadbetter, 2006), indicating that endogenous transaminases can convert VA to vanillin, which is

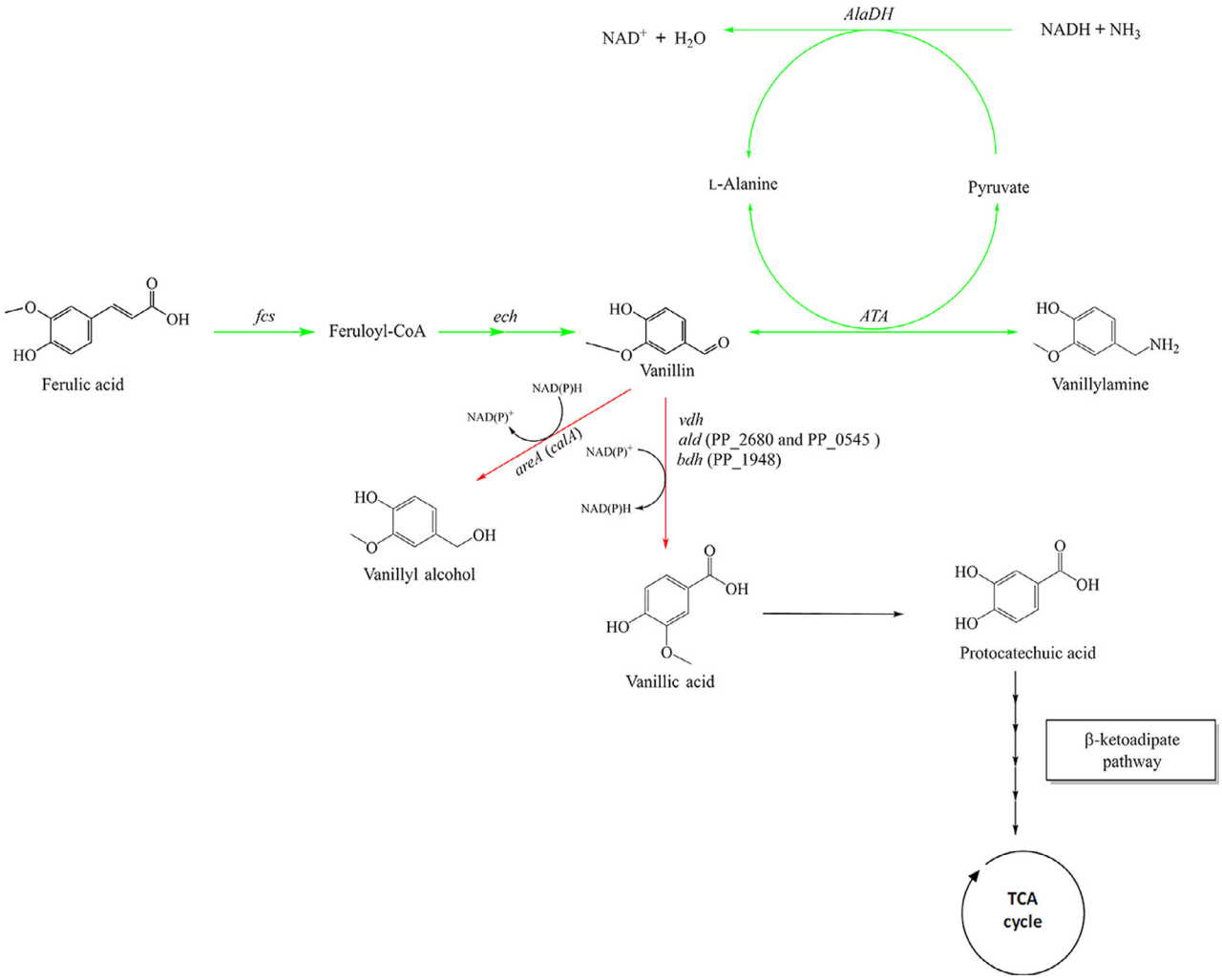


Fig. 1. Biochemical route for conversion of ferulic acid to vanillylamine, and major by-products vanillyl alcohol and vanillic acid. The first step is the synthesis of vanillin from ferulic acid by the two enzymes Feruloyl-CoA-synthetase and Enoyl-CoA-hydratase/aldolase encoded by the genes *fcs* and *ech*, respectively. The reversible transamination of vanillin is accomplished by an amine transaminase (ATA). The amino group is transferred from the amino donor (L-alanine) to vanillin with the formation of the corresponding amine and release of a co-product (pyruvate). Alanine dehydrogenase (AlaDH) recycles L-alanine from pyruvate consuming NADH and NH_3. Conversion of vanillin to vanillyl alcohol, and to vanillic acid is prevented by the deletion of vanillin dehydrogenase (*vdh*), and the aldehyde dehydrogenases (PP_2680 and PP_0545) and *bdh* (PP_1948; Graf and Altenbuchner, 2014), as well as *areA* (*calA*) encoding a vanillin reductase (García-Hidalgo *et al*., 2020). Black arrows: Not modified; Green arrows: overexpressed genes; Red arrows: deleted genes.

assimilated via the β-ketoadipate pathway. *In vivo* evaluation of ATAs by growth on specific amines as a carbon source is a powerful high-throughput screening tool to facilitate whole-cell biocatalyst development. Here, a *P. putida* KT2440 whole-cell assay was assessed by determining cell growth profiles in liquid M9 medium supplemented with glucose, vanillin or VA as the sole carbon source. No growth was observed for *P. putida* KT2440 (wild-type) during 48 h of cultivation with VA as the only carbon source, while significant growth with vanillin (final OD 0.8 $\pm$ 0.0) was evident (Fig. S1). This indicates that (i) the endogenous ATAs lack activity towards VA or (ii) that the endogenous ATAs are expressed at an insufficient level to sustain cell growth under the applied conditions. To investigate which of the two alternatives was the case, we searched for putative native transaminases in *P. putida* KT2440 (taxid: 160488) by using BLASTp analysis (Dalal and Atri, 2014) with two ATAs as query sequences; Cv-ATA from *Chromobacterium violaceum* (Kaulmann *et al*., 2007; Du *et al*., 2014) and Pp-SpuC from *P. putida* (taxid: 303; Galman *et al*., 2017; Galman *et al*., 2018). Cv-ATA (PDB id: 4BA5) was used as a query sequence since it has previously been found to catalyse the desired reaction (Kaulmann *et al*., 2007), and the putrescine transaminase Pp-SpuC (PDB id: 6HX9) was chosen because it has been shown to

recognize a broad spectrum of aromatic compounds as substrate (Galman *et al*., 2017; Galman *et al*., 2018). The top three ATAs displaying the lowest e-value and highest sequence homology to both Pp-SpuC and Cv-ATA were chosen for further investigations and denoted here as Pp-SpuC-I (gene PP_2180), Pp-SpuC-II (gene PP_5182) and Pp-ATA (gene PP_2588), respectively (Table S1). Genes encoding Pp-SpuC-II, Pp-SpuC-I and Pp-ATA were cloned individually in the expression vector pSEVA424 carrying the IPTG-inducible *lacI*^{q}*/Ptrc* system. Plasmids were then transformed in the strain KT2440 to generate strains TMB-NM011, TMB-NM012 and TMB-NM013, respectively (Tables S2 and S3). Also, the genes encoding Cc-ATA (Weber *et al*., 2014b) and Cv-ATA, both previously shown to accept vanillin as substrate, were cloned and transformed into KT2440 generating the strains TMB-JH001 and TMB-JH002, respectively (Tables S2 and S3). Constructed strains were evaluated for their ability to proliferate in liquid M9 medium supplemented with 5-25 mM VA as the sole carbon source (Fig. 2).

The strains TMB-NM011 and TMB-JH002, overexpressing *Pp-SpuC-II* and *Cv-ATA*, respectively, displayed significant growth, demonstrating that they carry the desired activity towards VA. However, no substantial growth was seen for the wild-type or strains overexpressing *Cc-ATA* (data not shown), *Pp-SpuC-I* and *Pp-ATA* (Fig. S2). For the plant transaminase Cc-ATA, previously found to be active when expressed in *E. coli* (Weber *et al*., 2014), this may be due to low expression level or protein misfolding in *P. putida* under the applied conditions. Furthermore, no negative impact on the growth with glucose in the medium was observed for the strains overexpressing the Cc-ATA encoding gene (Figs S3 and S5), indicating that the lack of activity is not related to potential toxic effects of Cc-ATA in *P. putida*.

Growth was similar for both strains overexpressing *Pp-SpuC-II* and *Cv-ATA* on VA as limiting carbon source, and followed apparent Monod kinetics up to 15 mM with a growth rate plateau between 10 and 15 mM (Fig. 2C). Above this concentration, the cell growth rate was reduced (Fig. 2D); however, the carrying capacity was not affected (Fig. 2C), demonstrating no effect on final biomass yield (Fig. 2A and B). The high activity for Pp-SpuC-II was confirmed in subsequent whole-cell bioconversion experiments where the conversion of VA in glucose-grown cells was measured. The strain TMB-NM011, overexpressing *Pp-SpuC-II*, displayed the highest specific transaminase activity (0.420 ± 0.03 mmol/h/OD compared to 0.020 ± 0.03 mmol/h/OD for the wild-type), resulting in almost complete conversion of VA to vanillin within 6 hours (Fig. S3A). The specific transaminase activity for TMB-NM012 and TMB-NM013 was also higher than the wild-type (0.052 ± 0.11 mmol/h/OD and 0.085 ± 0.00 mmol/h/OD, respectively) and reached complete conversion of 5 mM VA within 24 hours. For wild-type KT2440, only 28% of the substrate was converted over the same period, demonstrating that native vanillin transaminases are weakly expressed in KT2440 under the studied conditions (Fig. S3D). Overall, the results suggest that all three native transaminases recognize VA as substrate in *P. putida* and Pp-SpuC-II showed the highest *in vivo* specific activity. However, the basal activity in the wild-type strain is not enough to sustain growth on VA (Fig. S1).VA is not likely the natural substrate for the native ATAs presented here, although they have activity against VA in different levels. The Pp-SpuC-II, which is closely related to Pp-SpuC from other *P. putida* strain (taxid: 303), is a putrescine transaminase that has preference for aliphatic diamines (Galman *et al*., 2017; Galman *et al*., 2018). Wu *et al*. (2017b) demonstrated the potential of Pp-SpuC from *P. putida* NBRC 14164 for the kinetic resolution of several racemic amines and amino alcohols and corroborated its role as a promiscuous transaminase. In *P. aeruginosa* PAO1, the SpuC is the major putrescine transaminase and responsible for the conversion of putrescine to 4-aminobutyraldehyde (Lu *et al*., 2002). Therefore, this indicates that Pp-SpuC-II (PP_5182) and/or Pp-SpuC-I (PP_2180) are promiscuous transaminases that may be related to putrescine metabolism in *P. putida* KT2440. Pp-ATA (PP_2588) has been annotated as a Class III aminotransferase (Nelson *et al*., 2002), but has not been characterized previously. From sequence analysis in the conserved protein domain database (Lu *et al*., 2020), it is found to belong to the highly conserved aspartate aminotransferase family protein and it can be speculated to be a part of the amino acid metabolism in *P. putida* KT2440.

Next, the possibility to identify vanillin transaminases by growing the recombinant cells on a solid medium as an integrated part of the transformation protocol was investigated. The constructed plasmid harbouring *Cv-ATA* (pJH002) and the empty plasmid (pSEVA424) were again transformed into KT2440. The cells were poured directly onto M9 plates supplemented with 5 mM VA as the sole carbon source, streptomycin to select for the plasmid and IPTG to induce gene expression under the control of the *lacI*^{q}*/Ptrc* system. Agar plates were incubated at 30 °C for 3 days and then evaluated for colonies' presence (Fig. 3).

There were no colonies on the plate with transformants bearing the empty plasmid (Fig. 3A), while small colonies were visible in the plate containing the cells expressing *Cv-ATA* gene (Fig. 3B). The selection principle was thus found to be functional, that is only cells overexpressing the vanillin transaminase were able to sustain growth with VA as sole carbon source. However,

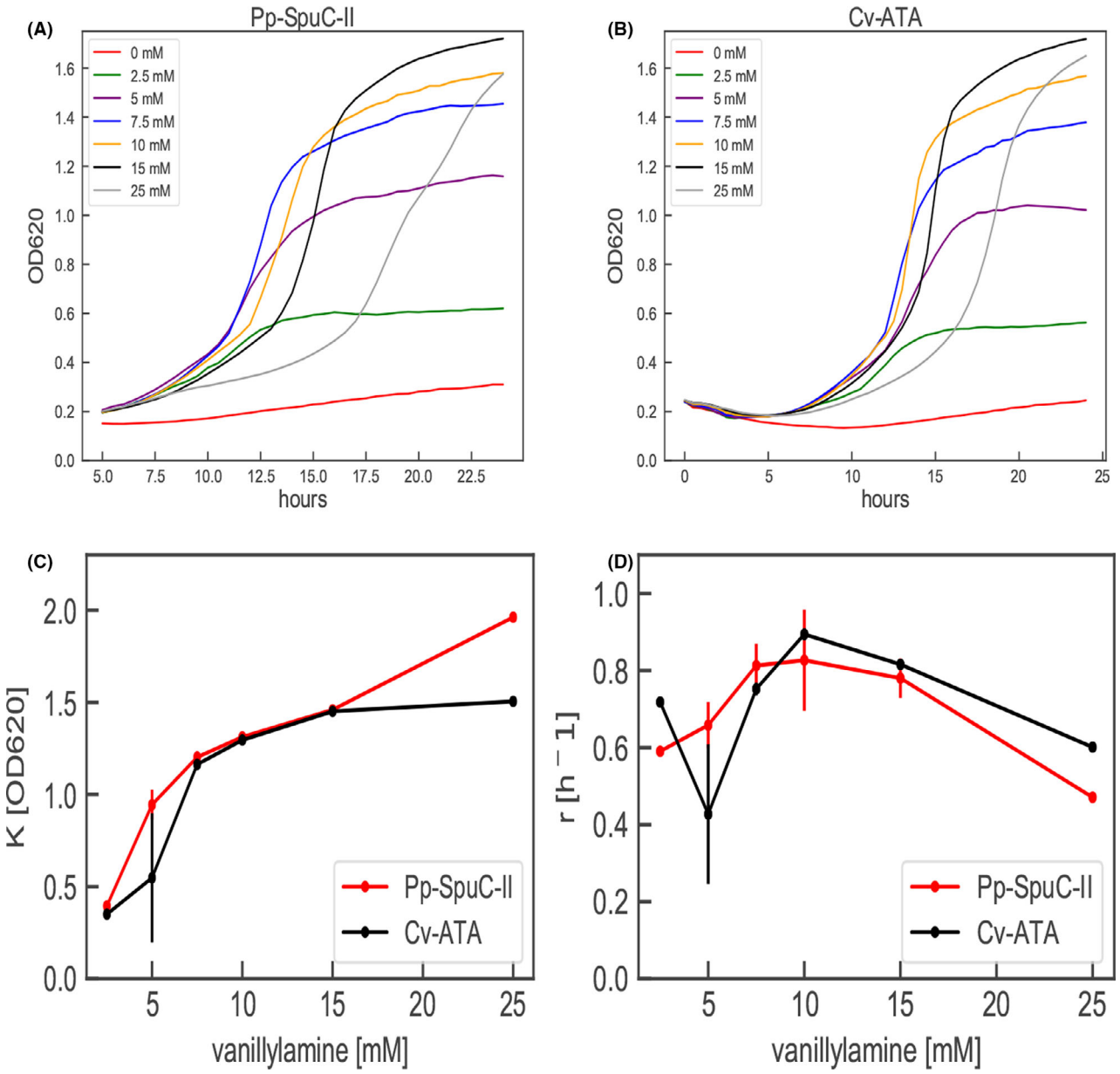


Fig. 2. Growth of *P. putida* KT2440 overexpressing (A) Pp-SpuC-II (TMB-NM011) or (B) Cv-ATA (TMB-JH002) in M9 medium supplemented with different concentrations of vanillylamine (0–25 mM) as sole carbon source. Growth curves are representative figures from one of the two biological replicates. After fitting the data in (A–B) to a logistic growth curve (described in Experimental Procedures) the estimated Carrying capacity (K; Panel C) and the maximal growth rates (r; Panel D) for each concentration are plotted against vanillylamine concentration for both Pp-SpuC-II (red line) and Cv-ATA (black line).

addition of streptomycin to the medium was found to be necessary (Fig. S4), which may be explained by a correlation between the presence of the antibioticum and a sufficient plasmid copy number to ensure high enough expression level of the transaminase. A positive correlation between the concentration of antibioticum used for selection, the plasmid copy number and the expression level of a recombinant gene has been found previously (Begbie *et al*., 2005; Lian *et al*., 2016).

Nevertheless, the results demonstrate that it is possible to couple the growth-based selection method directly to the transformation pipeline of *P. putida* KT2440. Hence, it can be used for high-throughput screening of novel ATAs accepting VA as substrate or even developing evolutionary engineering studies of these enzymes allowing a better understanding of the structure–activity relationship. It may be possible to extrapolate the method for identifying ATAs able to convert other amines

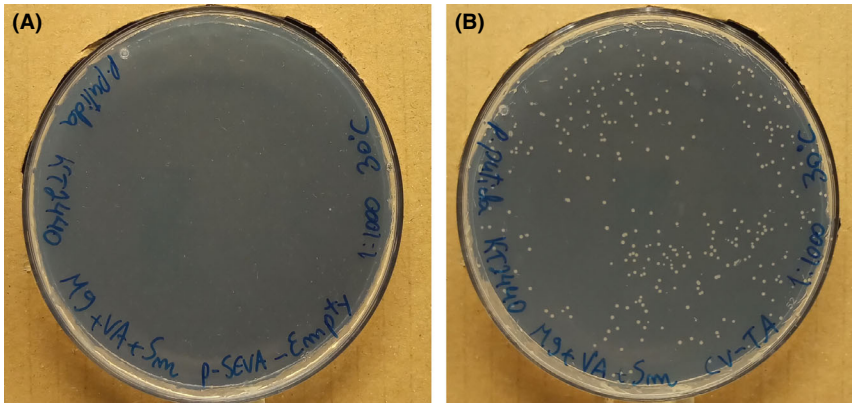


Fig. 3. M9 plates after 3 days incubated at 30 °C with 5 mM of vanillylamine as the sole carbon source, streptomycin (100 μg ml^{-1}) for plasmid selection and IPTG (1 mM) to induce gene expression. A: KT2440 cells with the empty pSEVA424 plasmid; B: TMB-JH002 cells expressing *Cv-ATA* gene.

as long as the carbonyl product is metabolized by *P. putida* KT2440. Growth-based screening methods for identifying genes coding for specific ATAs have to the best of our knowledge not been described previously. Other approaches using chemical analysis methods have been described (Baud *et al*., 2017; Pawar *et al*., 2018; Coscolín *et al*., 2019). However, they can be time-consuming or require high-cost reagents and/or equipment.

Overexpression of transaminases in GN442ΔPP_2426 enables whole-cell bioconversion of vanillin to VA

Genes encoding ATAs identified to carry the desired vanillin transaminase activity (*Cv-ATA*, *Pp-SpuC-II*, *Pp-SpuC-I* and *Pp-ATA*) and also *Cc-ATA* were overexpressed in GN442ΔPP_2426. *In vivo* transaminase activity of constructed strains was assessed by whole-cell bioconversion of VA to vanillin in M9 medium supplemented with glucose to verify if similar results as for KT2440 are obtained when the enzymes are produced by the strain GN442ΔPP_2426 (Fig. S5). Based on the *in vivo* assays, the initial conversion rates (first 6 h) were highest for strains overexpressing *Cv-ATA* (0.276 ± 0.09 mmol/h/OD) and *Pp-SpuC-II* (0.281 ± 0.10 mmol/h/OD; Table 1; Fig. S5), which is in correlation with the observed growth on VA for the KT2440 strain background. These results indicate that both ATAs showed the highest specific transaminase activity against VA. No hyperexpression was observed in the SDS-PAGE analysis performed in GN442ΔPP_2426 strains expressing different ATAs encoding genes (Fig. S8).

Next, whole-cell bioconversion of vanillin to VA was performed using resting cells in sodium phosphate buffer supplemented with L-alanine as an amine donor (Fig. S6). All overexpression strains, except for *Cc-ATA*, were capable of producing VA, although to a different degree, in correlation with specific transaminase activities measured in the reverse direction (Fig. S6F). The maximum VA concentration reached by TMB-NM014 (Pp-SpuC-II) was 0.700 ± 0.20 mM, almost 4-fold more than TMB-NM015 (Pp-SpuC-I; 0.180 ± 0.01 mM) and almost twice the amount compared to TMB-NM016 (Pp-ATA; 0.360 ± 0.01 mM), while TMB-JH004 (Cv-ATA) produced the highest amount of VA 0.915 ± 0.30 mM in the first 2 h (Table 1). The control strain did not produce a significant amount of VA, again supporting the hypothesis that the expression of native ATAs is not induced under the applied conditions.

After the initial production phase lasting for 2–7 h, a significant re-assimilation of the product and a small amount of vanillyl alcohol and vanillic acid were observed in the broth during the bioconversion using a resting cell setup (Fig. S6). Vanillyl alcohol was the main by-product for all strains excluding TMB-JH004 (CV-ATA), which did not produce significant amounts of by-products (Fig. S6D). The lowest concentration was obtained by TMB-NM0016 (0.280 ± 0.028 mM) and highest by the cells carrying the empty plasmid (0.490 ± 0.057 mM). In contrast, the bioconversion of VA using GN442ΔPP_2426 produced mainly vanillic acid as the main by-product ant its production started between 6–10 h for all strains overexpressing ATA with activity against VA (Fig. S5). From this, it is clear that GN442 ΔPP_2426 still carries endogenous dehydrogenase activity towards vanillin. Apparently, the ATA-based conversion of vanillin to VA competes with remaining vanillin dehydrogenase background activity converting it to by-products and further metabolization. When the rate of VA formation is slowed down (after the initial 2-7h production phase), the reaction equilibrium shifts and VA is re-assimilated. It is probably connected to a very low

Table 1. Summary of results obtained from whole-cell bioconversions performed using *P. putida* GN442ΔPP_2426 overexpressing different ATAs.

Strain	ATAs	*In vivo* specific transaminase activity in GN442ΔPP_2426[a] (mmol h^{-1} OD^{-1})	Fold-change in specific transaminase activity[b]	Maximum VA (mM)[c]/Time	VA Yield (mol mol^{-1}; %)[d]
GN442ΔPP_2426 carrying the empty pSEVA424 plasmid	Negative control	0.003 ± 0.00^{a}	1	0.010 ± 0.01^{a} 24 h	0.11 ± 0.16^{a}
TMB-NM014	Pp-SpuC-II	0.281 ± 0.10^{b}	94	$0.700 \pm 0.20^{b,c}$ 7 h	$7.3 \pm 2.6^{a,b}$
TMB-NM015	Pp-SpuC-I	0.032 ± 0.00^{a}	11	$0.180 \pm 0.01^{a,b}$ 24 h	$1.9 \pm 0.08^{a,b}$
TMB-NM016	Pp-ATA	0.047 ± 0.01^{a}	16	$0.360 \pm 0.01^{a,b,c}$ 6 h	$3.6 \pm 0.4^{a,b}$
TMB-JH004	Cv-ATA	0.276 ± 0.09^{b}	92	0.915 ± 0.30^{c} 2 h	9.9 ± 4.6^{b}
TMB-JH003	Cc-ATA	0.002 ± 0.00^{a}	ND	ND	ND

ND, Not determined.
Equal letters indicate no statistical difference between the samples ($p > 0.05$).
a. Values calculated based on the production of vanillin in the whole-cell bioconversion of VA to vanillin using growing cells.
b. Values calculated in relation to the negative control
c. Values obtained in the whole-cell bioconversion of vanillin to VA using resting cells. The time (h) indicates the period of bioconversion when the maximum VA titre was reached.
d. Calculated using the maximum VA obtained for each strain during the bioconversion of vanillin to VA (mol/mol) using resting cells.

amount of vanillin in the cell driving the equilibrium towards vanillin production from VA and further conversion to by-products, especially when glucose is added to the broth (Fig. S5).

Worth noting, however, is that only a fraction of the original vanillin was converted, indicating that resting cells may be suitable to avoid assimilation via vanillic acid and vanillyl alcohol due to limited regeneration of $NAD(P)^{+}$under the applied conditions. Supporting this claim is that the vanillin was fully converted when the strains were grown in M9 medium supplemented with glucose (Fig. S5). Another conclusion is that the transamination reaction equilibrium was dynamic under the applied batch process configuration. Further metabolic engineering and optimization of reaction conditions to reduce the intracellular concentration of endogenous amine acceptors (e.g. pyruvate) may prolong the production phase.

Effect of L-alanine regeneration by AlaDH on whole-cell bioconversion of vanillin to VA

The low titres obtained and re-assimilation of the product could be a consequence of unfavourable reaction equilibrium resulting in vanillin oxidation to vanillic acid being the dominating activity. Supplementation of amine donor proved to be indispensable when the bioconversion was starting with low cell density (OD_{620} 0.3~0.5) since, without supplementation, cells failed to produce VA (Fig. S7). However, even with supplementation, low intracellular availability of L-alanine and inefficient pyruvate removal by endogenous enzymes may still disadvantage the desired reaction's direction. The introduction of AlaDH from *B. subtilis* for the regeneration of L-alanine from pyruvate could assist in shifting the equilibrium, as previously shown for *E. coli* (Wu *et al*., 2016; Wu *et al*., 2017a; Patil *et al*., 2019). L-alanine regeneration was investigated by cloning the *AlaDH* gene in the plasmid pJH004 and inserting it in GN442ΔPP_2426. The resulting strain was named TMB-JH006 and overexpresses both *Cv-ATA* and *Bs-AlaDH,* under the same promoter control. Whole-cell bioconversion of vanillin was performed in M9 medium supplemented with glucose since the AlaDH requires efficient NADH regeneration to convert pyruvate to L-alanine. For this experiment, a high cell density ($OD_{620} = 10$) was used to enable higher volumetric productivity and minimize cell growth (Fig. 4). Also, the possibility to reach VA by supplementation of ammonium as the sole nitrogen source and to rely on AlaDH to produce L-alanine intracellularly was investigated. Indeed, the amine donor omission did not abolish the transamination of vanillin to VA by TMB-JH006 and TMB-JH004 (Fig. 4A and B, respectively), and coexpression of AlaDH resulted in increased production regardless if L-alanine was added or not.

In the absence of L-alanine, the maximum titre of VA reached by TMB-JH004 and TMB-JH006 were 0.27 ± 0.24 mM and 0.65 ± 0.15 mM, respectively, while in the presence of the amine donor, TMB-JH004 produced 0.70 ± 0.01 mM and TMB-JH006 0.84 ± 0.07 mM of VA. In L-alanine's presence, these values were obtained in the first 6 h of bioconversion, and after this point, the re-assimilation of VA occurred. In contrast, VA re-assimilation did not occur in the

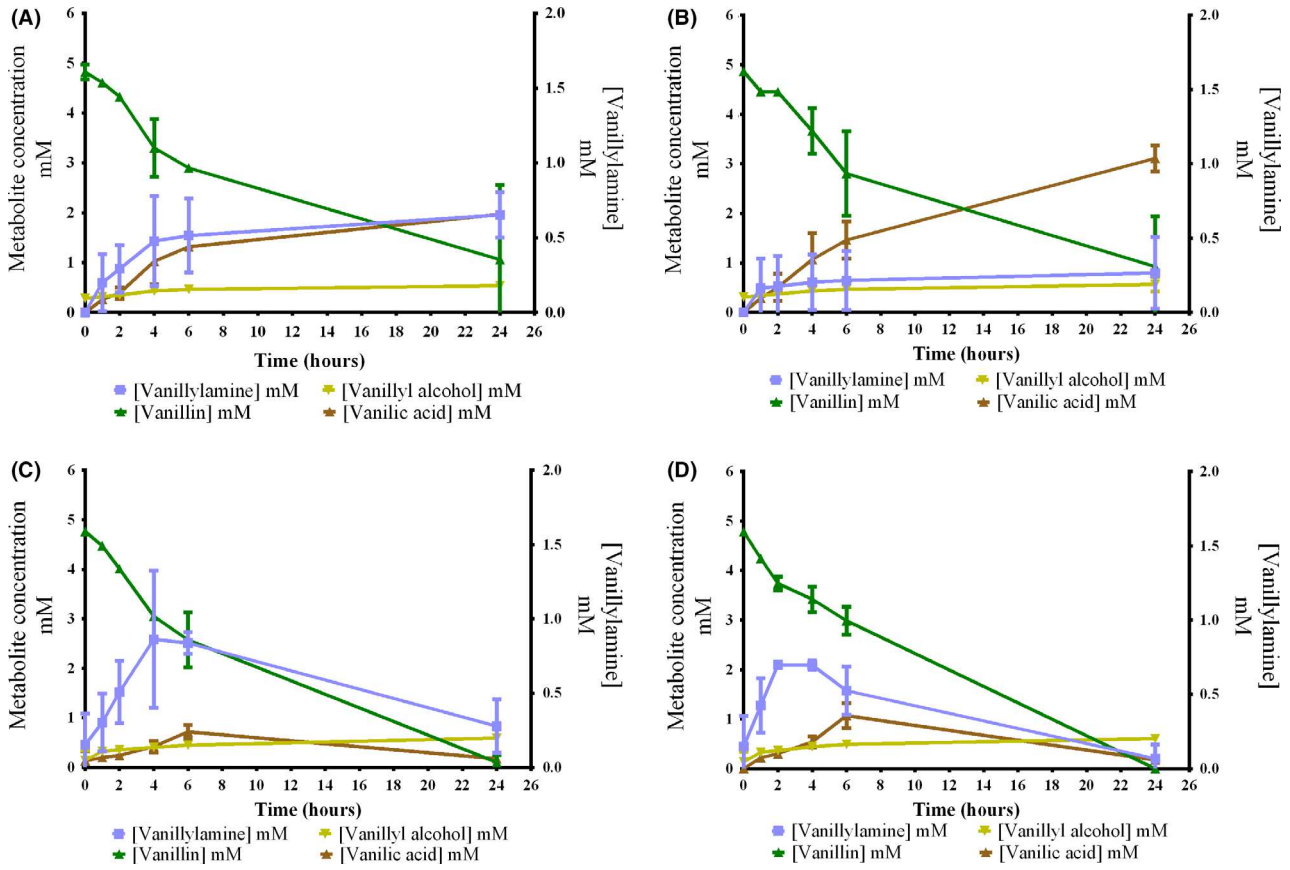


Fig. 4. Effect of L-alanine regeneration on whole-cell bioconversion of vanillin to vanillylamine using metabolically engineered *P.putida* GN442ΔPP_2426 strains as biocatalysts. Bioconversions were performed with a high cell density (OD_{620} = 10) for 24 h at 30 °C and 180 rpm in M9 medium supplemented with 5 mM vanillin, 10 g/L of glucose for NADH regeneration and 200 mM of NH_4Cl. L-Alanine was omitted for A and B and 50 mM of L-alanine was added to the medium for C and D. A and C: TMB-JH006 (Cv-ATA + AlaDH) and B and D: TMB-JH004 (Cv-ATA). Error bars indicate ± SD of two biological replicates.

absence of the amine donor, and more VA could be detected in the broth after 24 h, especially for the cell overexpressing both *Cv-ATA* and *AlaDH* (0.65 ± 0.15 mM versus 0.28 ± 0.18 mM, respectively). Therefore, the utilization of ammonium as the sole source of amine donors is efficient to avoid re-assimilation of VA, although a lower amount of product is formed in the initial phase of bioconversion (6 first hours). Furthermore, in the presence of L-alanine, the optimum sampling point appears to occur in 4 h of bioconversion, which precedes the re-assimilation of VA. While in the absence of L-alanine, a longer bioconversion time is advantageous for the transamination reaction. A more detailed profile regarding the *in vivo* transamination of vanillin could indicate the optimum harvest time, which may occur either between 6 and 24 h or after this. These results indicate that relying on endogenous amine donors, which can be supplied via the addition of ammonium, is an efficient strategy to prevent VA re-assimilation.

The observed difference for experiments with and without external supplementation of L-alanine may be due to endogenous transaminases background activity using other amine acceptors than vanillin as substrate and thereby increasing intracellular pyruvate levels from externally supplemented L-alanine. Pyruvate formed by this route could then be available as an amine acceptor for conversion of VA to vanillin by Cv-ATA or by additional native ATAs described in this study (Table 1). Although the pyruvate was not measured, it can be speculated that its intracellular levels are higher in the presence of L-alanine, and this impacted VA production. Nevertheless, the results indicate that it is advantageous to use ammonium for two reasons, (i) since it could make the addition of expensive amine donors redundant, although its utilization increases considerably the efficiency of the bioconversion in the initial phase, and (ii) since it avoids product re-assimilation.

Formation of the by-products vanillyl alcohol and vanillic acid was higher in a growing-cell setup compared

with using resting cells without glucose in the medium (Fig. S6), explaining why all vanillin was converted in the first 24 h when metabolically active cells were utilized (Fig. 4). The presence of glucose is likely to have maintained redox homeostasis, that is the NAD+/NADH ratio, which probably is the explanation for the higher by-product formation compared to the resting cell setup. Furthermore, a higher amount of cells was utilized in the bioconversion of vanillin by growing cells of TMB-JH004 and TMB-JH006 (OD_{620} = 10; Fig. 4) than the resting cells (OD_{620} = 3; Fig. S6), which may also have contributed to the higher amount of vanillic acid and vanillyl alcohol observed (Fig. 4B and D; and Fig. S6D).

Although the strain GN442ΔPP_2426 has significantly lower oxidoreductase activity towards vanillin compared to KT2440, other aldehyde dehydrogenases need to be deleted to increase the efficiency of vanillin biotransformation. In a previous proteomics study, the aldehyde dehydrogenases PP_3151, PP_5120 and PP_5258 showed an increased abundance in response to vanillin, representing possible targets to abolish the oxidation of vanillin into vanillic acid (Simon *et al*., 2014).

Whole-cell bioconversion of ferulic acid to VA

An additional series of experiments to investigate whether it was possible to convert ferulic acid to vanillylamine were performed. Similar to vanillin bioconversion, a significant amount of ferulic acid was assimilated, and the by-products vanillyl alcohol and vanillic acid were formed, resulting in a limited yield of VA (Fig. 5).

An increase of the VA production was observed in the strain coexpression Cv-ATA and AlaDH (0.473 ± 0.09 mM compared to 0.235 ± 0.18 mM, p = 0.2738) in the first 6 h of bioconversion. In contrast to vanillin bioconversions with L-alanine as amine donor, VA re-assimilation was remarkably not observed for the ferulic acid bioconversions (Fig. 4C and D). This indicates that a reduced intracellular level of pyruvate in combination with higher levels of L-alanine and/or additional endogenous metabolites acting as amine donors achieved the desired reaction equilibrium. Further investigation of intracellular metabolites under the reaction's progress is required to shed light on the matter. However, the possibility to use *P. putida* to reach VA from ferulic acid through whole-cell bioconversion is clear.

Conclusions

In this study, a novel method to select for recombinant vanillin transaminases was developed. The selection principle is based on the ability of *P. putida* KT2440 to utilize VA as the sole carbon source to sustain growth only if a suitable ATA is overexpressed. The method is easy-to-use and can be integrated into the transformation pipeline, thereby potentially being used for screening campaigns covering large sequence spaces. Here, the method was used to identify Pp-SpuC-II and Cv-ATA to possess high *in vivo* activity when overexpressed in *P. putida*. Whole-cell bioconversion of vanillin and ferulic acid to VA in batch mode was restricted by competing activity of endogenous enzymes producing vanillic acid, resulting in reduced yield and loss of product by re-assimilation. This points towards the need to knock out additional genes encoding oxidoreductases to make ω-transamination the dominating activity. When resting cells were used, the formation of by-products was significantly reduced; however, no increase in VA titres was

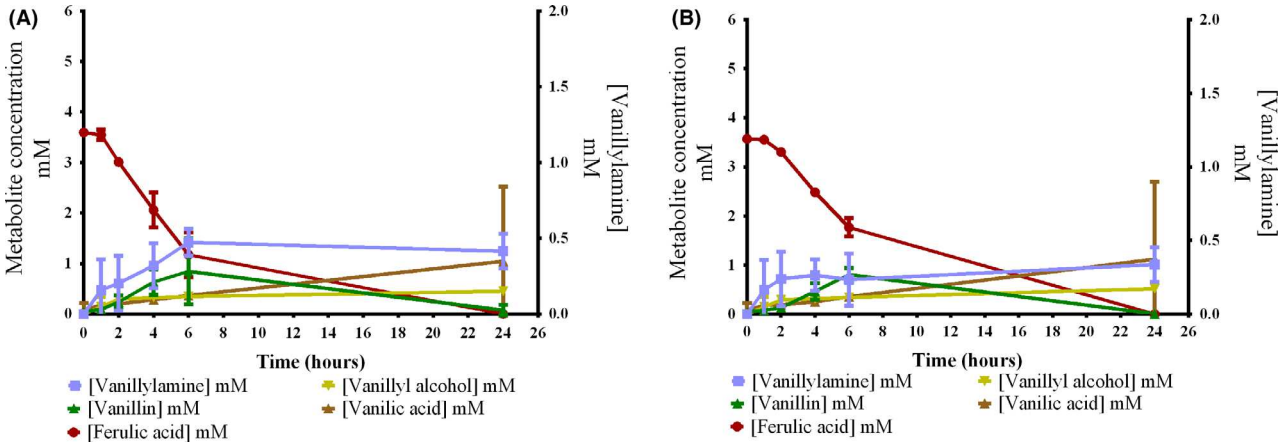


Fig. 5. Whole-cell bioconversion of ferulic acid to vanillylamine using metabolically engineered *P.putida* GN442ΔPP_2426 strains as biocatalysts. Bioconversions were performed with high cell density (OD_{620} = 10) for 24 h at 30 °C and 180 rpm in M9 medium supplemented with 3.5 mM ferulic acid, 10 g l^{-1} of glucose for NADH regeneration, 50 mM of L-Alanine and 200 mM of NH_4Cl. A: TMB-JH006 strain (Cv-ATA + AlaDH) and B: TMB-JH004 strain (Cv-ATA). Error bars indicate ± SD of two biological replicates.

observed. Also, using ammonium as the sole nitrogen source, instead of external supplementation of L-alanine, minimized the re-assimilation after the initial amine production phase. Yet, further investigation of intracellular reactants accepted by the recombinant transaminases under the different reaction phases and any potential influences of endogenous metabolites is needed to shed additional light into possible ways to shift the adverse reaction equilibrium. Also, optimization of other influencing parameters, such as PLP concentration, enzyme levels, ammonium concentration and external supply of co-substrates for NADH regeneration, are additional potential strategies for improving VA production.

Experimental procedures

Chemicals and enzymes

All reagents (vanillin, VA, vanillyl alcohol, vanillic acid, ferulic acid and L-alanine) were of analytical grade and were purchased from Merck KGaA (Darmstadt, Germany). The antibiotics ampicillin and streptomycin and isopropyl-β-D-thiogalactoside (IPTG) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Restriction enzymes, T4 DNA ligase and Phusion High-Fidelity DNA Polymerase utilized for gene cloning were ordered from Thermo Fisher Scientific (Vilnius, Lithuania). All DNA oligonucleotides (primers) utilized for cloning and sequencing were synthesized by Eurofins Genomics (Ebersberg, Germany).

Molecular biology procedures

All procedures in the cloning steps were performed according to the manufacturer's recommendations (Thermo Fisher Scientific). Plasmid DNA was isolated using GeneJET Plasmid Miniprep Kit from Thermofisher (Vilnius, Lithuania). Agarose gel extractions were performed using GeneJET Gel Extraction Kit from the same manufacturer. Genomic DNA extraction from *P. putida* KT2440 was performed using the GeneJET Genomic DNA Purification Kit from Thermo Fisher Scientific Baltics (Vilnius, Lithuania). Primers utilized for polymerase chain reactions (PCRs) are described in the Table S1. Phusion High-Fidelity DNA Polymerase was utilized for cloning and sequencing while DreamTaq Polymerase (Thermo Scientific, Germany) was utilized for colony PCR. GeneAmp PCR System 9700 (Applied Biosystems, Foster City, CA, USA) was used for the PCRs amplifications and the products were purified using GeneJET PCR Purification Kit (Thermofisher). Sequencing was performed with the amplified DNA fragments to verify the absence of mutations in the RBS sequence or coding sequence (Eurofins Genomics, Ebersberg, Germany).

Plasmids and bacterial strains

All plasmids and strains utilized in this work are described in Tables S2 and S3, respectively. The pSEVA424 plasmid was obtained from the Standard European Vector Architecture (SEVA) repository (Silva-Rocha *et al.*, 2013) and it was utilized as the backbone for the construction of the plasmids pJH001 (*Cc-ATA*), pJH002 (*Cv-ATA*), pJH004 (*Cv-ATA* + *AlaDH*), pNM011 (*Pp-SpuC-II*), pNM012 (*Pp-SpuC-I*) and pNM013 (*Pp-ATA*). The *AlaDH* gene (Sequence S1) from *B. subtilis* 168 (Genbank ID: 936557, Swiss-Prot: Q08352) was synthesized by GenScript (New Jersey, USA) and delivered in plasmid pUC57-AlaDH. *Cc-ATA* gene encoding vanillin transaminase from *C. chinense* (GenBank: AAC78480.1, Swiss-Prot: O82521) and *Cv-ATA* from *C. violaceum* (GenBank: WP011135573.1, Swiss-Prot: Q7NWG4) were amplified from pNW10 and pNW12 (Weber *et al.*, 2017), respectively (primers described in Table S4). Genes encoding the native ATAs Pp-SpuC-II (GenBank: AAN70747.1), Pp-SpuC-I (GenBank: AAN67793.1) and Pp-ATA (GenBank AAN68196.1) were amplified from *P. putida* KT2440 genome. PCR products were digested and ligated into backbone plasmid pSEVA424 digested with compatible restriction enzymes (Table S4). To construct the plasmid containing the *AlaDH* gene, the synthetic pUC57-AlaDH was digested with the restriction enzymes *PsTI* and *SpeI*. The resulting fragment was inserted into pJH002 digested with the same restriction enzymes, creating pJH004. The consensus Ribosomal Binding Site (RBS) sequence and the eight conserved nucleotides downstream to this sequence were added immediately upstream to the start codon of the *AlaDH* coding sequence (Silva-Rocha *et al.*, 2013). *Escherichia coli* DH5α strain was utilized for subcloning and plasmid maintenance. *P. putida* KT2440 (DSM 6125) was obtained from German Collection of Microorganisms and Cell Cultures (D.S.M.Z.; Braunschweig, Germany). The engineered strain GN442ΔPP_2426 was constructed previously (García-Hidalgo *et al.*, 2020) utilizing GN442 strain (Graf and Altenbuchner, 2014) as the scaffold. KT2440 was utilized as background for the construction of TMB-JH001, TMB-JH002, TMB-NM011, TMB-NM012 and TMB-NM013 strains, while GN442ΔPP_2426 strain was used to construct the other strains (Table 1). All bacterial strains were prepared and kept in glycerol stocks 20% at −80 °C until the experiments.

Electrocompetent cells preparation and transformation conditions

Heat-shock transformation of *E. coli* was performed as described previously (Green and Sambrook, 2018).

Transformed cells were selected in LB plates supplemented with ampicillin (100 μg ml^{-1}) or streptomycin (60 μg ml^{-1}). *P. putida* KT2440 and GN442ΔPP_2426 electrocompetent cells were prepared as described previously (Martínez-García and de Lorenzo, 2012) with slight modifications. Overnight cultures were transferred to a fresh LB medium with the OD_{620} adjusted to 0.1 and then cultivated in shake flasks until OD_{620} reached 0.8. Cells in mid-exponential-phase were washed two times with sucrose (300 mM) and re-suspended in the same solution. Freshly prepared electrocompetent cells and plasmids (100 ng) were mixed in an electroporation cuvette with 2 mm gap width. The electric pulse was applied using a Gene Pulser apparatus equipped with a Pulse Controller (Bio-Rad, Hercules, CA, USA). Subsequently, cells were plated on M9 plates with citrate (10 mM) or vanillylamine (5 mM) as sole carbon source and streptomycin (100 μg/ml) for the selection. For the growth-based ATA selection method, 1 mM IPTG was added to induce gene expression. Cells were diluted 1000x and 10 000x using a sodium chloride solution (0.9% NaCl) prior to plating. Plates were incubated at 30 °C for 3 days.

Media and cultivation conditions

Lysogeny broth (LB) medium (tryptone 10 g l^{-1}, yeast extract 5 g l^{-1}, NaCl 10 g l^{-1}) was used for routine cultivation of *E. coli* strain DH5α at 37 °C and *P. putida* strains at 30 °C. M9 medium (Sambrook and Russell, 2001; 6 g l^{-1} disodium phosphate, 3 g l^{-1} monopotassium phosphate, 0.5 g l^{-1} sodium chloride, 1 g l^{-1} ammonium chloride, 2 mM magnesium sulphate, 100 μM calcium chloride, pH 7) with 1% (v/v) trace elements solution 100× (0.5 g l^{-1} EDTA disodium dihydrate, 0.2 g l^{-1} $FeSO_4 \cdot 7\ H_2O$, 0.01 g l^{-1} $ZnSO_4 \cdot 7\ H_2O$, 0.003 g l^{-1} $MnCl_2 \cdot 4\ H_2O$, 0.03 g l^{-1} H_3BO_3, 0.02 g l^{-1} $CoCl_2 \cdot 6\ H_2O$, 0.001 g l^{-1} $CuCl_2 \cdot 2\ H_2O$, 0.002 g l^{-1} $NiCl_2 \cdot 6\ H_2O$, 0.003 g l^{-1} $Na_2MoO_4 \cdot 2\ H_2O$; Pfennig and Lippert, 1966) and appropriate carbon source was used for cultivation of *P. putida*. Stock solutions of VA 10× (50 mM), vanillin 10× (50 mM), D-glucose 10× (100 g l^{-1}) and citrate 50× (0.5 M) were sterilized by filtration using 0.2 μm filter and added to the M9 media immediately before use. Ampicillin was utilized for *E. coli* selection (100 μg ml^{-1}) and streptomycin in the selection of *E. coli* (60 μg ml^{-1}) as well for *P. putida* strains (100 μg ml^{-1}). IPTG was added to induce the expression of ATA genes regulated by promoter system *lacIq*-P*trc*. To evaluate the capacity of *P. putida* KT2440 to metabolize VA, one single colony was picked in a LB plate and grown overnight on M9 medium with glucose (10 g l^{-1}). On the following day, the pre-inoculum was washed with a sodium chloride solution (0.9% NaCl) to remove residual glucose. Optical density at 620 nm (OD_{620}) was measured using a Ultrospec 2100pro spectrophotometer (Amersham Biosciences, Sweden) to prepare the inoculum (initial $OD_{620} = 0.3$) at 250 ml baffled shake flask containing 25 ml of M9 medium supplemented with 5 mM vanillylamine, 5 mM vanillin or 10 g l^{-1} D-glucose (55 mM) as carbon source. Cells were cultivated for 48 h at 30 °C and under shaking (180 rpm) using a Innova 43 Shaking Incubator (New Brunswick Scientific, Edison, NJ, USA) and the OD_{620} was monitored periodically. All shake flask cultivations were carried out in three biological replicates. For the plate reader growth experiments, an individual colony was used to inoculate 5 ml of LB containing 10 μg ml^{-1} of streptomycin. Overnight culture was then washed twice in M9 medium and diluted to a final OD of ~0.25–0.4 into a 96-well microtitre plate (Sarstedt, Germany) supplemented with vanillylamine ranging from 0 to 25 mM. Plates were incubated in a Multiskan™ FC Microplate Photometer (Thermo Fisher Scientific, USA) plate reader at 30 °C with shaking before each measurement of OD_{620} for a time period of 24 h. Growth curves were fitted to a logistic curve (Equation 1) using Python SciPy Optimize (Oliphant, 2007).

$$f(t) = \frac{K}{1 + \left(\frac{K - x_0}{x_0}\right) \cdot e^{-rt}} + b, \quad (1)$$

where K is the carrying capacity, r is the growth rate, x_0 is the initial cell concentration, t is the time and b is the floor term (accounting for measurement background).

Whole-cell bioconversions

Pseudomonas putida strains were pre-grown overnight at 30 °C with shaking in 10 ml LB supplemented with streptomycin to avoid plasmid loss. Subsequently, cells were harvested by centrifugation (4000 rpm, 5 min), washed and transferred to M9 media supplemented with 1 mM IPTG, 100 μg ml^{-1} streptomycin and either 10 g l^{-1} glucose, 5 mM vanillylamine, 5 mM vanillin, in a 250 ml baffled shake flask to an OD of ~0.3–0.5. For preparation of cells for bioconversions using high cell density (OD_{620} = 3–10 as indicated in the results and discussion), the precultures were instead grown in 50–100 ml LB supplemented with streptomycin in a 500 ml baffled shake flask and grown until OD ~ 1. At this point, 5 mM IPTG was added to induce gene expression, and the culture was grown at 23 °C overnight. Subsequently, cells were harvested and washed, and re-suspended in 5 ml sodium phosphate buffer (50 mM, pH 7.2) or 5 ml M9 medium supplemented with 5–10 mM vanillin or 3–10 mM ferulic acid, 0–10 g l^{-1} glucose, 0–100 mM L-alanine and/or 0-200 mM of ammonium chloride. Whole-cell

bioconversions were performed at 30 °C under aerobic conditions in shake flasks or glass vials with stirring for 24 h, and 0.5 ml samples were frequently withdrawn for OD620 measurements, centrifuged and frozen at −20 °C prior to HPLC analysis.

HPLC analysis

For HPLC metabolite analysis, a Select C18 column (4.6 × 150 mm) was used with a Waters HPLC system (Waters Binary HPLC pump 1525, UV/Vis detector 2489, Auto sampler 2707, All Waters Corporation, Milford, USA). Reverse-phase chromatography was performed using two mobile phases: millipore water with 0.1% trifluoroacetic acid (TFA; A) and acetonitrile (B). An isocratic method with 65% A and 35% B for 10 min with a flow of 1 ml min^{-1}, at a monitored wavelength at 281 nm was used. The analysis was performed at room temperature. A standard curve was made to calculate the concentration of the product in the sample.

Data analysis

All bioconversions curves graphs were prepared using Graphpad Prism software version 6.0 (GraphPad Software, San Diego, CA, USA) and display the concentration of metabolites obtained from HPLC analysis and/or the OD620 measured by spectrophotometer. The error bars represent the standard deviation from at least two biological replicates. The statistical analyses were performed using Graphpad Prism software (version 6.0). The one-way ANOVA test followed by Tukey's post-test ($P < 0.05$) or the unpaired *t*-test with Welch's correction ($P < 0.05$) were performed when three and two samples were compared, respectively.

Acknowledgements

The authors would like to thank Javier García-Hidalgo (Applied Microbiology, Lund university) for kindly providing the *P. putida* strain GN442ΔPP_2426, and Esteban Martínez-García and Tomás Aparicio from the National Center for Biotechnology (CSIC, Madrid, Spain) for kindly providing the plasmid pSEVA424.

Conflict of interest

The authors declare that they have no conflict of interest.

Author contributions

J.H.C.M.N. contributed to the design of the study, performed the bioinformatics analysis, design and performed the experiments, analysed the data and drafted the manuscript. F.L. and N.M. helped to perform the bioconversions experiments, the growth curves and helped with the data interpretation and write the paper. N.M. developed the HPLC setup used and helped with the analytical analyses, the data interpretation and to write the Experimental Procedures section. E.M.L. performed the growth curves in microplates and prepared the parameter fitting for the growth curves, helped to perform the bioconversions experiments and write the paper. N.S.P. and MC revised and helped to write the manuscript. MC conceived and designed the study and participated in the experimental design and data interpretation. All authors read and approved the final manuscript.

References

Abdelaziz, O.Y., Brink, D.P., Prothmann, J., Ravi, K., Sun, M., García-Hidalgo, J., *et al*. (2016) Biological valorization of low molecular weight lignin. *Biotechnol Adv* **34:** 1318–1346.

Anderson, M., Afewerki, S., Berglund, P., and Cõrdova, A. (2014) Total synthesis of capsaicin analogues from lignin-derived compounds by combined heterogeneous metal, organocatalytic and enzymatic cascades in one pot. *Adv Synth Catal* **356:** 2113–2118.

Aparicio, T., Nyerges, A., Martínez-García, E., and de Lorenzo, V. (2020) High-efficiency multi-site genomic editing of pseudomonas putida through thermoinducible ssDNA recombineering. *iScience* **23:** 100946.

Arce-Rodríguez, M.L., and Ochoa-Alejo, N. (2019) Biochemistry and molecular biology of capsaicinoid biosynthesis: recent advances and perspectives. *Plant Cell Rep* **38:** 1017–1030.

Batianis, C., Kozaeva, E., Damalas, S.G., Martín-Pascual, M., Volke, D.C., Nikel, P.I., and Martins dos Santos, V.A.P. (2020) An expanded CRISPRi toolbox for tunable control of gene expression in *Pseudomonas putida*. *Microb Biotechnol* **13:** 368–385.

Baud, D., Jeffries, J.W.E., Moody, T.S., Ward, J.M., and Hailes, H.C. (2017) A metagenomics approach for new biocatalyst discovery: application to transaminases and the synthesis of allylic amines. *Green Chem* **19:** 1134–1143.

Becker, J., and Wittmann, C. (2019) A field of dreams: lignin valorization into chemicals, materials, fuels, and healthcare products. *Biotechnol Adv* **37:** 107360.

Begbie, S., Dominelli, G., Duthie, K., Holland, J., and Jitratkosol, M. (2005) The effects of sub-inhibitory levels of chloramphenicol on pBR322 plasmid copy number in *Escherichia coli* DH5α cells. *J Exp Microbiol Immunol* **7:** 82–88.

Bhatia, S.K., Gurav, R., Choi, T.R., Han, Y.H., Park, Y.L., Park, J.Y., *et al*. (2019) Bioconversion of barley straw lignin into biodiesel using *Rhodococcus* sp. YHY01. *Bioresour Technol* **289:** 121704.

Blondiaux, E., Bomon, J., Smoleń, M., Kaval, N., Lemière, F., Sergeyev, S., *et al*. (2019) Bio-based aromatic amines

from lignin-derived monomers. *ACS Sustain Chem Eng* **7:** 6906–6916.

Brink, D.P., Ravi, K., Lidén, G., and Gorwa-Grauslund, M.F. (2019) Mapping the diversity of microbial lignin catabolism: experiences from the eLignin database. *Appl Microbiol Biotechnol* **103:** 3979–4002.

Bryan, M.C., Dunn, P.J., Entwistle, D., Gallou, F., Koenig, S.G., Hayler, J.D., *et al*. (2018) Key Green Chemistry research areas from a pharmaceutical manufacturers' perspective revisited. *Green Chem* **20:** 5082–5103.

Coscolín, C., Katzke, N., García-Moyano, A., Navarro-Fernández, J., Almendral, D., Martínez-Martínez, M., *et al*. (2019) Bioprospecting reveals class ω-transaminases converting bulky ketones and environmentally relevant polyamines. *Appl Environ Microbiol* **85:** 1–20.

Dalal, A., and Atri, A. (2014) An introduction to sequence and series. *Int J Res* **1:** 1286–1292.

Du, C.J., Rios-Solis, L., Ward, J.M., Dalby, P.A., and Lye, G.J. (2014) Evaluation of CV2025 ω-transaminase for the bioconversion of lignin breakdown products into value-added chemicals: Synthesis of vanillylamine from vanillin. *Biocatal Biotransformation* **32:** 302–313.

van Duuren, J.B.J.H., de Wild, P.J., Starck, S., Bradtmöller, C., Selzer, M., Mehlmann, K., *et al*. (2020) Limited life cycle and cost assessment for the bioconversion of lignin-derived aromatics into adipic acid. *Biotechnol Bioeng* **117:** 1381–1393.

Fache, M., Montérémal, C., Boutevin, B., and Caillol, S. (2015) Amine hardeners and epoxy cross-linker from aromatic renewable resources. *Eur Polym J* **73:** 344–362.

Flagan, S.F., and Leadbetter, J.R. (2006) Utilization of capsaicin and vanillylamine as growth substrates by *Capsicum* (hot pepper)-associated bacteria. *Environ Microbiol* **8:** 560–565.

Froidevaux, V., Negrell, C., Caillol, S., Pascault, J.P., and Boutevin, B. (2016) Biobased amines: from synthesis to polymers; present and future. *Chem Rev* **116:** 14181–14224.

Galman, J.L., Gahloth, D., Parmeggiani, F., Slabu, I., Leys, D., and Turner, N.J. (2018) Characterization of a putrescine transaminase from *Pseudomonas putida* and its application to the synthesis of benzylamine derivatives. *Front Bioeng Biotechnol* **6:** 1–11.

Galman, J.L., Slabu, I., Weise, N.J., Iglesias, C., Parmeggiani, F., Lloyd, R.C., and Turner, N.J. (2017) Biocatalytic transamination with near-stoichiometric inexpensive amine donors mediated by bifunctional mono- and di-amine transaminases. *Green Chem* **19:** 361–366.

García-Hidalgo, J., Brink, D.P., Ravi, K., Paul, C.J., Lidénb, G., and Gorwa-Grauslund, M.F. (2020) Vanillin production in *Pseudomonas*: whole-genome sequencing of *Pseudomonas* sp. strain 9.1 and reannotation of *Pseudomonas putida* CalA as a vanillin reductase. *Appl Environ Microbiol* **86:** e02442-19.

Graf, N., and Altenbuchner, J. (2014) Genetic engineering of *Pseudomonas putida* KT2440 for rapid and high-yield production of vanillin from ferulic acid. *Appl Microbiol Biotechnol* **98:** 137–149.

Green, M.R., and Sambrook, J. (2018) The hanahan method for preparation and transformation of competent *Escherichia coli*: high-efficiency transformation. *Cold Spring Harb Protoc* **2018:** 183–191.

Johnson, C.W., Salvachúa, D., Rorrer, N.A., Black, B.A., Vardon, D.R., St. John, P.C., *et al*. (2019) Innovative chemicals and materials from bacterial aromatic catabolic pathways. *Joule* **3:** 1523–1537.

Kaulmann, U., Smithies, K., Smith, M.E.B., Hailes, H.C., and Ward, J.M. (2007) Substrate spectrum of ω-transaminase from *Chromobacterium violaceum* DSM30191 and its potential for biocatalysis. *Enzyme Microb Technol* **41:** 628–637.

Kelly, S.A., Pohle, S., Wharry, S., Mix, S., Allen, C.C.R., Moody, T.S., and Gilmore, B.F. (2018) Application of ω-transaminases in the pharmaceutical industry. *Chem Rev* **118:** 349–367.

Kohlstedt, M., Starck, S., Barton, N., Stolzenberger, J., Selzer, M., Mehlmann, K., *et al*. (2018) From lignin to nylon: Cascaded chemical and biochemical conversion using metabolically engineered *Pseudomonas putida*. *Metab Eng* **47:** 279–293.

Lian, J., Jin, R., and Zhao, H. (2016) Construction of plasmids with tunable copy numbers in *Saccharomyces cerevisiae* and their applications in pathway optimization and multiplex genome integration. *Biotechnol Bioeng* **113:** 2462–2473.

Lu, C.D., Itoh, Y., Nakada, Y., and Jiang, Y. (2002) Functional analysis and regulation of the divergent spuABC-DEFGH-spuI operons for polyamine uptake and utilization in *Pseudomonas aeruginosa* PAO1. *J Bacteriol* **184:** 3765–3773.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M.K., Geer, R.C., Gonzales, N.R., *et al*. (2020) CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* **48:** D265–D268.

Luengo, J.M., and Olivera, E.R. (2020) Catabolism of biogenic amines in *Pseudomonas* species. *Environ Microbiol* **22:** 1174–1192.

Masuo, S., Zhou, S., Kaneko, T., and Takaya, N. (2016) Bacterial fermentation platform for producing artificial aromatic amines. *Sci Rep* **6:** 1–9.

Mogheiseh, M., Karimian, R., and Khoshsefat, M. (2020) Vanillin-derived epoxy monomer for synthesis of bio-based epoxy thermosets: effect of functionality on thermal, mechanical, chemical and structural properties. *Chem Pap* **74:** 3347–3358.

Molnár, Z., Farkas, E., Lakó, Á., Erdélyi, B., Kroutil, W., Vértessy, B.G., *et al*. (2019) Immobilized whole-cell transaminase biocatalysts for continuous-flow kinetic resolution of amines. *Catalysts* **9:** 1–11.

Nelson, K.E., Weinel, C., Paulsen, I.T., Dodson, R.J., Hilbert, H., Martins dos Santos, V.A.P., *et al*. (2002) Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ Microbiol* **4:** 799–808.

Nikel, P.I., and de Lorenzo, V. (2018) *Pseudomonas putida* as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism. *Metab Eng* **50:** 142–155.

Oliphant, T.E. (2007) Python for scientific computing. *Comput Sci Eng* **9:** 10–20.

Patil, M.D., Grogan, G., Bommarius, A., and Yun, H. (2018) Recent advances in ω-transaminase-mediated biocatalysis for the enantioselective synthesis of chiral amines. *Catalysts* **8:** 254.

Patil, M.D., Yoon, S., Jeon, H., Khobragade, T.P., Sarak, S., Pagar, A.D., *et al*. (2019) Kinetic resolution of racemic amines to enantiopure (S)-amines by a biocatalytic cascade employing amine dehydrogenase and alanine dehydrogenase. *Catalysts* **9:** 254.

Pawar, S.V., Hallam, S.J., and Yadav, V.G. (2018) Metagenomic discovery of a novel transaminase for valorization of monoaromatic compounds. *RSC Adv* **8:** 22490–22497.

Pfennig, N., and Lippert, K.D. (1966) Über das Vitamin B12-Bedürfnis phototropher Schwefelbakterien. *Arch Mikrobiol* **55:** 245–256.

Salvachúa, D., Rydzak, T., Auwae, R., De Capite, A., Black, B.A., Bouvier, J.T., *et al*. (2020) Metabolic engineering of *Pseudomonas putida* for increased polyhydroxyalkanoate production from lignin. *Microb Biotechnol* **13:** 290–298.

Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., De Las Heras, A., *et al*. (2013) The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res* **41:** 666–675.

Simon, O., Klaiber, I., Huber, A., and Pfannstiel, J. (2014) Comprehensive proteome analysis of the response of *Pseudomonas putida* KT2440 to the flavor compound vanillin. *J Proteomics* **109:** 212–227.

Straathof, A.J.J., Wahl, S.A., Benjamin, K.R., Takors, R., Wierckx, N., and Noorman, H.J. (2019) Grand research challenges for sustainable industrial biotechnology. *Trends Biotechnol* **37:** 1042–1050.

Sudhakar Johnson, T., Ravishankar, G.A., and Venkataraman, L.V. (1996) Biotransformation of ferulic acid and vanillylamine to capsaicin and vanillin in immobilized cell cultures of *Capsicum frutescens*. *Plant Cell Tissue Organ Cult* **44:** 117–121.

Tufvesson, P., Lima-Ramos, J., Jensen, J.S., Al-Haque, N., Neto, W., and Woodley, J.M. (2011) Process considerations for the asymmetric synthesis of chiral amines using transaminases. *Biotechnol Bioeng* **108:** 1479–1493.

Tufvesson, P., Lima-Ramos, J., Nordblad, M., and Woodley, J.M. (2011) Guidelines and cost analysis for catalyst production in biocatalytic processes. *Org Process Res Dev* **15:** 266–274.

Wang, H., Pu, Y., Ragauskas, A., and Yang, B. (2019) From lignin to valuable products–strategies, challenges, and prospects. *Bioresour Technol* **271:** 449–461.

Weber, N., Gorwa-Grauslund, M., and Carlquist, M. (2014a) Exploiting cell metabolism for biocatalytic whole-cell transamination by recombinant *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* **98:** 4615–4624.

Weber, N., Gorwa-Grauslund, M., and Carlquist, M. (2017) Improvement of whole-cell transamination with *Saccharomyces cerevisiae* using metabolic engineering and cell pre-adaptation. *Microb Cell Fact* **16:** 1–12.

Weber, N., Ismail, A., Gorwa-Grauslund, M., and Carlquist, M. (2014b) Biocatalytic potential of vanillin aminotransferase from *Capsicum chinense*. *BMC Biotechnol* **14:** 1–6.

Weimer, A., Kohlstedt, M., Volke, D.C., Nikel, P.I., and Wittmann, C. (2020) Industrial biotechnology of *Pseudomonas putida*: advances and prospects. *Appl Microbiol Biotechnol* **104:** 7745–7766.

Wirth, N.T., Kozaeva, E., and Nikel, P.I. (2020) Accelerated genome engineering of *Pseudomonas putida* by I-SceI-mediated recombination and CRISPR-Cas9 counterselection. *Microb Biotechnol* **13:** 233–249.

Wu, H.L., Zhang, J.D., Zhang, C.F., Fan, X.J., Chang, H.H., and Wei, W.L. (2017) Characterization of four new distinct ω-transaminases from *Pseudomonas putida* NBRC 14164 for kinetic resolution of racemic amines and amino alcohols. *Appl Biochem Biotechnol* **181:** 972–985.

Wu, S., Liu, J., and Li, Z. (2017) Biocatalytic formal anti-markovnikov hydroamination and hydration of aryl alkenes. *ACS Catal* **7:** 5225–5233.

Wu, S., Zhou, Y., Wang, T., Too, H.P., Wang, D.I.C., and Li, Z. (2016) Highly regio- and enantioselective multiple oxy- and amino-functionalizations of alkenes by modular cascade biocatalysis. *Nat Commun* **7:** 11917.

Xu, Z., Lei, P., Zhai, R., Wen, Z., and Jin, M. (2019) Recent advances in lignin valorization with bacterial cultures: Microorganisms, metabolic pathways, and bio-products. *Biotechnol Biofuels* **12:** 1–19.

Zhou, Y., Wu, S., Mao, J., and Li, Z. (2018) Bioproduction of benzylamine from renewable feedstocks via a nine-step artificial enzyme cascade and engineered metabolic pathways. *Chemsuschem* **11:** 2221–2228.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Growth profiles of *P. putida* KT2440 cultivated for 48 hours at 30°C and 180 rpm in shake flasks with 50 ml M9 medium. Vanillin, vanillylamine and glucose were utilized as carbon sources. Error bars indicate ± SD of three biological replicates.

Fig. S2. Growth profiles of *P. putida* KT2440 overexpressing ATAs encoding genes on M9 medium supplemented with a range of vanillylalmine concentrations (0-25 mM).

Fig. S3. Whole-cell bioconversion of vanillylamine using growing-cells of *P. putida* KT2440 strains over-expressing different ATAs encoding genes.

Fig. S4. Effect of the antibiotic in the selection of positive clones containing *in vivo* transaminase activity against vanillylamine.

Fig. S5. Whole-cell bioconversion of vanillylamine using growing-cells of metabolically engineered *P. putida* GN442ΔPP_2426 strains over-expressing different ATAs encoding genes.

Fig. S6. Whole-cell bioconversion of vanillin to vanillylamine using resting-cells of metabolically engineered *P.putida* GN442ΔPP_2426 strains over-expressing different ATAs encoding genes.

Fig. S7. Whole-cell bioconversion of vanillin to vanillylamine without amine donor using growing-cells of metabolically

engineered *P. putida* strains as *P.putida* GN442ΔPP_2426 strains over-expressing different ATAs encoding genes.

Fig. S8. SDS-PAGE gel from cell crude extract of GN442ΔPP_2426 strains overexpressing different ATA encoding genes.

Table S1. BLASTp results for each ATA utilized as query sequences against *P. putida* KT2440 database.

Table S2. List of plasmids used in this work.

Table S3. List of strains used in this work.

Table S4. List of PCR primers utilized in this work.

Sequence S1. Nucleotide sequence of the *AlaDH* gene from *B. subtilis* 168.

Supporting information 1 for research article: supplementary tables and sequences

Metabolic engineering of *Pseudomonas putida* for production of vanillylamine from lignin-derived substrates

João Heitor Colombelli Manfrão-Netto[1], Fredrik Lund[2], Nina Muratovska[2], Elin M. Larsson[2,3], Nádia Skorupa Parachin[1], Magnus Carlquist[2,*]

[1]Grupo Engenharia de Biocatalisadores, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília-DF, Brazil

[2]Division of Applied Microbiology, Department of Chemistry, Faculty of Engineering, Lund University, PO Box 124, 221 00 Lund, Sweden

[3]Department of Bioengineering, California Institute of Technology, 1200 East California Blvd, Pasadena, CA, 91125, USA

*Correspondence: magnus.carlquist@tmb.lth.se

Supporting information 1 for research article: supplementary tables and sequences

Table S1. BLASTp results for each ATA utilized as query sequences against *P. putida* KT2440 database.

Table S2. List of plasmids used in this work.

Table S3. List of strains used in this work.

Table S4. List of PCR primers utilized in this work.

Sequence S1. Nucleotide sequence of the *AlaDH* gene from *B. subtilis* 168

Table S1 . BLASTp results for each ATA utilized as query sequences against the protein database for *P. putida* KT2440.

Query ATA	GenBank Accession number	Locus tag	Pp-ATAs	Query cover (%)	Identity (%)	E-value	Description
Pp-SpuC (PDB id: 6HX9)	AAN70747.1	PP_5182 (*spuC-II*)	Pp-SpuC-II	99	98.01	0.0	Polyamine:pyruvate transaminase
	AAN67793.1	PP_2180 (*spuC-I*)	Pp-SpuC-I	99	84.99	0.0	Polyamine:pyruvate transaminase
	AAN68196.1	PP_2588	Pp-ATA	98	51.89	2e-170	Aminotransferase, class III
Cv-ATA (PDB id: 4BA5	AAN70747.1	PP_5182 (*spuC-II*)	Pp-SpuC-II	96	57.82	0.0	Polyamine:pyruvate transaminase
	AAN67793.1	PP_2180 (*spuC-I*)	Pp-SpuC-I	97	56.05	0.0	Polyamine:pyruvate transaminase
	AAN68196.1	PP_2588	Pp-ATA	94	54.23	1e-169	Aminotransferase, class III

Supporting information 1 for research article: supplementary tables and sequences

Table S2.List of plasmids used in this work

Plasmid	**Description***	**Reference**
pSEVA424	Broad host range plasmid regulated by expression promoter system *lacIq*-P*trc* under IPTG induction; Resistance marker Sm; Replication oriV(RK2)	Silva-Rocha et al. 2013
pUC57-AlaDH	pUC57-derived plasmid bearing the synthetic *ald* gene encoding AlaDH from *Bacillus subtilis* 168 strain	This study
pNW10	Yeast plasmid bearing the *Cc-ATA* gene from *Capsicum chinense*	Weber et al., 2017
pNW12	Yeast plasmid bearing the *Cv-ATA* gene from *Chromobacterium violaceum*	Weber et al., 2017
pJH001	Constructed plasmid derived from pSEVA424 backbone containing *Cc-ATA* gene under IPTG control	This study
pJH002	Constructed plasmid derived from pSEVA424 backbone containing *Cv-ATA* gene under IPTG control	This study
pJH004	Constructed plasmid derived from pSEVA424 backbone containing both *Cv-ATA* and *AlaDH* gene under IPTG regulation	This study
pNM011	Constructed plasmid derived from pSEVA424 backbone containing *Pp-SpuC-II* gene under IPTG regulation	This study
pNM012	Constructed plasmid derived from pSEVA424 backbone containing *Pp-SpuC-I* gene under IPTG regulation	This study
pNM013	Constructed plasmid derived from pSEVA424 backbone containing *Pp-ATA* gene under IPTG regulation	This study

*Antibiotic marker: Sm: Streptomycin.

Table S3. List of strains used in this work

Strain	Description*	Sourve or reference
Escherichia coli DH5α	Strain for plasmid maintenance and replication F- ϕ*dlacZ* Δ*M*15 Δ*lacZYA-argF U169 deoR supE44 hsdR17 recA1 endA1 gyrA96 thi-1 relA1*	Taylor *et al*. (1993)
Pseudomonas putida KT2440	Wild-type strain, derived from P. putida mt-2 (Worsey and Williams, 1975) cured of the TOL plasmid pWW0	DSM 6125 Regenhardt *et al*. (2002)
GN442ΔPP_2426	*P. putida* vanillin-accumulator strain Δ*upp* ΔPP_0166-0168 Δ*vdh* ΔPP_3827-3832 ΔPP_2680 ΔPP_0545 ΔPP_1948 *lacIq-Ptac-ech-fcs* ΔPP_2426	García-Hidalgo *et al.*, 2020
TMB-JH001	KT2440 carrying the plasmid pJH001 overexpressing *Cc-ATA* gene under IPTG induction, Sm^R	This study
TMB-JH002	KT2440 carrying the plasmid pJH002 overexpressing *Cv-ATA* gene under IPTG induction, Sm^R	This study
TMB-JH003	GN442ΔPP_2426 carrying the plasmid pJH001 overexpressing *Cc-ATA* gene under IPTG induction, Sm^R	This study
TMB-JH004	GN442ΔPP_2426 carrying the plasmid pJH002 overexpressing *Cv-ATA* gene under IPTG induction, Sm^R	This study
TMB-JH006	GN442ΔPP_2426 carrying the plasmid pJH004 overexpressing both *Cv-ATA* and *AlaDH* genes under IPTG induction, Sm^R	This study
TMB-NM011	KT2440 carrying the plasmid pNM011 overexpressing *Pp-SpuC-II* gene under IPTG induction, Sm^R	This study
TMB-NM012	KT2440 carrying the plasmid pNM012 overexpressing *Pp-Spuc-I* gene under IPTG induction, Sm^R	This study
TMB-NM013	KT2440 carrying the plasmid pNM013 overexpressing *Pp-ATA* gene under IPTG induction, Sm^R	This study
TMB-NM014	GN442ΔPP_2426 carrying the plasmid pNM011 overexpressing *Pp-SpuC-II* gene under IPTG induction, Sm^R	This study
TMB-NM015	GN442ΔPP_2426 carrying the plasmid pNM012 overexpressing *Pp-SpuC-I* gene under IPTG induction, Sm^R	This study
TMB-NM016	GN442ΔPP_2426 carrying the plasmid pNM013 overexpressing *Pp-ATA* gene under IPTG induction, Sm^R	This study

*Antibiotic marker: Sm: Streptomycin.

Table S4. List of PCR primers utilized in this work

Primer name	Gene to amplify	Orientation	[1]Restriction enzyme 5'	[2]Primer sequence 5'→ 3'
Fw_CC-TA	*Cc-ATA*	Forward	XbaI	TCG*TCTAGA***AGGAGGAAAAAC**AT**ATGGCAAACATTACAAACG
Rv_ CC-TA	*Cc-ATA*	Reverse	PsTI	AAAA*CTGCAG*TTATTGCTTTTGGGACTTCA
Fw_CV-TA	*Cv-ATA*/	Forward	SacI	AAC*GAGCTC***AGGAGGAAAAAC**AT**ATGCAAAAACAAAGAACAAC
Rv_ CV-TA	*Cv-ATA*	Reverse	BamHI	CGC*GGATCC*TTATGCTAAACCTCTAGCCTT
Fw_PP-TA1	*Pp-SpuC-II*	Forward	SacI	AAT*GAGCTC***AGGAGGAAAAAC**AT**ATGAGCGTCAACAACCCG
Rv_PP-TA1	*Pp-SpuC-II*	Reverse	XbaI	CCCC*TCTAGA*TTATTGAATCGCCTCAAGGG
Fw_PP-TA2	*Pp-Spuc-I*	Forward	SacI	AAT*GAGCTC***AGGAGGAAAAAC**AT**ATGAGTGAACAGAATTCGCA
Fw_PP-TA2	*Pp-Spuc-I*	Reverse	XbaI	CCC*TCTAGA*TTACCGAACAGCCTCATAGG
Fw_PP-TA3	*Pp-ATA*	Forward	SacI	AAT*GAGCTC***AGGAGGAAAAAC**AT**ATGAACGCGCCTTTCGCC
Rv_PP-TA3	*Pp-ATA*	Reverse	XbaI	ACA*TCTAGA*TTACAGCTTGCCGACCAGCC

[1]In *italic* the recognizing sequence for the respective enzyme.

[2]In **bold** the RBS sequence utilized

Sequence S1. Nucleotide sequence of the synthetic *AlaDH* gene from *B. subtilis* 168 (Genbank ID: 936557, Swiss-Prot: Q08352). The sites for the restriction enzymes *PsTI* (5`) and *SpeI* (3`) to clone *AlaDH* gene into the vector pSEVA424 are shown in red and the ribosome binding site (RBS) in italic. The start codon and the stop codon of the open reading frame (ORF) from *AlaDH* gene are shown in bold.

CTGCAG*AGGAGGAAAAACCC***ATG**ATCATCGGTGTCCCAAAAGAGATTAAGAACAAC
GAAAACAGAGTTGCTTTGACTCCAGGTGGTGTTTCTCAATTGATTTCTAACGGTCAC
AGAGTTTTGGTTGAAACTGGTGCTGGTTTAGGTTCTGGTTTTGAAAATGAAGCTTAC
GAATCTGCTGGTGCCGAAATTATTGCTGATCCAAAACAAGTTTGGGATGCCGAAATG
GTTATGAAGGTAAAAGAACCATTGCCAGAGGAATACGTCTACTTCAGAAAAGGTTT
GGTCTTGTTCACCTACTTGCATTTGGCTGCTGAACCAGAATTGGCTCAAGCTTTGAA
AGATAAGGGTGTTACTGCTATTGCTTACGAAACTGTTTCAGAAGGTAGAACCTTGCC
ATTATTGACTCCAATGTCTGAAGTTGCTGGTAGAATGGCTGCTCAAATTGGTGCTCA
ATTTTTGGAAAAACCCAAAGGTGGTAAGGGTATTTTGTTGGCTGGTGTTCCAGGTGT
TTCTAGAGGTAAGGTTACTATTATTGGTGGTGGTGTTGTAGGTACTAATGCTGCTAA
AATGGCTGTTGGTTTGGGTGCTGATGTTACCATTATAGATTTGAACGCCGACAGATT
GAGACAATTGGATGATATTTTCGGTCACCAGATCAAGACGTTGATCTCTAATCCAGT
TAACATTGCTGATGCTGTTGCTGAAGCTGATTTGTTGATTTGCGCTGTTTTAATTCCA
GGTGCTAAAGCTCCAACTTTGGTTACTGAAGAAATGGTCAAACAAATGAAGCCAGG
TTCCGTTATAGTTGATGTTGCTATAGATCAAGGTGGTATCGTTGAAACCGTTGATCAT
ATTACTACCCATGATCAACCCACTTACGAAAAACATGGTGTTGTTCATTACGCTGTT
GCTAATATGCCAGGTGCTGTTCCAAGAACTTCTACTATTGCATTGACTAACGTTACT
GTTCCATACGCCTTGCAAATTGCTAACAAAGGTGCAGTTAAGGCTTTGGCTGATAAT
ACTGCTTTGAGAGCTGGTTTGAATACCGCTAATGGTCATGTTACTTATGAAGCTGTT
GCAAGAGATTTGGGTTACGAATATGTTCCAGCTGAAAAAGCCTTGCAAGATGAATCT
TCTGTTGCTGGTGCT**TGA**ACTAGT

References

García-Hidalgo, J., Brink, D.P., Ravi, K., Paul, C.J., Lidénb, G., and Gorwa-Grauslund, M.F. (2020) Vanillin Production in *Pseudomonas*: Whole-genome sequencing of *Pseudomonas* sp. strain 9.1 and reannotation of *Pseudomonas putida calA* as a vanillin reductase. *Appl Environ Microbiol* **86**:.

Graf, N. and Altenbuchner, J. (2014) Genetic engineering of *Pseudomonas putida* KT2440 for rapid and high-yield production of vanillin from ferulic acid. *Appl Microbiol Biotechnol* **98**: 137–149.

Regenhardt, D., Heuer, H., Heim, S., Fernandez, D.U., Strömpl, C., Moore, E.R.B., and Timmis, K.N. (2002) Pedigree and taxonomic credentials of *Pseudomonas putida* strain KT2440. *Environ Microbiol* **4**: 912–915.

Taylor, R.G., Walker, D.C., and Mclnnes, R.R. (1993) E.coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Res* **21**: 1677–1678.

Worsey, M.J. and Williams, A.P. (1975) Metablism of toluene and xylenes by *Pseudomonas putida* (*arvilla*) mt 2: evidence for a new function of the TOL plasmid. *J Bacteriol* **124**: 7–13.

Metabolic engineering of *Pseudomonas putida* for production of vanillylamine from lignin-derived substrates

João Heitor Colombelli Manfrão-Netto[1], Fredrik Lund[2], Nina Muratovska[2], Elin M. Larsson[2,3], Nádia Skorupa Parachin[1], Magnus Carlquist[2,*]

[1]Grupo Engenharia de Biocatalisadores, Instituto de Ciências Biológicas, Universidade de Brasília, Brasília-DF, Brazil

[2]Division of Applied Microbiology, Department of Chemistry, Faculty of Engineering, Lund University, PO Box 124, 221 00 Lund, Sweden

[3]Department of Bioengineering, California Institute of Technology, 1200 East California Blvd, Pasadena, CA, 91125, USA

*Correspondence: magnus.carlquist@tmb.lth.se

Figure S1. Growth profiles of *P. putida* KT2440 cultivated for 48 hours at 30 °C and 180 rpm in shake flasks with 50 ml M9 medium. Vanillin, vanillylamine and glucose were utilized as carbon sources. Error bars indicate ± SD of three biological replicates.

Figure S2. Growth profiles of *P. putida* KT2440 overexpressing ATAs encoding genes on M9 medium supplemented with a range of vanillyalmine concentrations (0-25 mM).

Figure S3. Whole-cell bioconversion of vanillylamine using growing-cells of *P. putida* KT2440 strains over-expressing different ATAs encoding genes.

Figure S4. Effect of the antibiotic in the selection of positive clones containing *in vivo* transaminase activity against vanillylamine.

Figure S5. Whole-cell bioconversion of vanillylamine using growing-cells of metabolically engineered *P. putida* GN442ΔPP_2426 strains over-expressing different ATAs encoding genes.

Figure S6. Whole-cell bioconversion of vanillin to vanillylamine using resting-cells of metabolically engineered *P.putida* GN442ΔPP_2426 strains over-expressing different ATAs encoding genes.

Figure S7. Whole-cell bioconversion of vanillin to vanillylamine without amine donor using growing-cells of metabolically engineered P. putida strains as *P.putida* GN442ΔPP_2426 strains over-expressing different ATAs encoding genes.

Figure S8.SDS-PAGE gel from cell crude extract of GN442ΔPP_2426 strains overexpressing different ATA encoding genes.

Figure S1. Growth profiles of *P. putida* KT2440 cultivated for 48 hours at 30 °C and 180 rpm in shake flasks with 50 ml M9 medium. Vanillin, vanillylamine and glucose were utilized as carbon sources. Error bars indicate ± SD of three biological replicates.

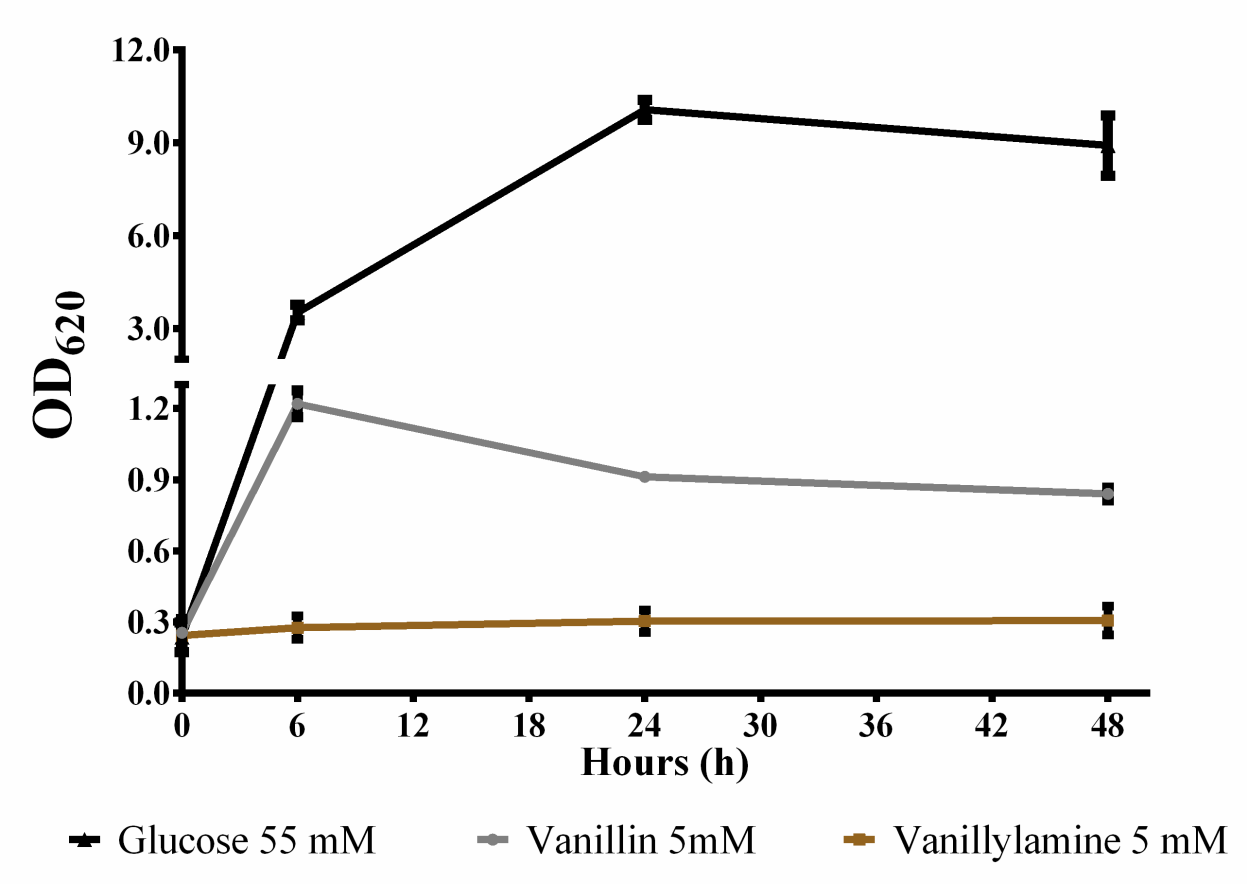


Figure S2. Growth profiles of *P. putida* KT2440 overexpressing ATAs encoding genes *P. putida* KT2440 wild-type (A), TMB-NM012 (*Pp-SpuC*-I) (B), and TMB-NM013 (*Pp-ATA*) (C), displayed no significant growth in M9 medium supplemented with a range of vanillylamine concentrations (0-25 mM) as the sole carbon source for 24 hours at 30C. In a pilot experiment performed in shake flasks, the strain overexpressing *Cc-ATA* showed no significant growth on 5 mM of VA as the sole carbon source and it was not used for microplate assays.

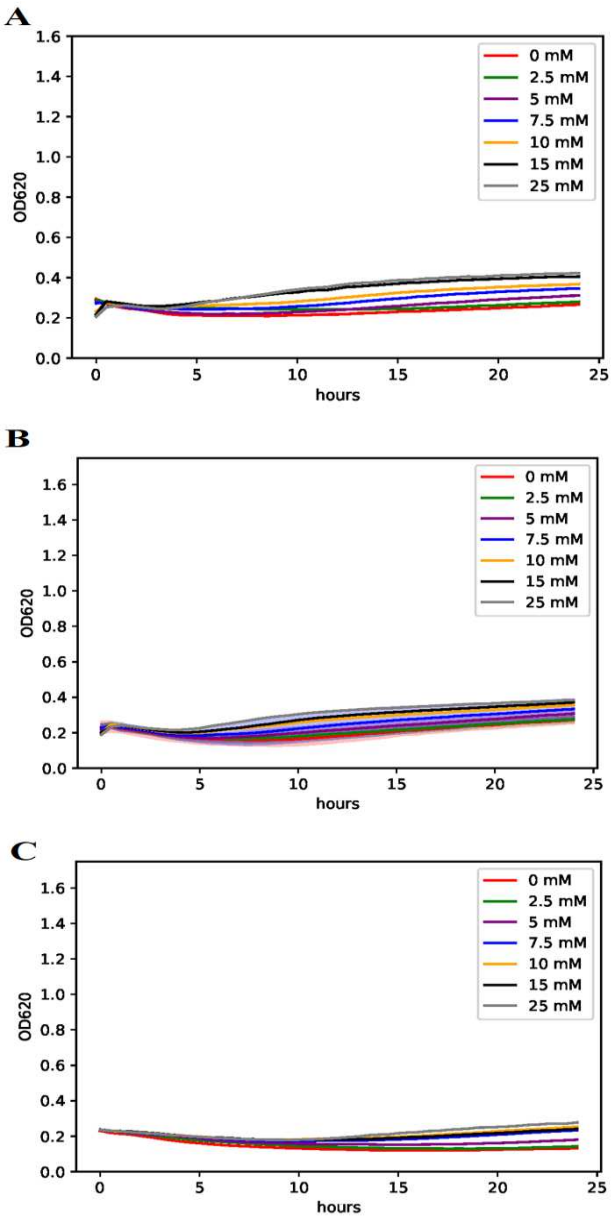


Figure S3. Whole-cell bioconversion of vanillylamine using growing-cells of *P. putida* KT2440 strains over-expressing (A) TMB-NM011 (Pp-SpuC-II), (B) TMB-NM012 (Pp-SpuC-I), (C) TMB-NM013 (Pp-ATA), (D) TMB-JH001 (Cc-ATA), (E) TMB-JH002 (Cv-ATA) or (F) Wild-type KT2440 (negative control). The bioconversions were performed for 24 hours at 30 °C and 180 rpm in M9 medium supplemented with 5 mM vanillylamine, 10 g/L glucose and 1 mM IPTG. For the Wild-type strain streptomycin was omitted. Error bars indicate ± SD of two biological replicates.

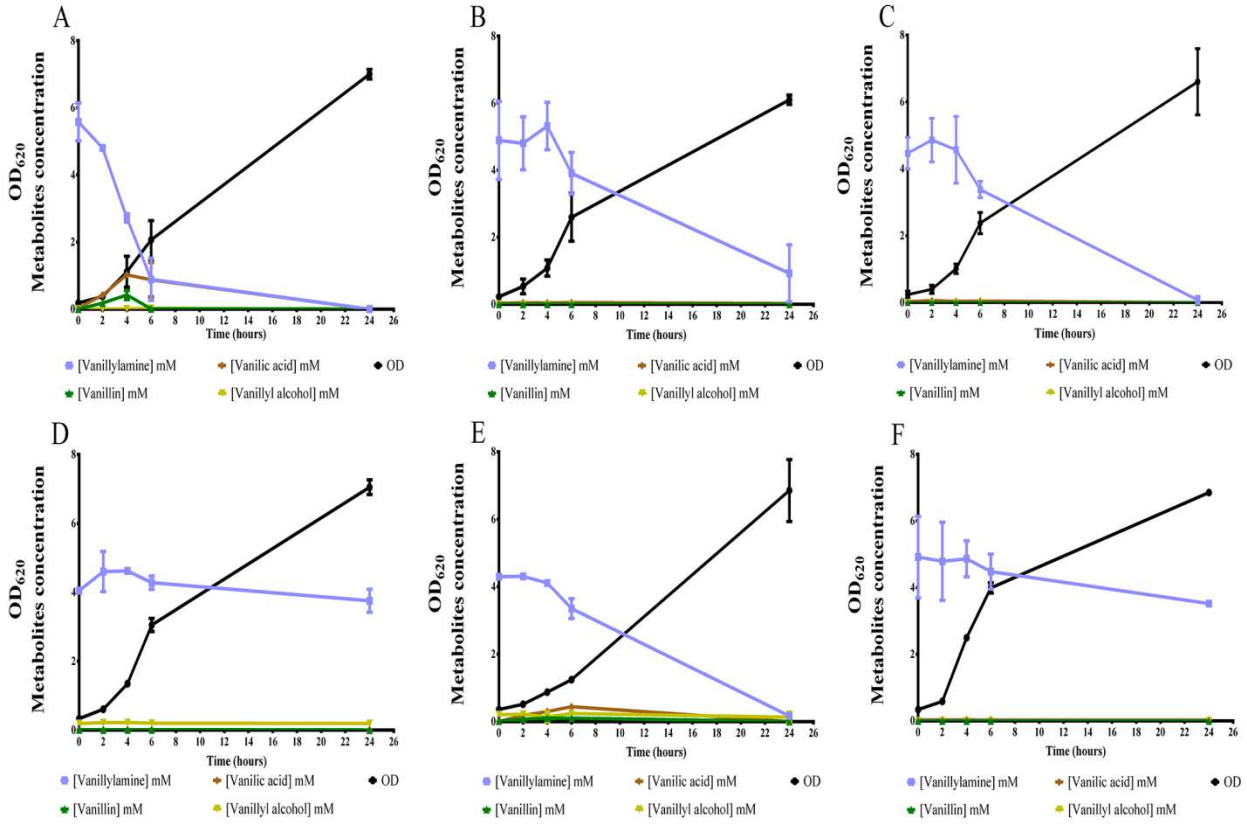


Figure S4. M9 plates after 3 days incubated at 30° C with 5 mM of VA as the sole carbon source and IPTG (1 mM) to induce the *Cv-ATA* gene expression. Streptomycin (100 µg/ml) was added to the plates in A and B while C and D only VA was utilized for the selection. A and C: *P. putida* KT2440 WT without pSEVA424 plasmid. B and D cells, TMB-JH002 cells expressing *Cv-ATA* gene.

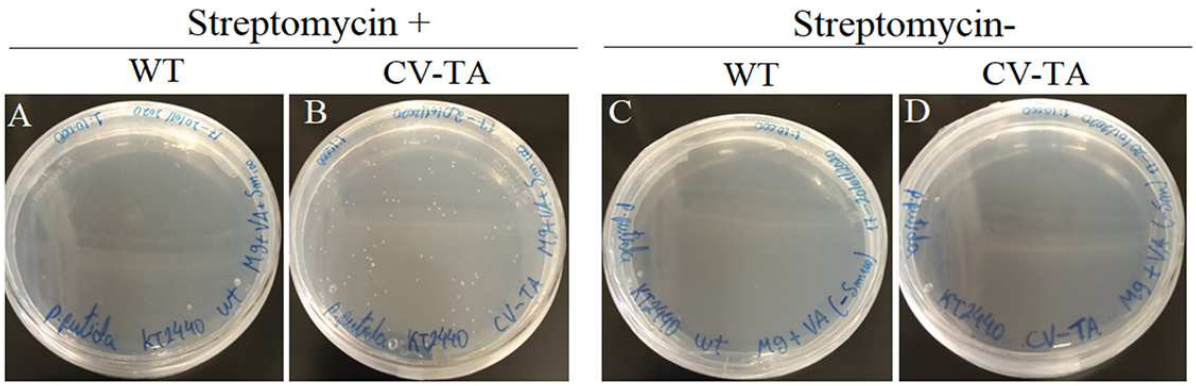


Figure S5. Whole-cell bioconversion of VA to vanillin using growing-cells of metabolically engineered *P.putida* GN442ΔPP_2426 strains as biocatalysts. The cells were cultured for 24 hours at 30 °C and 180 rpm in M9 medium with 5 mM of vanillylamine as the substrate for the bioconversion and 10 g/L of glucose as the carbon source. IPTG (1 mM) was added to induce the genes expression and 100 µg/ml streptomycin was added for plasmid maintenance. (A) TMB-NM014 (Pp-SpuC-II), (B) TMB-NM015 (Pp-SpuC-I), (C) TMB-NM016 (Pp-ATA),(D) TMB-JH003 (Cc-ATA), (E) TMB-JH004 (Cv-ATA) and (F) *P. putida* GN442ΔPP_2426 harboring the empty plasmid. This experiment was performed in two biological replicates, excluding TMB-JH003 (Cc-ATA).

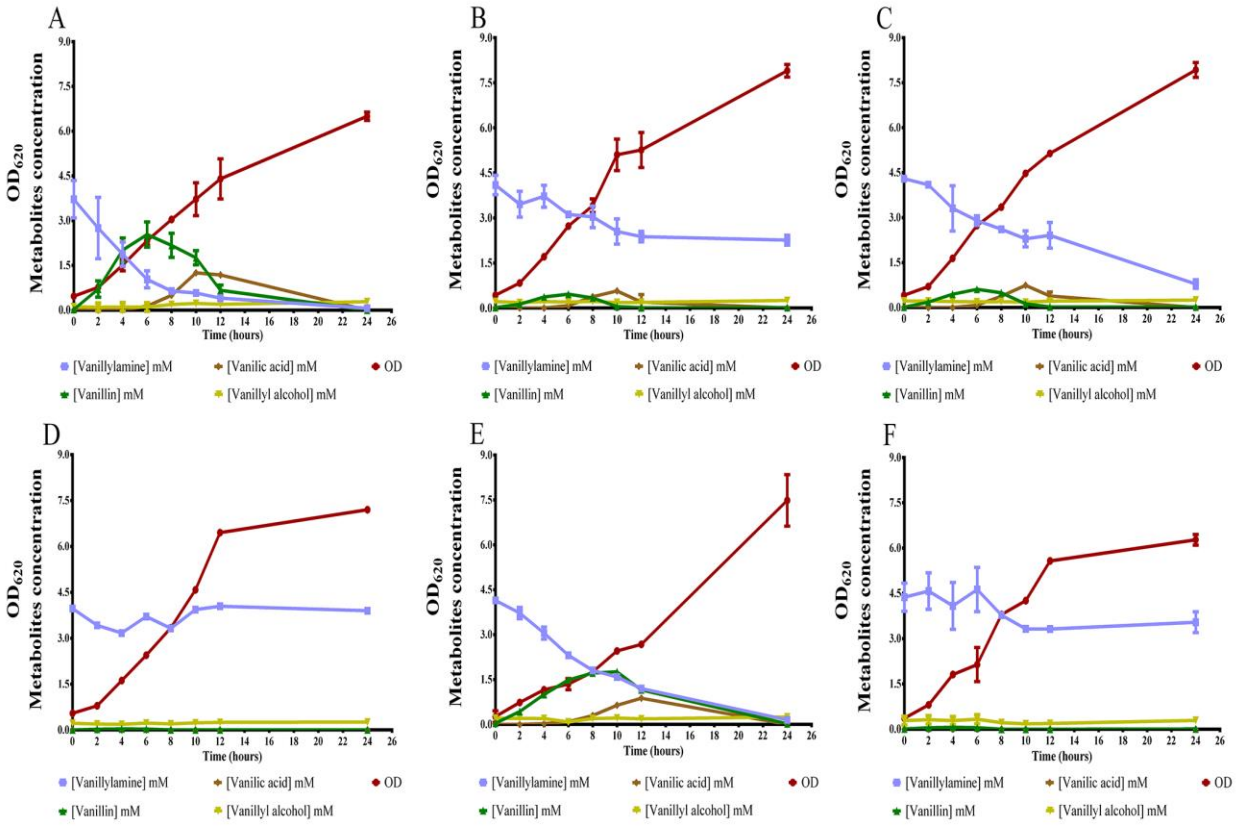


Figure S6. Whole-cell bioconversion of vanillin to vanillylamine using resting-cells of metabolically engineered *P.putida* GN442ΔPP_2426 strains as biocatalysts (OD_{620} =3). The assays were performed for 24 hours at 30 °C and 180 rpm using sodium phosphate buffer without the addition of glucose. Vanillin (10 mM) was utilized as the substrate and 100 mM of alanine as the amine donor. (A) TMB-NM014 (Pp-SpuC-II), (B) TMB-NM015 (Pp-SpuC-I), (C) TMB-NM016 (Pp-ATA),(D) TMB-JH004 (Cv-TA) and (E) *P. putida* GN442ΔPP_2426 harboring the empty plasmid. Error bars indicate ± SD of two biological replicates. **F.** Correlation between Specific transaminase activity* and maximum production of VA from vanillin.

*Specific activity was quantified by following the initial rate of whole-cell conversion of vanillylamine in M9 medium supplemented with glucose, and in turn is directly correlated with cell growth with VA as sole carbon source.

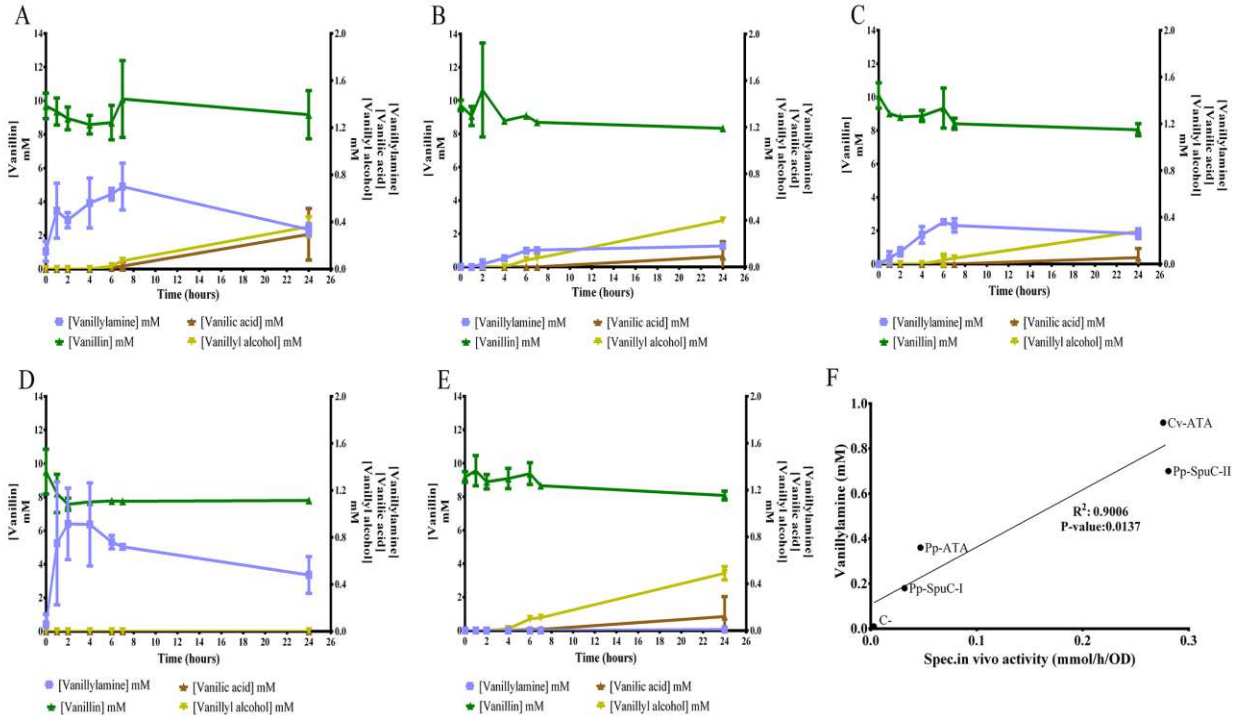


Figure S7. Whole-cell bioconversion of vanillin to VA without amine donor using growing-cells of metabolically engineered *P. putida* strains as biocatalysts. The cells were cultured for 24 hours at 30 °C and 180 rpm in M9 medium with 5 mM of vanillin as the substrate for the bioconversion and 10 g/L of glucose as the carbon source. IPTG (1 mM) was added to induce the genes expression, 100 µg/ml streptomycin was added for plasmid maintenance and the amine donor was omitted. (A) TMB-NM014 (Pp-SpuC-II), (B) TMB-NM015 (Pp-SpuC-I) or (C) TMB-NM016 (Pp-ATA), (D) TMB-JH004 (Cv-TA) and (E) *P. putida* GN442ΔPP_2426 harboring the empty plasmid. This experiment was performed in a single biological replicate.

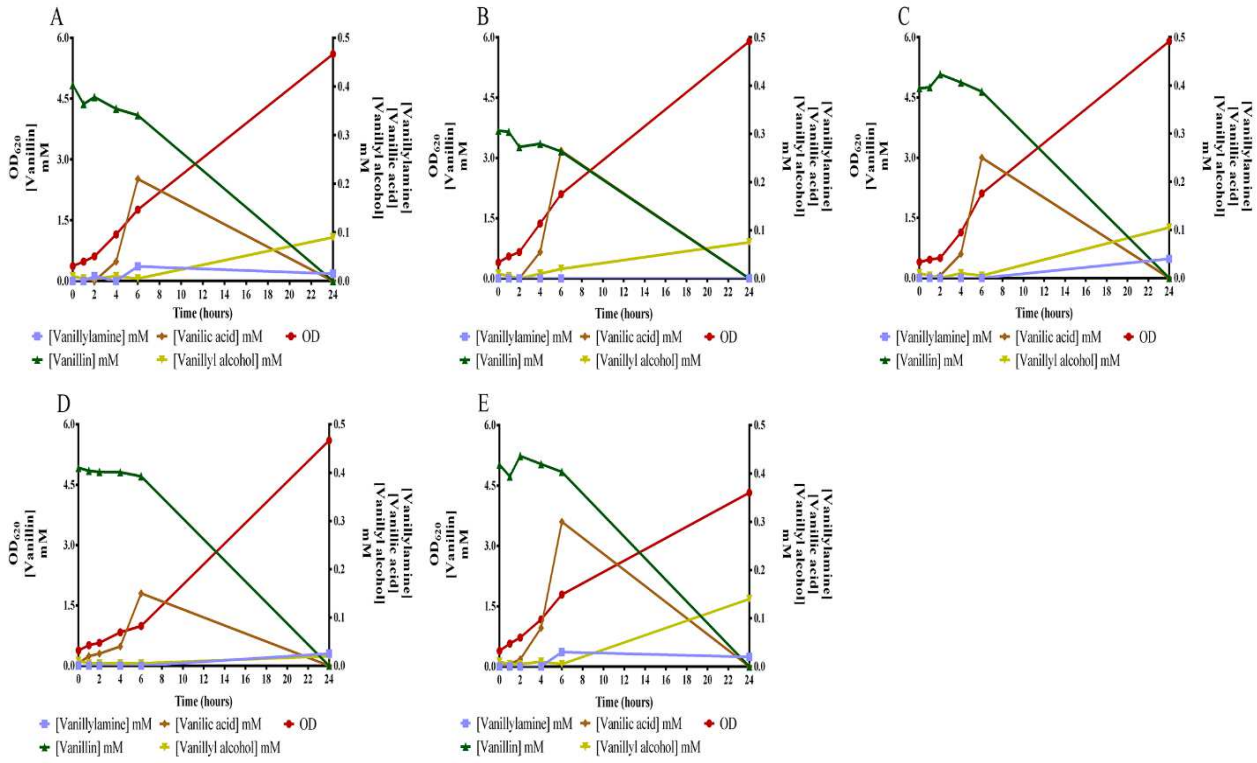
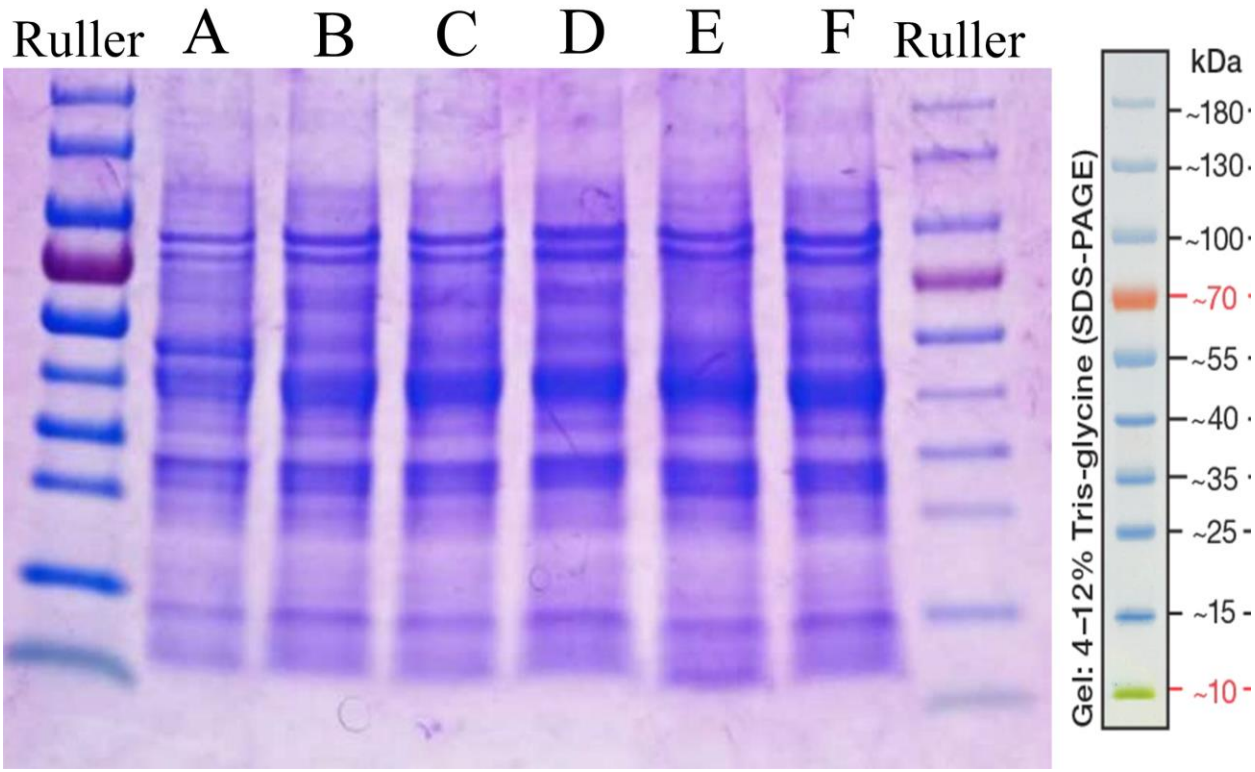


Figure S8. The pattern of proteins expressed by GN442ΔPP_2426 was assessed by SDS-PAGE analysis using cell crude extract. Pellets from bioconversion of VA to vanillin using growing cells of GN442ΔPP_2426 (experiment reported in Figure S5) after 24h were collected and frozen at -80 C. The proteins were extracted with YPER. Bradford assay was performed at 595 nm in a microplate and the concentration of total protein was estimated using a standard curve from BSA. Sample volume for SDS-PAGE was calculated so that 10µg of protein was loaded for each sample onto a 4-20% polyacrylamide Mini-PROTEAN TGX gel. (A) TMB-NM014 (Pp-SpuC-II; Mass: 49.88 kDA), (B) TMB-NM015 (Pp-SpuC-I; Mass: 49.64 kDA), (C) TMB-NM016 (Pp-ATA, Mass: 50.86 kDA),(D) TMB-JH003 (Cc-ATA, Mass: 50.74 kDA), (E) TMB-JH004 (Cv-ATA; Mass: 51.22 kDA) and (F) *P. putida* GN442ΔPP_2426 harboring the empty plasmid.



Genetic strategies for improving Hyaluronic Acid production in recombinant bacterial culture

Journal:	*Applied Microbiology*
Manuscript ID	JAM-2020-2600.R1
Journal Name:	Journal of Applied Microbiology
Manuscript Type:	JAM - Review Article
Date Submitted by the Author:	24-Apr-2021

Complete List of Authors:	Manfrão Netto, João Heitor; University of Brasilia, Cell Biology Queiroz, Enzo; Universidade de Brasilia, Departamento de Biologia Celular Junqueira, Ana; Universidade de Brasilia, Departamento de Biologia Celular Gomes, Antônio ; Universidade de Brasilia, Departamento de Biologia Celular de Morais, Daniel; Universidade Católica de Brasília, Pós-Graduação em Ciências Genômicas e Biotecnologia Paes, Hugo; Universidade de Brasilia, Clinical Medicine Division Parachin, Nádia; Universidade de Brasilia, Departamento de Biologia Celular

Key Words:	Biopolymers, Metabolic processes, Molecular genetic, Bacillus, Lactococcus