ENZO BENTO QUEIROZ

Metabolic engineering of the yeast *Yarrowia lipolytica* for the monoterpene geraniol heterologous production

Revised version

Dissertation presented to the Polytechnic School of the University of São Paulo (Escola Politécnica da Universidade de São Paulo) in partial fulfillment of the requirements for the Degree of Master in Sciences from the Graduate Program in Chemical Engineering.

Area: Chemical Engineering

Supervisor: Prof. Dr. Thiago Olitta Basso

São Paulo
2023

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Este exemplar foi revisado e corrigido em relação à versão original, sob responsabilidade única do autor e com a anuência de seu orientador.

São Paulo, 28 de Agosto de 2023.

Assinatura do autor Enzo Bento Queiroz

Assinatura do orientador Thiago Olitta Basso.

Catalogação-na-publicação

Queiroz, Enzo Bento

Metabolic engineering of the yeast Yarrowia lipolytica for the monoterpene geraniol heterologous production / E. B. Queiroz -- versão corr. -- São Paulo, 2023.

118 p.

Dissertação (Mestrado) - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia Química.

1.Biotecnologia 2.Bioprocessos 3.Engenharia Genética 4.Terpenos 5.Fungos I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia Química II.t.

Name: Enzo Bento Queiroz

Title: Metabolic engineering of the yeast *Yarrowia lipolytica* for the monoterpene geraniol heterologous production

Dissertation presented to the Polytechnic School of the University of São Paulo (Escola Politécnica da Universidade de São Paulo) in partial fulfillment of the requirements for the Degree of Master in Sciences from the Graduate Program in Chemical Engineering.

Approved in: June 28th, 2023

Approved by:

Prof. PhD.: Thiago Olitta Basso
Institution: University of São Paulo, Department of Chemical Engineering of the Polytechnic School

Prof. PhD.: Luiziana Ferreira da Silva
Institution: University of São Paulo, Biomedical Sciences Institute

Prof. PhD.: Andreas Karoly Gombert
Institution: University of Campinas, School of Food Engineering

Acknowledgments

This work has received support from the São Paulo Research Foundation: master's internship grant #2021/02155-8 and *Auxílio à pesquisa - Projeto inicial* grant #2022/03052-0, São Paulo Research Foundation (FAPESP).

The institution *Amigos da Poli* also supports the project which this work is part of (*Edital 2022 – Engenharia metabólica e evolutiva de leveduras não-convencionais para produção de terpenos bioativos*).

This study and the Chemichal Engineering Graduate Program from the University of São Paulo Polytechnic School was also financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil* (CAPES), Finance Code 001.

Abstract

QUEIROZ, E. B. **Metabolic engineering of the yeast *Yarrowia lipolytica* for the monoterpene geraniol heterologous production**. 2023. Dissertação (Mestrado) – Programa de Pós-Graduação em Engenharia Química da Escola Politécnica, Universidade de São Paulo, São Paulo, 2023.

Geraniol is a monoterpene, a class of secondary metabolites of plants, used as a fragrance in various hygiene and cosmetic products. In addition, this molecule also has potential, for example, as an antimicrobial and antitumor drug and as a chemical precursor of other compounds relevant to industry and human health. Despite its growing market, obtainment of geraniol is restricted to the extraction of plant essential oils or chemical synthesis, which are processes with inefficient, expensive or polluting steps. In this context, the use of metabolic engineering for the production of terpenes by microorganisms is a promising alternative that has been explored by the scientific community. *Yarrowia lipolytica*, a non-conventional oleaginous yeast, has attributes that can be exploited for the sustainable and efficient production of geraniol. These characteristics include tolerance to organic compounds and high levels of cytosolic acetyl-CoA, the molecule that initiates the monoterpene biosynthesis pathway. Thus, the present project aimed to develop a *Y. lipolytica* strain engineered for the biosynthesis of geraniol. For this, the *Y. lipolytica* strain to be transformed was analyzed for its growth and cellular morphology and its growth physiology was studied. Then, the plasmids required for geraniol synthase (GES) expression and its integration into the *Y. lipolytica* genome were replicated and confirmed. Finally, the CRISPR-Cas9 genome editing system was applied for the expression of the GES enzyme from *Catharanthus roseus* in the *Y. lipolytica* strain with modifications for the production of monoterpenes. Integration of expression cassette at the target *locus* was confirmed with diganosis PCR with genomic DNA as template. Production of 35 ± 10 mg L^{-1} geraniol (which represents a yield of 0.08 % C-mol geraniol C-mol^{-1} glucose) by the genetically modified BYa3105 strain was verified via GC-MS. Therefore, the basis for a bioprocess of geraniol production have been established and may be further explored in the next stages of the project. This is the first known report of *de novo* geraniol biosynthesis by *Y. lipolytica*.

Keywords: Yeast, Metabolic engineering, Terpene, Bioprocess engineering, Geraniol, Yarrowia

Resumo

QUEIROZ, E. B. **Engenharia metabólica da levedura *Yarrowia lipolytica* para a produção heteróloga do monoterpeno geraniol**. 2023. Dissertação (Mestrado) – Programa de Pós-Graduação em Engenharia Química da Escola Politécnica, Universidade de São Paulo, São Paulo, 2023.

Geraniol é um monoterpeno, classe de metabólitos secundários de plantas, usado como fragrância em diversos produtos de higiene e cosméticos. Além disso, essa molécula apresenta potencial como fármaco antimicrobiano e antitumoral, por exemplo, e é precursor químico de outros compostos relevantes para a indústria e saúde humana. Apesar de seu mercado crescente, a obtenção de geraniol se restringe à extração de óleos essenciais vegetais ou à síntese química, processos com etapas pouco eficientes, caras ou poluentes. Nesse contexto, o uso da engenharia metabólica para a produção de terpenos por microrganismos é uma alternativa promissora que tem sido explorada no meio científico. *Yarrowia lipolytica*, uma levedura oleaginosa não convencional, apresenta atributos que podem ser explorados para a produção sustentável e eficiente de geraniol. Essas características incluem tolerância a compostos orgânicos e altos níveis de acetil-CoA citosólico, molécula que inicia a via de biossíntese do monoterpeno. Assim, o presente projeto teve como objetivo o desenvolvimento de linhagem de *Y. lipolytica* capaz de realizar a biossíntese de geraniol. Para isso, a linhagem de *Y. lipolytica* que seria transformada foi analisada quanto ao seu crescimento e morfologia celular e sua fisiologia de crescimento foi estudada. Então, os plasmídeos necessários para expressão da geraniol sintase (GES) e sua integração no genoma de *Y. lipolytica* foram replicados e confirmados. Finalmente, foi aplicado o método de CRISPR-Cas9 para a expressão da enzima GES de *Catharanthus roseus* na linhagem de *Y. lipolytica* com modificações para a produção de monoterpenos. Integração do cassete foi confirmada por PCR com DNA genômico como molde. Produção de 35 ± 10 mg L^{-1} de geraniol (que representa um rendimento de 0.08 % C-mol C-mol^{-1}) pela linhagem BYa3105 geneticamente modificada foi verificada por GC-MS. Portanto, as bases para produção de geraniol foram estabelecidas e poderão ser adicionalmente exploradas nas próximas etapas do projeto. Esse é o primeiro relato conhecido da biossíntese *de novo* de geraniol por *Y. lipolytica*.

Palavras-chave: Levedura, Engenharia metabólica, Terpeno, Engenharia de bioprocessos, Geraniol, Yarrowia

List of Figures

List of Tables

List of acronyms and abbreviations

ADH	Alcohol-Dependent Hemiterpene	IPA	Isoprenoid Alcohol
AMP	Adenosine Monophosphate	IPCC	Intergovernmental Panel on Climate Change
ATP	Adenosine Triphosphate	IPP	Isopentenyl Diphosphate
BDM	Barth Defined Medium	IUP	Isopentenol Utilization Pathway
Cas9	CRISPR-associated 9	LB	Lysogeny Broth
CFU	Colony Forming Unit	LS	Limonene Synthase
CRISPR	Clustered and Regularly Insterspaced Short Palindromic Repeats	MDM	Matthäus Defined Medium
crRNA	CRISPR RNA	MEP	2-C-methyl-D-erythritol-4-

	phosphate
DMAPP	Dimethylallyl Diphosphate
DNA	Desoxirribonucleic Acid
dNTP	Deoxyribonucleoside Triphosphate
DSB	Double Strand Break
DXP	1-deoxy-D-xylulose-5-phosphate
EPA	Eicosapentaenoic Acid
FDA	Food and Drug Administration
FPP	Farnesyl Diphosphate
GC-MS	Gas Chromatography-Mass Spectrometry
GEM	Genome-scale Model
GES	Geraniol Synthase
GGPP	Geranylgeraniol Diphosphate
GPP	Geranyl Diphosphate
GRAS	Generally Recognized As Safe
HMG-CoA	3-hydroxy-3-methylglutaryl-CoA
HPLC	High-Performance Liquid Chromatography
HR	Homologous Repair
MIA	Monoterpenoid Indole Alkaloid
MVA	Mevalonate ((3R)-3,5-Dihydroxy-3-methylpentanoic acid)
NHEJ	Non-Homologous End Joining
OD_{600}	Optical Density at 600nm
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PK-PTA	Phosphoketolase-Phosphotransacetylase
RID	Refractive Index Detector
RNA	Ribonucleic Acid
sgRNA	Synthetic Guide Ribonucleic Acid
TALEN	Transcription Activator-Like Effector Nucleases
TCA	Tricarboxylic Acid
tracrRNA	Trans-Activating CRISPR RNA
USER	Uracil-Specific Excision Reagent
UV	Ultraviolet
YPD	Yeast extract, Peptone and Dextrose medium

List of symbols

bp	Base Pair(s)
DCW	Dry Cell Weight (g L^{-1})
g	Gram(s)
k	10^3
L	Liter(s)
^{o} C	Celsius
Pa	Pascal
RCF	Relative Centrifugal Force
rpm	Rotation(s) per minute
ton	Tonne or Tons

m	10^{-3}	u	Unified atomic mass
m/z	Mass-charge Ratio	μ	10^{-6}
m^3	Cubic meter(s)		

Summary

1 INTRODUCTION AND CONTEXT

During the last centuries, as a consequence of the rapid technological development, the average well-being of humanity has increased exponentially. Social inequalities still restrict most of this prosperity to the small rich fraction, but even the proportion of people living in absolute poverty has declined from 60 % in 1950 to 10 % in 2020. This is translated, for example, into the fast growth of the global population and into the unprecedented increase of life expectancy in the last 70 years. Therefore, it may seem that humanity has devised an efficient socioeconomic system capable of nourishing and sustaining life on Earth. However, this model of development enforced nowadays does not consider the welfare of the natural systems in which humanity is plunged into (DASGUPTA, 2021).

Projections regarding the global exploitation of fossil fuels point to a dire situation: the environmental impacts of their usage threaten survival on the planet (BECKER et al., 2007). As pointed out by the IPCC (Intergovernmental Panel on Climate Change), humanity's actions might impose an increment of 3º C on the average global temperature by 2100 compared to the pre-industrial era, which would mean the increase in social inequalities, let alone the loss of many lives ("Climate Change 2022: Impacts, Adaptation and Vulnerability", 2022). Moreover, the climate is not the only natural system endangered, as the loss of biodiversity and the change in biogeochemical flows, for example, also point to the end of the Holocene, the only geological era known to nurture human life (STEFFEN et al., 2015). Therefore, the unsustainable ways of achieving development must be readily reconsidered and replaced to defend the existence and equally distributed prosperity of the peoples of Earth.

In this context, the question of how to responsibly produce chemicals for fuels, materials, nutraceuticals and pharmaceuticals arises. These molecules play important roles in modern society, but acquiring them, more often than not, is prejudicial for the environment (SANDERSON, 2011; ZADEK et al., 2021). Thus, the biosynthesis of chemicals using biotechnological processes is shown as an alternative. Known as industrial biotechnology, this field employs metabolic and bioprocess engineering, for example, and has been growing globally with the development of microbiology and its application of biocatalysts in biorefineries (MAURY et al., 2005; FRIEDMAN; BEALL, 2020). For its acknowledged potential as a cheaper and more sustainable option, industrial biotechnology has been explored and improved not only at the scientific level, but also at industrial application scale (NIELSEN, 2001).

2 BIBLIOGRAPHIC REVIEW

2.1 Metabolic and Bioprocess Engineering

Systems metabolic engineering is a growing field in industrial biotechnology which incorporates genetic engineering, metabolic engineering, bioprocess engineering and synthetic biology. It is devised as a cell engineering method in which omic approaches are employed to scan for genetic targets, then synthetic biology is applied to generate strains with desired metabolic modifications and lastly evolutionary engineering is conducted to further improve the host suitability to an industrial setting (LEE et al., 2012). During each stage, consequences of strain design in bioprocess and purification are considered. Therefore, systems metabolic engineering is an iterative process that implies cycles of development in which the host is gradually and rationally modified for cheaper and higher production of biochemicals. In fact, this holistic approach also contributes to the decrease in time and manpower necessary for strain design, one of the main weaknesses of industrial biotechnology approaches (CHOI et al., 2019).

Nowadays, there is a societal and environmental demand for more sustainable production models. Environmentally harmful processes of chemical synthesis may be substituted by methods that employ microorganisms as biocatalysts (SCHMIDT-DANNERT, 2017). Inside the broad systems metabolic engineering field, this substitution is economically feasible and has the potential to reach industrial applications. As examples, diols such as 1,2-propanediol and 1,4-butanediol and plant secondary metabolites like vanillin and resveratrol are biomolecules which turned to microbial biosynthesis for commercialization (CHOI et al., 2019). Another example is the anti-malarial drug precursor and sesquiterpene artemisinic acid (PADDON et al., 2013), although Amyris and Sanofi seem to have abandoned the biosynthesis route due to market uncompetitiveness (KUNG et al., 2018). Nevertheless, related to flavor and fragrance molecules, in the last years some important biotechnology and synthetic biology companies have expanded their businesses to research and development in this field (JIANG; WANG, 2023). In all these examples and in many others, strain modification and process development are entwined in two, but not only, essential research fields: metabolic engineering and bioprocess engineering.

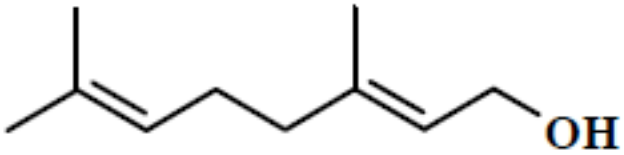
Through the heterologous expression of proteins, metabolic engineering allows the reconstruction and control of exogenous metabolic pathways in microbial hosts. Moreover, this field of knowledge may be employed to increase the metabolic flux from carbon sources to the desired product and to reduce the formation of undesirable co-products. This is possible due to the variety and efficiency of genetic manipulation tools available and in development. (NIELSEN, 2001). For example, genome editing using CRISPR (Clustered and Regularly Interspaced Short Palindromic Repeats) is a now well-developed system in many microbial hosts, and accordingly is transforming the field of metabolic engineering.

In regard to bioprocess engineering, the culture of microorganisms in bioreactors displays advantages compared to traditional chemical synthesis, like less harsh conditions and optically pure products that are deemed as natural (MOSER; PICHLER, 2019; YUAN; ALPER, 2019). As for extraction from natural sources, bioprocesses offer reduced requirement for production area and more feasible scale-up (TAI; STEPHANOPOULOS, 2013). Furthermore, the variety of carbon sources available for microbial growth may be explored to further reduce production costs (MA et al., 2019; SOONG et al., 2019). Among these renewable feedstock sources, using industrial waste and sub-products, which are cheap and abundant, contributes to sustainable development in its environmental, social and economic aspects (MAURY et al., 2005).

Thus, metabolic engineering and bioprocess engineering are entwined areas which present the basis for application of microorganisms as biocatalysts in industrial biotechnology. Research in these fields of knowledge works towards the development and industrial application of chemical production processes that are sustainable.

2.2 Geraniol

Figure 1: Molecular structure of geraniol.



Source: (MĄCZKA; WIŃSKA; GRABARCZYK, 2020).

Geraniol ((2E)-3,7-dimethylocta-2,6-dien-1-ol, molecular mass 154.25 u) (Figure 1) is an acyclic alcohol firstly isolated from the palmarosa oil (*Cymbopogon martinii*) in 1871 by

O. Jacobsen (CLARK IV, 1998). In older sources, the name geraniol may also ambiguously refer to an unspecified mixture of geraniol and its stereoisomer nerol (SELL, 2007). The molecule of geraniol belongs to the monoterpene class of plant secondary metabolites, as it is formed by two isoprene units and thus has 10 carbon atoms (EGGERSDORFER, 2000; RUZICKA, 1953). Geraniol is a volatile substance, insoluble in water, with melting and boiling points of -15º C and 230º C (at 101.3 kPa), respectively, and 889.4 kg m^{-3} of density (at 20º C) (FAHLBUSCH et al., 2003; LAPCZYNSKI et al., 2008; NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION, 2004). Its appearance is that of a transparent or light-yellow oil, with a flowery smell and light waxy taste (MĄCZKA; WIŃSKA; GRABARCZYK, 2020).

Table 1: Geraniol-rich oil sources in plants.

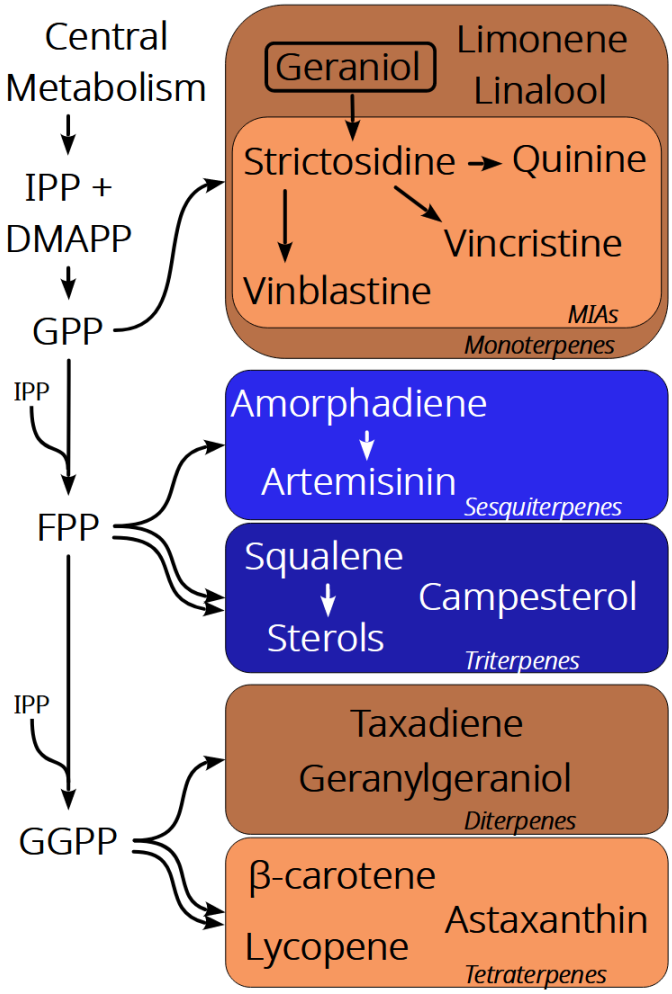
Plant species rich in essential oils	Plant parts from which the essential oils are extracted	Percentage of geraniol in the volatile fraction of the extracted oil (%)
Aeollanthus myrianthus	Flower	66
Aframomum citratum	Seed	70
Boesenbergia pandurata	Rhizome	26
Campomanesia adamantium	Leaf	18.1
Cymbopogon distans	Aerial parts	18.6
Cymbopogon nardus	Whole plant	24.2 / 40.5
Cymbopogon martinni	Whole plant	74.2
Cymbopogon martini var. *martini*	Whole plant	61.4
Cymbopogon martinii var. *motia*	Leaf	93.25
Cymbopogon martinii (Roxb.) Wats. var. *martinii*	Seed	88.06
Cymbopogon winterianus	Whole plant	25.1
Cymbopogon winterianus	Leaf	25.5
Elettariopsis elan	Leaf, rhizome and root	71.6
Neofinetia falcata	Flower	53
Ocimum basilicum	Inflorescences	18.3
Pelargonium graveolens L'Her ex Ait	Fresh herb	34.6
Rosa bourboniana	Flower	15.8
Rosa brunonii Lindl.	Flower	19.2
Rosa damascena	Flower	29.3
Thymus daenensis	Aerial parts	37.2 / 75.7
Thymus longicaulis	Aerial parts	56.8
Thymus longicaulis	Leaf, stem and calyx	27.35
Thymus pulegioides	Whole plant	23.5
Thymus tosevii var. *tosevii*	Whole plant	37.8

Source: adapted from MĄCZKA; WIŃSKA; GRABARCZYK, 2020.

Naturally-occurring geraniol is mostly found in plants, but there is also published evidence of its synthesis by insects, bacteria and fungi (BURSE; BOLAND, 2017; CARRAU et al., 2005; DALOZE; PASTEELS, 1994; REDDY et al., 2020; RUAN et al., 2021). In plant essential oils, it coexists with nerol and its oxidation products, neral and geranial (MĄCZKA; WIŃSKA; GRABARCZYK, 2020). So far, more than 250 geraniol producing plants have been identified. Part of this diversity is shown in Table 1, with the plant part used for extraction of the oil and the proportion of geraniol obtained from it (MĄCZKA; WIŃSKA; GRABARCZYK, 2020). Components of essential oils are considered plant secondary metabolites, as they are not required for the organism survival (YANG et al., 2018). Their functions are related to defense, stress tolerance, communication and competition (LEE DÍAZ et al., 2022; YANG et al., 2018). Indeed, geraniol presents antimicrobial and phytotoxic activities (LEI et al., 2019; LOŽIENĖ; VAIČIULYTĖ, 2022). Thus, it can be speculated that its natural synthesis by plants may be linked to protection from pathogens and to territorial dominance, for example.

Geraniol biosynthesis requires the terpene precursors isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) (MAHMOUD; CROTEAU, 2002). These two molecules are condensed to form geranyl diphosphate (GPP), common precursor to all monoterpenes and monoterpenoids, classes that comprise from 1000 to 55000 compounds (DEHSHEIKH et al., 2020; ZEBEC et al., 2016). GPP is then converted to geraniol in a single enzymatic step (DONG et al., 2013; IIJIMA et al., 2004; SIMKIN et al., 2013). By the sequential addition of IPP molecules to GPP, the immediate precursors of sesquiterpenes, diterpenes, triterpenes and tetraterpenes are also generated (MOSER; PICHLER, 2019). An overview of the biodiversity of terpenes is shown in Figure 2.

Figure 2: Common classes of terpenes with some noteworthy examples and biosynthesis pathway overview.



DMAPP: Dimethylallyl diphosphate; FPP: Farnesyl diphosphate; GGPP: Geranylgeranyl diphosphate; GPP: Geranyl diphosphate; IPP: Isopentenyl diphosphate; MIAs: Monoterpene Indole Alkaloids. Source: Author. Based in (ARNESEN et al., 2020; MOSER; PICHLER, 2019).

2.2.1 Market, applications and consequences

In 1993, essential oil total market size was estimated as 0.6 billion dollars (FAHLBUSCH et al., 2003). This figure skyrocketed to 7.5 billion by 2018. Additionally, this global market was expected to grow an additional 9 % between 2019 and 2026 (SHARMEEN et al., 2021). All this information indicates the expanding trade of essential oils in the last decades. Many of these largely commercialized oils display geraniol in their composition. Essential oils from geranium plant (*Pelargonium* sp.), with approximately 75 % (v / v) geraniol, has been employed since 1819 due to its characteristic scent (CLARK IV, 1998). As

another example, the oil from the citronella grass (*Cymbopogon winterianus*), is used in insect repellents and is applied as a food additive, and shows around 13 % of geraniol in its composition (WANY et al., 2014).

As for geraniol, by the beginning of the 21st century its global market demand was estimated at 1000 annual tons (LAPCZYNSKI et al., 2008). Another number from the same time period points out that 6000 metric tons of geraniol, nerol and its esters are used every year (SELL, 2007). Geraniol cost nowadays ranges from US$ 15 to 100 per litre (ALIBABA.COM, 2022; MADE-IN-CHINA.COM, 2022). Although fragrance prices lie close to those of petrochemical bulk chemicals, for some products the fragrance makes most of the final cost. Therefore, fragrance costs play an important role in cleaning products market balance and competitiveness (SELL, 2007).

Geraniol is mostly applied as a fragrance and was a component of over 76 % of deodorants and 41 % of cleaning products in the European market by the beginning of the century (RASTOGI et al., 1998, 2001). This monoterpene is also found in shampoos, facial creams and other cosmetics (LAPCZYNSKI et al., 2008). Furthermore, it has industrial importance for the synthesis of other terpenes, such as citronellol and citral, and in small quantities it may be employed in foodstuff to enhance citric flavors (FAHLBUSCH et al., 2003; SELL, 2007).

The geraniol molecule in plants may also be directed to the synthesis of strictosidine, through the 11 steps of the seco-iridoid pathway (MIETTINEN, 2014). Therefore, this molecule is relevant as it is the precursor of thousands of monoterpenoid indole alkaloids (MIAs) (O'CONNOR; MARESH, 2006). Molecules from this diverse group are extracted in small quantities from plants (MIETTINEN, 2014), in which they are synthesized by complex and not yet completely explored pathways (DE LUCA et al., 2012). Moreover, these secondary metabolites are applied in a vast array of medical fields, as drugs for neurological disorders, cancer and cardiovascular conditions, among others. As examples, vinblastine, vincristine, irinotecan and their derivatives are approved antitumor drugs with markets that sum up a billion dollar global figure (ZHANG et al., 2022). Another important molecule is quinine, employed as the main drug used in the treatment of malaria (BROWN et al., 2015; MIETTINEN, 2014; O'CONNOR; MARESH, 2006; WILSON; ROBERTS, 2012).

Geraniol may cause contact skin allergic dermatitis in humans (HAGVALL et al., 2018; SILVESTRE et al., 2019), despite being approved for use in low concentrations as

flavoring and fragrance by different international organizations, including the USA FDA (LAPCZYNSKI et al., 2008). Another issue is that geraniol may be found in the environment as a consequence of human activities, given its wide employment as a fragrance. In a study analyzing the waters of Portuguese rivers (HOMEM et al., 2022), geraniol overall concentration in samples was one of the highest among personal care product-related compounds. In three points of two rivers, geraniol concentration indicated potential ecotoxicological risk. Furthermore, it has been proven in laboratory that extracts containing mostly geraniol may impair seed germination of economically relevant plants (LOŽIENĖ; VAIČIULYTĖ, 2022). This poses a problem if uncontrolled disposal of large quantities of the human-made monoterpene is carried out. Therefore, in regard to scaled-up production of this chemical, studies on the geraniol life cycle should be conducted, aiming at its safe treatment or disposal.

2.2.2 Novel applications at the research level

Lately, geraniol biotransformation has been an increasingly developed scientific field. It consists of employing cells to metabolize the monoterpene and convert it to other fragrance molecules with commercial value (CHEN; VILJOEN, 2010). Plant cell culture may be employed, such as *Anethum graveolens* roots, which presented conversion of geraniol to 10 products, including citral, citronellol and linalool (FARIA et al., 2009). There are also reports of bioconversion of geraniol, especially into linalool, by *Kluyveromyces* sp., *Pichia* sp., *Debaryomyces* sp. and *Aspergillus niger* (DEMYTTENAERE; DEL CARMEN HERRERA; DE KIMPE, 2000; PONZONI et al., 2008). In a recent example, researchers employed *Saccharomyces cerevisiae* to achieve the conversion of geraniol to citronellol and to the geranic and citronellic acids (OHASHI; HUANG; MAEDA, 2021). Microbial production of MIAs using geraniol as a starting point has also been researched in *S. cerevisiae* (MISA et al., 2022).

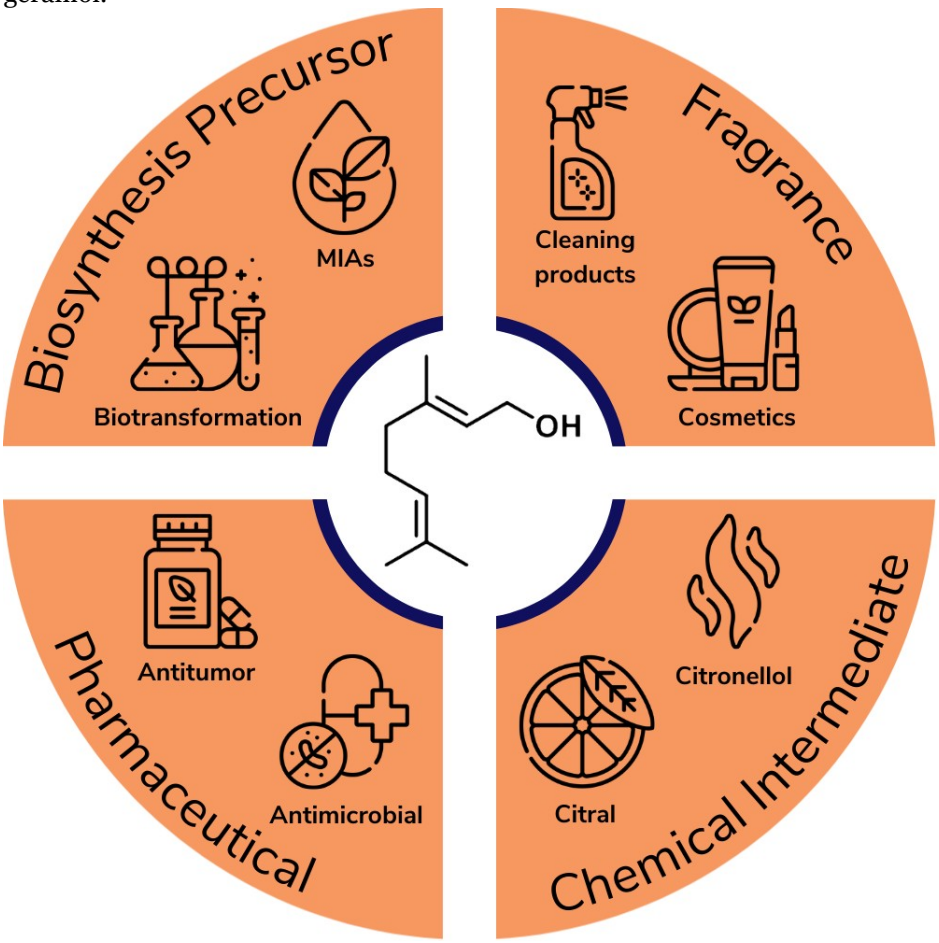
Despite this interest in bioconversion by part of the scientific community, most recent published papers about geraniol focus mainly in the evaluation of this monoterpene as a pharmaceutical (LEI et al., 2019; MĄCZKA; WIŃSKA; GRABARCZYK, 2020). In fact, the keyword search for “pharmacological geraniol” in PubMed gives about 578 results in the period of 1962 to 2021. Of those, 45 publications are from 2021, while in 2001 there were only nine. This may indicate that it is an area of study currently in expansion.

Geraniol has been tested against a plethora of tumoral cells, both *in vitro* and *in vivo*. It presented relevant effects in human lung, skin, pancreatic, liver, mammary, prostate and colon tumor cells (CHEN; VILJOEN, 2010; FATIMA et al., 2021; LEI et al., 2019; MĄCZKA; WIŃSKA; GRABARCZYK, 2020). Taken together, these research works show that the monoterpene has the potential to be used in anticancer therapies. Geraniol antiproliferative action against tumoral cells may be related to its antioxidant and anti-inflammatory activities, which have been extensively studied (LEI et al., 2019). In this regard, it is shown that geraniol aids in the control of inflammation in ulcerative colitis in mice (MEDICHERLA et al., 2015). In another example, the protective activity of geraniol against age-related neurodegenerative diseases is demonstrated in rats (ATEF et al., 2022). Indeed, molecular dynamic simulations that show possible interaction of geraniol with Alzheimer's disease-related ligands have been conducted (LIU et al., 2022d).

Lastly, one of the most studied properties of geraniol is its antibiotic activity. Among bacterial species with recorded impaired growth under geraniol, is *Escherichia coli*, *Salmonella enterica*, *Streptococcus pneumoniae*, *Streptococcus pyogenes* and *Staphylococcus aureus*, for example. Meanwhile, geraniol also shows relevant results against *Candida albicans* and *Trichophyton rubrum*, which are infectious fungi (CHEN; VILJOEN, 2010; LEI et al., 2019; MĄCZKA; WIŃSKA; GRABARCZYK, 2020). Possible applications range from food packages with antimicrobial action to antiperspirants and human medicine. In mice, *Helicobacter pylori* infections can be controlled through geraniol administration, as described in (BHATTAMISRA et al., 2019), which paves the way for the development of an alternative therapy for humans.

Thus, while the geraniol molecule itself, at present, is applied mostly as a fragrance and as chemical intermediate, it may come to be utilized as an antimicrobial and chemotherapeutic drug, among others. Moreover, its formation is an important limiting step in the biosynthesis of MIAs (ZHANG et al., 2022), molecules which chemical synthesis processes are hardly feasible (MISA et al., 2022). These novel applications are at scientific investigation level, and additional production process development and clinical trials are yet to be conducted. Notwithstanding, establishment of any of these potential applications in the next years will undoubtedly further increase geraniol demand. Therefore, it is imperative that a cost-effective and environment-responsible production process is developed and employed in the industry setting. Some of the present and potential commercial uses of geraniol are represented in Figure 3.

Figure 3: Examples of present and potential commercial applications of geraniol.



Source: Author. Geraniol molecule structure from (LIU et al., 2021c). Icons designed by Freepik from Flaticon.

2.2.3 Geraniol production

Geraniol may be acquired through plants that synthesize it. Among the methods for essential oil extraction, it is possible to submit citronella or palmarosa plant tissue to supercritical CO_2, vapor distillation and hidrodistillation (WANY et al., 2014). After extraction, fractional distillation is employed for separation and isolation of the target molecule. As an example, extracts from *Java citronella* oil are composed of up to 60 % geraniol (FAHLBUSCH et al., 2003).

Extraction may be a comparatively simple technique, but it is a low yield and unsustainable process due to some fundamental issues (LIANG et al., 2021). Firstly, great quantities of slow-growing plant biomass are required for plant extraction of terpenes. The yield of essential oils, which may be low in the target terpene, represents a few percent of the plant tissue at best (FAHLBUSCH et al., 2003). As an illustration, obtaining 1 g of the MIA chemotherapeutics vinblastine and vincristine requires up to 500 and 2000 kg, respectively, of

Catharanthus roseus leaves (ZHANG et al., 2022). Similarly, geraniol low-volume production imposes a restriction to further market development (LEI et al., 2019). Indeed, 500 kg of rose petals yield only 1 kg of geraniol-rich oil (CLARK IV, 1998). Secondly, the extracts obtained are complex mixtures with compounds similar in structure, which interfere with isolation and purification of the target molecule, increasing the time, labor and cost requirements of the overall process (BROWN et al., 2015; KRIVORUCHKO; NIELSEN, 2015).

Alternatively, it is possible to cultivate plant cells in bioreactors and then extract the desired plant metabolite. Indeed, this has been commercially applied for terpene production, including geraniol, by Mitsui Chemicals (KOLEWE; GAURAV; ROBERTS, 2008). However, geraniol is not listed in this company's product portfolio anymore (MITSUI CHEMICALS, INC, 2022). This might be due to culture heterogeneity, low growth rates, stress susceptibility and aggregation, which translates in low and variable product titer, complications that plant cell culture faces (KOLEWE; GAURAV; ROBERTS, 2008; WILSON; ROBERTS, 2012). Moreover, metabolic engineering of such cells to increase yields, although feasible, encounters technical issues related to difficult genetic manipulation of plants (KRIVORUCHKO; NIELSEN, 2015; MUHAMMAD et al., 2020).

Nowadays, chemical synthesis presents greater commercial importance than plant extraction for industrial sourcing of geraniol (EGGERSDORFER, 2000; FAHLBUSCH et al., 2003; SELL, 2007). One such process, called semi-synthesis, achieves a mixture with 98 % geraniol from β-pinene via myrcene. Feedstock for this chemical process is costly, as it is obtained from the wood and paper industry through energy inefficient operations. Nevertheless, using the less expensive α-pinene generates plinols in a side reaction. Therefore, the incomplete conversion of starting materials and a technically difficult subsequent distillation must be conducted. Finally, chemical interconversion between geraniol, citronellol, citronellal, linalool and citral (which is an unspecified mixture of geranial and neral) is also possible. This proffers process versatility, as citral synthesis may be achieved through total synthesis, an easily conducted, although unsustainable route with petrochemical precursors. Nevertheless, the interconversion is more often operated in the inverse, with geraniol yielding citral, given citral higher prices as synthetic precursor for vitamins E and A (FAHLBUSCH et al., 2003; SELL, 2007).

Therefore, chemical synthesis is a more suitable process for geraniol production than plant extraction. Comparatively, chemical routes are more reproducible processes, as they are independent from climate, soil and crop growth stage conditions, for example (BAKKALI et

al., 2008). Notwithstanding, they require intermediates that are either expensive or unsustainable. Moreover, temperatures of up to 500º C are employed and remnant chlorinated compounds must be removed from the obtained mixtures (SELL, 2007).

2.3 Heterologous biosynthesis of geraniol

The hurdles that terpene extraction and chemical synthesis suffer from indicate the necessity of designing microbial biocatalysts for geraniol acquisition (CHANG; KEASLING, 2006). Thus, biotechnology may be utilized for the development of a process of products regarded as natural with less harsh requirements and higher potential of achieving process sustainability.

Monoterpene biosynthesis bioprocesses still lack viability analyses that tackle the economic, social and environmental aspects of sustainability. It has not been found such analyses for the microbial synthesis of geraniol. As examples, some works have described techno-economic analyses focusing on the *E. coli* production of the monoterpene limonene from glucose. In a first article, limonene is the closest to achieve viable production and commercialization among twelve terpenes, including monoterpenes linalool and pinene. Nonetheless, it would still be imperative that limonene yields increase to at least 30 % of the maximum theoretical (WU; MARAVELIAS, 2018). In another report, the authors describe that, unless this number increases to 45 % of the maximum theoretical yield, industrial production of microbial limonene would not be feasible. As means of comparison, it is stated that the highest yield achieved at the time was approximately 1 % of the maximum theoretical (SUN; THEODOROPOULOS; SCRUTTON, 2020). Thus, both reports agree that one of the disadvantages of biotechnological processes when compared to chemical and extractive ones is low yield, issue that is being addressed with the great number of inventive systems metabolic engineering research (LEE et al., 2012). Accordingly, geraniol biosynthesis research, as of that of limonene, focus on fine tuning the heterologous host to increase product yield.

Currently, the highest titer of geraniol achieved in any heterologous host is 5.5 g L^{-1}, with recombinant *S. cerevisiae* with peroxisome-bound precursor pathway employed in fed-batch operation (DUSSÉAUX et al., 2020). Although not reported in the article, the geraniol yield of this strain in batch culture was estimated to be approximately 0.9 % C-mol C-mol^{-1}. Other noteworthy works with the baker's yeast have achieved more than 1.6 g L^{-1} in fed-batch operation (JIANG et al., 2017; ZHAO et al., 2017). Considering the highest titers in batch

operation for yield estimation, the most efficient strain was the one from (JIANG et al., 2017), which achieved 5.1 % C-mol geraniol C-mol^{-1} glucose, while in (ZHAO et al., 2017) the highest yield was approximately 2.8 % C-mol C-mol^{-1}. As for genetically modified *E. coli*, it was achieved 2.12 g L^{-1} geraniol (with an estimated yield of 21.1 % C-mol C-mol^{-1}, the highest yet) (WANG et al., 2021). Considering the production of the geraniol ester geranyl acetate, the titer value goes up to 4.8 g L^{-1} (with a yield estimated at 6.6 % C-mol geranyl-acetate C-mol^{-1} substrate) (CHACÓN et al., 2019). All these concentrations are similar to the ones obtained for the production of other monoterpenes on these hosts (LIU et al., 2021b). However, these are still considered low numbers for industrial settings, especially when compared to the 20~100 g L^{-1} achieved with sesquiterpene production (CHACÓN et al., 2019; MOSER; PICHLER, 2019; PADDON et al., 2013).

Other than the attempts of heterologous geraniol production in *E. coli* and *S. cerevisiae*, the bacteria *Azospirillum brasilense* (0.184 mg L^{-1} geraniol) (MISHRA et al., 2020) and *Corynebacterium glutamicum* (15.2 mg L^{-1} geraniol) (LI; XU; LU, 2021), the archaeon *Methanococcus maripaludis* (0.5 mg L^{-1} geraniol) (LYU et al., 2016) and the alga *Phaeodactylum tricornutum* (0.309 mg L^{-1} geraniol) (FABRIS et al., 2020) have been employed as microbial chassis. The fungi *Aspergillus oryzae* was also used as host to produce geraniol as a means to achieve the synthesis of the MIA precursor nepetalactol (7.2 mg L^{-1} nepetalactol) (DUAN et al., 2021). Among these alternative hosts, the highest titer to date (1.194 g L^{-1}, estimated to represent 8.9 % C-mol C-mol^{-1}) was achieved by *Candida glycerinogenes*, in which a dual pathway engineering approach was allied with the control of geraniol synthase expression by an n-decane inducible promoter (ZHAO et al., 2022). Nevertheless, scientific publications describing heterologous geraniol production employ mostly *E. coli* and *S. cerevisiae* (CHACÓN et al., 2019; GERKE et al., 2020; JIANG et al., 2017; LIU et al., 2016; TASHIRO et al., 2017; WANG et al., 2021; YEE et al., 2019; ZHAO et al., 2017; ZHOU et al., 2014). This is due to the accumulated knowledge and the plethora of molecular biology and genome editing tools accessible for these model organisms. In these hosts, geraniol biosynthesis is achieved by the heterologous expression of only the geraniol synthase (GES) enzyme, but it is hindered by the low availability of precursors and by the product toxicity (ZHAO et al., 2017). However, these hurdles may be ameliorated in other microorganism expression platforms which present advantageous attributes for the production of monoterpenes.

2.3.1 Metabolic pathways of geraniol biosynthesis

The IPP and DMAPP isoprene units, which are the precursor molecules of all terpenes, may be generated by two different naturally occurring pathways. On the one hand, the 2-C-methyl-D-erythritol-4-phosphate (MEP) or 1-deoxy-D-xylulose-5-phosphate (DXP) pathway is found in plant plastids and bacteria. Monoterpenes are synthesized through this metabolic pathway in plants (LUAN, 2002; MAHMOUD; CROTEAU, 2002). On the other hand, the mevalonate (MVA) pathway is natural to plant citosol and other eukariotic cells. This pathway starts from the main catabolism product acetyl-CoA and is required for synthesis of different essential compounds, including sterols such as cholesterol in animals and ergosterol in yeasts, which are necessary for cell membranes (MAURY et al., 2005; MOSER; PICHLER, 2019). Synthetic, non-natural pathways have also been established, like the Isopentenol Utilization Pathway (IUP), Alcohol-Dependent Hemiterpene (ADH) pathway and the Isoprenoid Alcohol (IPA) pathway (CLOMBURG et al., 2019; LI et al., 2021b; ZHU et al., 2021). These pathways are allegedly simpler to manipulate, carbon efficient or require less cofactors than the MEP or MVA pathways. All three of them start from isopentenol and its isomers prenol or isoprenol, but only the IPA pathway was found to have been linked with central metabolism (CLOMBURG et al., 2019). A simplified overview of these metabolic pathways is shown in Table 2.

Table 2: Pathways comparison for 1 mol of IPP formation considering glucose as carbon source.

Pathway	Nº of steps	Cofactors	Co-products	Max. theor. Yield (% C-mol C-mol^{-1})	References
MVA	17	2 (NADPH + H^+)	2 $NADP^+$ + 4 CO_2 + 6 (NADH + H^+)	56	(MIZIORKO, 2011)
MEP	17	(NADPH + H^+) + CTP + ATP	$NADP^+$ + CMP + ADP + CO_2 + (NADH + H^+)	83	(BANERJEE; SHARKEY, 2014)
IPA	19	2 (NADPH + H^+) + ADP	2 $NADP^+$ + 4 CO_2 + 6 (NADH + H^+) + ATP	56	(CLOMBURG et al., 2019)

The plant MEP pathway is fed with pyruvate and glyceraldehyde 3-phosphate from central metabolism. As the pathway is localized in plastids, these compounds come mostly from the Calvin-Benson cycle. They undergo seven enzymatic reactions which lead to the IPP precursor (BANERJEE; SHARKEY, 2014). The MEP pathway IPP maximum theoretical yield in carbon is higher compared to the MVA pathway, but they both present similar results

when considering microbial growth and metabolic pathway networks, with MEP maximum IPP yield at and MVA at (GRUCHATTKA et al., 2013).

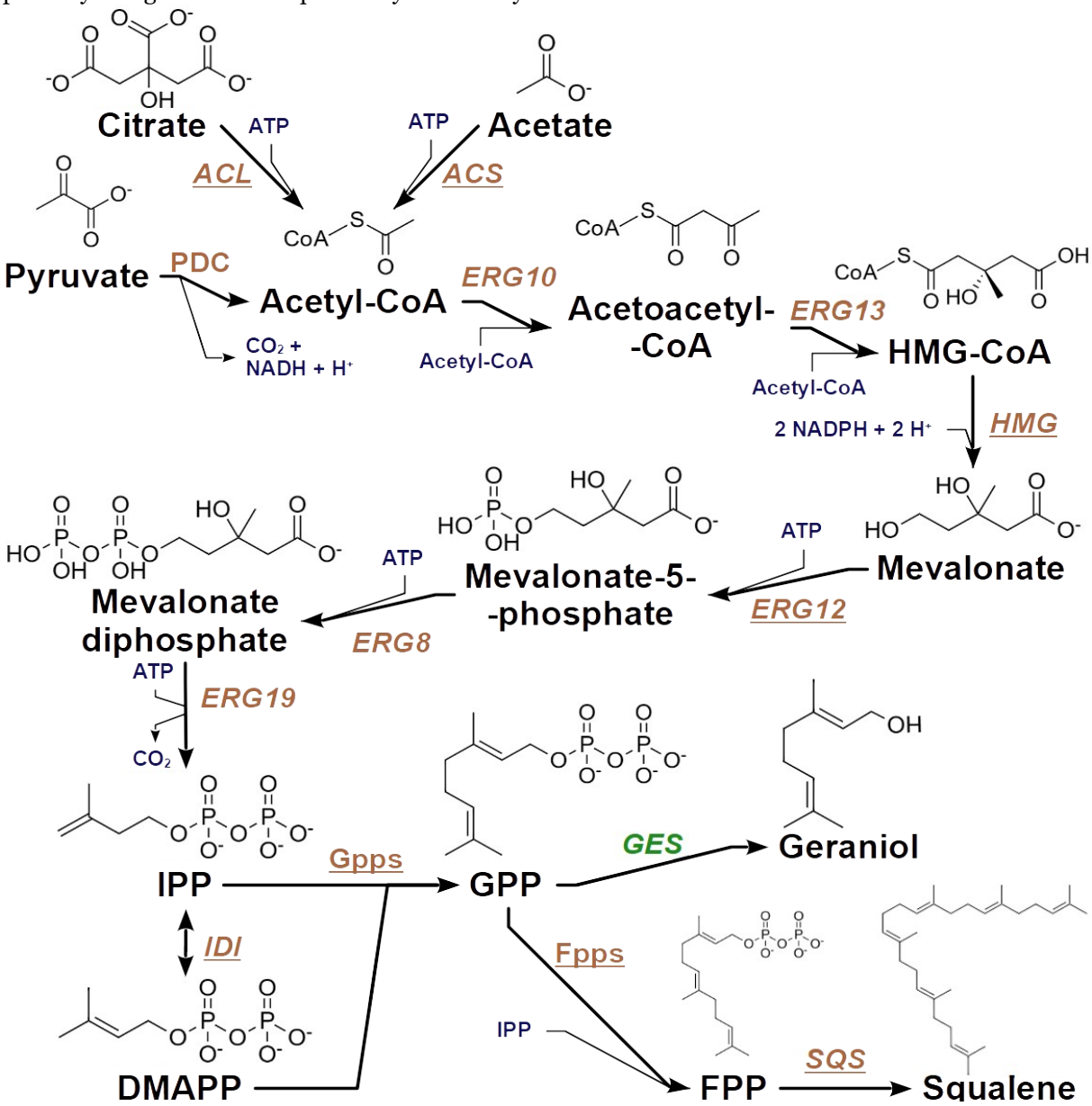
For the biosynthesis of geraniol in yeasts, the endogenous cytosolic MVA pathway, which generates the membrane component ergosterol from squalene, might be exploited (Figure 4). In its first reaction, two molecules of acetyl-CoA are condensed by acetyl-CoA acetyltransferase (*ERG10* gene) to generate acetoacetyl-CoA. This molecule undergoes a fusion with a third acetyl-CoA molecule by 3-hydroxy-3-methylglutaryl-CoA synthase (*ERG13*), generating 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA). Through reduction of HMG-CoA by 3-hydroxy-3-methylglutaryl-CoA reductase (*HMG*), mevalonate is obtained. Mevalonate is phosphorilated into mevalonate-3-phosphate by mevalonate kinase (*ERG12*) and then phosphorilated again by phosphomevalonate kinase (*ERG8*), producing mevalonate diphosphate. On its turn, mevalonate diphosphate is the substrate for mevalonate diphosphate decarboxylase (*ERG19*), releasing CO_2 and an IPP. Finally, isopentyl diphosphate isomerase (*IDI*) may turn IPP into DMAPP (MIZIORKO, 2011).

However, the condensation of IPP and DMAPP into GPP, the monoterpene precursor, encounters an obstacle in yeasts. The protein codified by the *ERG20* gene exerts consecutively both the geranyl diphosphate synthase (Gpps) and the farnesyl diphosphate synthase (Fpps) catalytic activities. Thus, most of the GPP formed is readily converted to farnesyl diphosphate (FPP) and doesn't leave the enzyme active site (FISCHER et al., 2011). Therefore, expression of a mutated *ERG20* protein, with trytophan residues in the place of a phenylalanine residue at the position 96 and of an asparagine residue at the position 127, achieves higher release of GPP in *S. cerevisiae* (IGNEA et al., 2014). Such modification was achieved in the non-conventional yeast *Yarrowia lipolytica* altering the phenylalanine residue at the position 88 and the asparagine residue at the 119 position to tryptophan residues. By expressing this mutated *ERG* $20^{F88W-N119W}$ protein in addition to *HMG* and *IDI* overexpression, the heterologous production of the monoterpene linalool reached the highest titer in yeasts yet (CAO et al., 2017).

Lastly, for geraniol biosynthesis, heterologous expression of *GES* enzyme (EC 3.1.7.11) is required. The sequences of orthologous *GES* enzymes are relatively conserved. Nevertheless, they display distinct activities in *S. cerevisiae* and *E. coli* (JIANG et al., 2017). Moreover, the natural protein exhibits a plastid-bound N-terminal transit peptide after its translation (IIJIMA et al., 2004; SIMKIN et al., 2013). As heterologous hosts do not remove this sequence, the geraniol synthase activity may be hindered. Indeed, expression of the *GES*

gene without the sequence that codifies the transit peptide leads to higher geraniol production in yeasts (JIANG et al., 2017; ZHAO et al., 2017).

Figure 4: Representation of the metabolic pathways for acetyl-CoA synthesis, mevalonate (MVA) pathway and geraniol and squalene synthesis in yeasts.



ACL: ATP citrate lyase; *ACS*: Acetyl-CoA synthetase; CoA: Coenzyme A; DMAPP: Dimethylallyl diphosphate; *ERG8*: Phosphomevalonate kinase; *ERG10*: Acetyl-CoA acetyltransferase; *ERG12*: Mevalonate synthase; *ERG13*: 3-hydroxy-3-methylglutaryl-CoA synthase; *ERG19*: Mevalonate diphosphate decarboxylase; FPP: Farnesyl diphosphate; Fpps: Farnesyl diphosphate synthase; *GES*: Geraniol synthase; GPP: Geranyl diphosphate; Gpps: Geranyl diphosphate synthase; *HMG*: 3-hydroxy-3-methylglutaryl-CoA reductase; HMG-CoA: 3-hydroxy-3-methylglutaryl-CoA; *IDI*: Isopentyl diphosphate isomerase; IPP: Isopentenyl diphosphate; *PDC*: Pyruvate dehydrogenase complex; *SQS*: Squalene synthase; Genes with underlined names were genetic engineering targets in the *Y. lipolytica* ST9202 strain. Gene in green has to be heterologously expressed for geraniol synthesis. Source: Author. Based in (ARNESEN et al., 2020; MIZIORKO, 2011).

2.3.2 Metabolic engineering for geraniol production

The highest production of geraniol obtained in recombinant *E. coli*, the second highest in any heterologous host so far, has achieved less than 30 % of the maximum theoretical yield (WANG et al., 2021). Therefore, there is a vast room for improvement which may be achieved by fine-tuning the microbial catalysts metabolism. Scientific literature describes different genetic modification of microbial hosts aiming at increased monoterpene biosynthesis.

Some common targets of metabolic engineering include pathway genes superexpression and increase in precursor molecule pools. In regard to the MVA pathway, superexpression of *HMG* is often conducted, as both the Hmg1p and Hmg2p enzymes coded by this gene family are highly regulated by pathway products (LIU et al., 2021b; ZHU et al., 2021). For example, squalene and ergosterol are shown to downregulate Hmg1p and Hmg2p activity in *S. cerevisiae* (KLUG; DAUM, 2014). Another successful strategy is the overexpression of *IDI*, which in one case increased production by 1.45 fold in *S. cerevisiae* (LIU et al., 2013). Other MVA pathway genes, such as *ERG10* and *ERG12*, have also been targeted and showed promising results at increasing monoterpene yields in different hosts (ZHU et al., 2021). For whole pathway upregulation, impaired transcription of the *ALD6* gene (codes for a aldehyde dehydrogenase related to acetate formation from acetaldehyde in the cytosol) increases carbon flux through the MVA pathway and thus is a target for elevated geraniol synthesis by *S. cerevisiae* (CHEN et al., 2019). *In silico* analysis pointed out that ATP and reducing agents are limiting resources to both MVA and MEP, so terpene biosynhtesis would also benefit from metabolic engineering strategies that target these molecules (GRUCHATTKA et al., 2013). As for higher acetyl-CoA availability, overexpression of *ACL* (encodes ATP citrate-lyase) and downregulation of fatty-acid synthesis genes are some of the strategies employed (ZHU et al., 2021).

Other than these modifications, it is also explored the identification and deletion of genes that code for geraniol-metabolizing enzymes or enzymes that drain the precursors downstream the MVA pathway. A strategy of pairing the degradation of the Fpps enzyme (coded by *ERG20*) with sterol production was devised to diminish GPP sinking into FPP without compromising essential ergosterol synthesis (PENG et al., 2018). This work was successful in increasing the metabolic flux to monoterpenes, albeit with a toll in *S. cerevisiae* growth. Also in *S. cerevisiae*, the knockout of the *OYE2* and *ATF1* genes diminishes the transformation of geraniol into citronellol (BROWN et al., 2015; ZHANG et al., 2022; ZHAO et al., 2017). In an example study, *OYE2* deletion resulted in the increase of geraniol

production 1.7 times, without altering yeast growth. Knockout of *ATF1* showed similar results, but *ATF1* and *OYE2* double knockout culminated in growth impairments (ZHAO et al., 2017).

Compartmentalization of the pathway in the mitochondria may further increase geraniol synthesis. This strategy has been tested with promising results, as six-fold geraniol synthesis was achieved compared to cytosolic localization (YEE et al., 2019). It is based on the hypothesis that the mitochondria-constrained pathway would not only have more of the acetyl-CoA precursor at disposal but would also show decreased conversion of GPP to FPP. In a similar fashion, directing the MVA pathway to peroxisomes was also a successful strategy applied (DUSSÉAUX et al., 2020; GERKE et al., 2020), which increased geraniol production by 13 % in yeast in one example (GERKE et al., 2020). It is stated that, compared to the mitochondria, peroxisome engineering may have less impact in cell fitness and is more favorable for substrate transport, as this organelle possesses a single membrane (DUSSÉAUX et al., 2020).

Aside from modifications in the MVA pathway, insertion of the whole MEP pathway in *S. cerevisiae* has also been described in an attempt to increase product titers (KIRBY et al., 2016). In this work, however, yeast growth relying solely on the MEP pathway was greatly impaired. Nevertheless, this is a promising field of research as *in silico* predictions indicate that MEP pathway expression would benefit IPP production (GRUCHATTKA et al., 2013). This is partially due to the higher maximum theoretical yield of MEP. It is also stated that the introduced pathway would be benefited by the lack of intrinsic regulation (GRUCHATTKA et al., 2013; KIRBY et al., 2016). Likewise, introduction of MVA in *E. coli* has been achieved in some works with promising results (CHACÓN et al., 2019; LIU et al., 2016; ZHOU et al., 2014). Related to whole metabolic pathway expression, there are reports of expression of an alternative, non-oxidative, four-step glycolytic pathway in *S. cerevisiae* and *Y. lipolytica* (BERGMAN et al., 2016; LU et al., 2020). Called Phosphoketolase-Phosphotransacetylase (PK-PTA), this pathway increases the carbon efficiency of acetyl-CoA formation, thus benefiting terpene biosynthesis through the MVA pathway.

2.3.3 Bioprocess engineering for geraniol production

In regard to increasing geraniol biosynthesis, most of the focus of the scientific literature is directed to metabolic engineering and synthetic biology. Nevertheless, promising results may also be achieved through optimization of culture conditions.

Aside from low-volume shake-flasks, the fed-batch operation mode is the most commonly observed among the heterologous geraniol production literature (CHACÓN et al., 2019; DUSSÉAUX et al., 2020; JIANG et al., 2017; LIU et al., 2016; YEE et al., 2019; ZHAO et al., 2017). Due to geraniol toxicity, fed-batch operation is often allied with inducible promoters to restrain *GES* expression and geraniol production to the stationary phase of microbial growth, after biomass has been formed (JIANG et al., 2017; YEE et al., 2019).

Use of alternative and secondary carbon sources is another parameter for geraniol biosynthesis optimization. For example, it was shown that geraniol production in *S. cerevisiae* benefited from ethanol feeding (ZHAO et al., 2017). Cultures fed with 400 g L^{-1} ethanol at a rate of 0.1 h^{-1}, when compared to ones fed with mixtures of glucose and ethanol, resulted in the highest geraniol production without compromising cell growth. It is discussed that, as ethanol provides a direct source of acetyl-CoA, there is a higher carbon flux through the MVA pathway. Similarly, supplementation of medium with 1.2 g L^{-1} acetic acid increased geraniol production by 60 % in heterologous MVA-bearing *E. coli* (CHACÓN et al., 2019). These works indicate that the MVA pathway suffers from low precursor availability in geraniol mainstream heterologous hosts.

Furthermore, culture conditions are also targets for geraniol production process improvement. For example, it has been reported that geraniol may suffer from pH-dependent conversion to nerol (FISCHER et al., 2011). In this study, while only 5.9 % of geraniol remained in the pH 7 medium at the end of the experiments, 27.5 % remained in the pH 3.4 medium. Therefore, product stability must be considered during geraniol biosynthesis process engineering.

2.3.4 Geraniol cytotoxicity

The antimicrobial activity of geraniol is pointed out as another hurdle to its heterologous biosynthesis, aside from precursor availability. As geraniol concentration increases in culture, both yeasts and bacteria show hindered growth (CHACÓN et al., 2019; GERKE et al., 2020; ZHAO et al., 2016). The main mechanism of geraniol-related cell death in *S. cerevisiae* is damage to the cell membrane due to the hydrophobic nature of geraniol (SCARIOT et al., 2021). The geraniol concentration limit that completely inhibited *E. coli* cell growth was determined as 300 mg L^{-1} (CHACÓN et al., 2019), with slightly impaired growth observed with as low as 75 mg L^{-1} of geraniol in the medium. For *S. cerevisiae*,

antifungal activity of geraniol was observed only at 150 mg L^{-1}, but 200 mg L^{-1} was sufficient to completely impair cell growth (ZHAO et al., 2016).

This issue is commonly circumvented by employment of appropriate organic solvents in the culture media. This way, the monoterpene is captured by this organic phase, which is then used for isolation and quantification of the molecule by gas-chromatography coupled with mass spectrometry. Another point addressed is geraniol volatility, which has a lessened effect on geraniol final titers by the presence of the organic phase. The most frequently described compounds in scientific literature are dodecane and isopropyl myristate (CAO et al., 2017; IGNEA et al., 2014; JIANG et al., 2017; WANG et al., 2021). Other compounds, like hexane and decane, have also been tested (LYU et al., 2016), but these present less capturing efficiency. In a systematic approach, a study found that isopropyl myristate was the most suitable solvent for geraniol sequestering among 18 liquid and solid compounds, although dodecane was not tested (PRIEBE et al., 2018).

Moreover, there are records of metabolic modifications aiming at increased geraniol tolerance in heterologous hosts. One such strategy is the truncation of the Bul1 α-Arrestin-Like Adaptor in *S. cerevisiae*. In (GERKE et al., 2020), it was observed that the loss of function by this protein after deletion of genes related to peroxisome metabolism increased the engineered strains tolerance to geraniol. In another report, conversion of geraniol to geranyl acetate, which is a less toxic molecule, was conducted in *E. coli* (CHACÓN et al., 2019) through alcohol acyltransferase expression. This strategy led to the biosynthesis of 4.8 g L^{-1} geranyl acetate. It is stated that, after biosynthesis, geraniol may be recovered from geranyl acetate by ester hydrolysis.

Finally, strategies applied for alleviating cytotoxicity of other monoterpenes may be applied to geraniol biosynthesis processes. The truncated version of Tcb3p, the tricalbin protein in *S. cerevisiae*, is related to higher cell wall integrity and resilience in the presence of limonene, β-pinene and myrcene (LIANG et al., 2021). Another example showed that, in *Y. lipolytica*, YALI0F19492p, a protein of unknown function, had its transcription upregulated during growth in toxic concentrations of limonene (LI et al., 2021a). When overexpressed under the control of the hp4d promoter, this protein was correlated with higher limonene titers and higher growth rates in the presence of linalool. Testing these approaches for geraniol biosynthesis might be fruitful for increasing product titers.

2.4 *Yarrowia lipolytica* as a host in industrial biotechnology

Y. lipolytica is an aerobic and oleaginous yeast capable of growing in different industrial-relevant carbon sources, such as hexoses, glycerol, acetate, fatty acids and cellobiose (MA et al., 2019). It is commonly found in soil, cheese and meat samples and thus is a safe microorganism to human health, with GRAS (Generally Recognized As Safe) strains. As for its taxonomic position, *Y. lipolytica* had been named *Candida lipolytica* and then *Saccharomycopsis lipolytica* and is a constituent of the Saccharomycetales order (ABDEL-MAWGOUD et al., 2018; GROENEWALD et al., 2014). Identification through physiological growth characteristics places it near *Candida deformans*, for example (GROENEWALD et al., 2014). An historical overview of *Y. lipolytica* research, as well as of its industrial interest, is given in different review papers (MADZAK, 2021b; WOLF, 1996). *Y. lipolytica* presents a dimorphic morphology, and cells may organize themselves in free, yeast form or in pseudohyphae and septate hyphae. Predominant morphology in a given cell culture is dictated by strain and culture conditions (BARTH; GAILLARDIN, 1997; GÁLVEZ-LÓPEZ et al., 2019).

Y. lipolytica was named after its notorious ability to produce lipids, a trait that is vastly explored in biotechnology. High cytosolic acetyl-CoA pool and accumulation of fatty acids during growth in nutrient-limiting conditions are the main characteristics of this yeast lipid metabolism. It is stated that AMP is used as a nitrogen source in nitrogen-limiting media. Consequently, the isocitrate dehydrogenase enzyme in the tricarboxylic acid (TCA) cycle is inhibited due to low levels of AMP. Thus, citric acid accumulates, is transported to the cytoplasm and then converted to acetyl-CoA by ATP citrate lyase (POORINMOHAMMAD; KERKHOVEN, 2021). The resulting acetyl-CoA pool is drained mainly by synthesis of fatty acids, which are stored in the yeast lipid bodies as triacylglycerols, but also for the mevalonate pathway for ergosterol synthesis. Additionally, *Y. lipolytica* fatty acid turnover phase, that is, fatty acid degradation, may occur before depletion of carbon source in the medium, differently from *S. cerevisiae*, for example (MAKRI; FAKAS; AGGELIS, 2010). Because of this traits, *Y. lipolytica* has been used in lipid metabolism function studies and peroxissome and lipid body research (BEOPOULOS; CHARDOT; NICAUD, 2009; BREDEWEG et al., 2017; FAKAS, 2017).

This non-conventional yeast has increasingly gained visibility in the field of chemical biosynthesis process development (BAGHBAN et al., 2019; GÜNDÜZ ERGÜN et al., 2019;

VIEIRA GOMES et al., 2018), given not only its intrinsic high-flux lipid metabolism but also its distinct tolerance to organic compounds, high salt concentrations and low pH (MILLER; ALPER, 2019). Accordingly, there is a plethora of economically relevant molecules and proteins produced at laboratory level in this platform, both natural and heterologous, aiming to explore these traits (DARVISHI et al., 2018; MILLER; ALPER, 2019). As a non-conventional yeast, *Y. lipolytica* is less often employed in the industry than *S. cerevisiae*. Nevertheless, this oleaginous yeast has been employed since the 1950's as single cell protein grown in cheap substrates for animal feed (GROENEWALD et al., 2014). Other products obtained from *Y. lipolytica* in industrial processes are citrate, mannitol, erythritol and lipase as a therapeutic protein, for example (GROENEWALD et al., 2014; POORINMOHAMMAD; KERKHOVEN, 2021). Heterologous products, namely eicosapentaenoic acid (EPA) and carotenoid terpenes, have also reached commercialization and there are therapeutic heterologous proteins in preclinical trials (MADZAK, 2021a). Therefore, *Y. lipolytica* is an important biocatalyst platform for commercially relevant bioprocesses.

2.4.1 Genetic tools and resources

Although only a number of proteins are properly annotated, the genome of a number of *Y. lipolytica* strains have been sequenced, and their proteins have already been partially annotated (DEVILLERS et al., 2016; DEVILLERS; NEUVÉGLISE, 2019; MAGNAN et al., 2016; POORINMOHAMMAD; KERKHOVEN, 2021). Furthermore, the growing research investment in this microorganism is related to a recent increase in the number of developed and ready-to-use genetic manipulation tools (ABDEL-MAWGOUD et al., 2018; MUHAMMAD et al., 2020; SOONG et al., 2019). These tools allow for the exploitation of the attractive metabolic and physiologic traits of this host verified by the boost in heterologous proteins expressed (DARVISHI et al., 2018). A brief overview of the available resources for genetic manipulation of *Y. lipolytica* is described in this section.

Metabolic models may be explored in systems metabolic engineering as a means to glance into changes in the metabolism of the host cell before committing to laborious and costly genetic engineering. *Y. lipolytica* Genome-scale Metabolic models (GEMs) have been developed with a general focus on lipid or citrate and isocitrate production (ABDEL-MAWGOUD et al., 2018; GUO et al., 2022b; MADZAK, 2021b; POORINMOHAMMAD; KERKHOVEN, 2021), but the terpene β-carotene biosynthesis has also been considered (AMARADIO et al., 2022). These GEMs are important tools for evaluating candidates for

gene knockout, overexpression and heterologous expression targets, as the metabolic fluxes, product yields and overall cell performance can be assessed with high accuracy.

After choosing, with the aid of models or through prior knowledge-based design, the target genes to be manipulated in metabolic engineering, their expression cassettes must be constructed according to the genetic background of the host. Protein expression in *Y. lipolytica* may be dictated by strong constitutive promoters, like TEFin (TAI; STEPHANOPOULOS, 2013), or by inductive promoters that allow the control of transcript levels, such as XPR2 (MADZAK; GAILLARDIN; BECKERICH, 2004). On the one hand, constitutive expression is not recommended for the production of molecules toxic to the host. On the other hand, the requirement of induction factors and conditions is not attractive in industrial settings, as they increase costs (VIEIRA GOMES et al., 2018). An alternative for *Y. lipolytica* expression systems is the hybrid promoter hp4d (MADZAK; TRÉTON; BLANCHIN-ROLAND, 2000), widely employed in the scientific literature (CHUANG et al., 2010). This synthetic sequence was constructed using four Upstream Activating Sequences (UAS) from the XPR2 promoter in tandem and the minimal LEU2 promoter. The hp4d promoter is classified as quasi-constitutive, in which the transcription is conducted mostly at the beginning of the stationary phase of microbial growth (MADZAK; TRÉTON; BLANCHIN-ROLAND, 2000; NICAUD et al., 2002). Although weaker than the also frequently employed TEFin promoter (TAI; STEPHANOPOULOS, 2013), the hp4d synthetic promoter is a strong tool to partially decouple, without the need of induction, biomass formation in the exponential phase from the biosynthesis of a toxic product (MADZAK; GAILLARDIN; BECKERICH, 2004). However, the cellular mechanism involved in this activity related to the depletion of a limiting nutrient in the medium has not been described in the literature. Furthermore, some sources even consider the hp4d as a constitutive promoter (DULERMO et al., 2017). Nevertheless, promoter strength is a relevant issue when considering heterologous protein expression due to the reduction of cell fitness it imposes, often called metabolic burden. In *Y. lipolytica*, the relation between protein size, promoter used and metabolic burden has been investigated (KORPYS-WOŹNIAK et al., 2020). It was not found a general rule, although the authors pointed out that bigger proteins imposed higher metabolic burden than smaller ones, independently from their transcription rates.

In *Y. lipolytica*, the expression cassette is often cloned into an integrative vector, which fuses itself with the host genome in a random or predicted region, depending on the strategy employed. A plethora of integrative vectors is available, with a range of homology arms for

genome integration in different loci, dominant and auxotrophy-based selection markers and cloning regions for cassette insertion (MADZAK, 2021b). Episomal or replicative plasmids, which are replicated independently from the endogenous chromosomes, have only recently been developed for this host, and are still unstable compared to integrative vectors (WOLF, 1996). Due to low copy and mitosis instability (ABDEL-MAWGOUD et al., 2018; MADZAK, 2021b), it is reported difficulty in maintaining more than one *Y. lipolytica* episomal plasmid in the cell and therefore this type of vectors is used in specific cases for transient expression (HOLKENBRINK et al., 2018). Whether it is integrative or episomal, modular tools for assembly of these vectors with the desirable expression cassettes are available to facilitate genetic engineering. For example, GoldenGate and USER Cloning assembly methods have been developed and applied for *Y. lipolytica* genetic manipulation (DARVISHI et al., 2018; MADZAK, 2021b).

As for genome editing tools, there are records of the TALEN genome editing tool and the efficient CRISPR-Cas9 system adaptation to *Y. lipolytica*. These will be discussed below, with a focus on the latter one.

2.4.1.1 CRISPR and its use in *Yarrowia lipolytica*

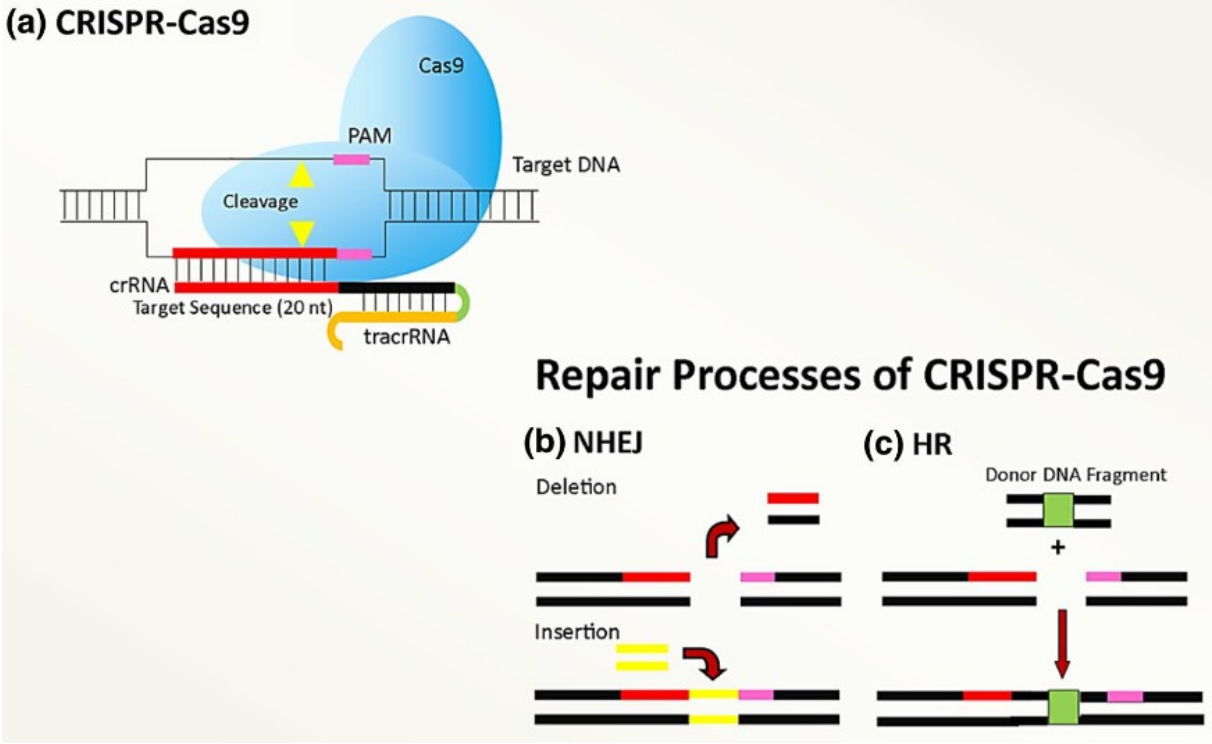
CRISPR was originally discovered in *E. coli* in 1987, but it wasn't until 2012 that its mechanisms got fully elucidated (ISHINO et al., 1987; JINEK et al., 2012; JAVED et al., 2018). Considered to be akin in function to the adaptive immune system, the CRISPR system activates to interrupt the infection of viruses in the cells of bacteria and archeae. It does that by inducing a double strand break (DSB) in the invading viral nucleic acid, thus disrupting it (JINEK et al., 2012). However, the CRISPR system is best known for its application in genetic engineering, for which it has been adapted to a myriad of hosts.

The most common usage of CRISPR for genome editing is based on two main components: the Cas9 (CRISPR-associated 9) protein and the sgRNA (synthetic guide ribonucleic acid). The sgRNA is composed of a tracrRNA (trans-activating CRISPR RNA) and a crRNA (CRISPR RNA) sequences, necessary for its coupling with the Cas9 protein and with the target DNA, respectively. As for the Cas9 enzyme, it is an endonuclease that requires the direction given by the sgRNA to generate the double-strand break in the DNA. On a mechanistic level, the Cas9 forms a complex with the sgRNA, which binds to the target DNA downstream of a NGG protospacer-adjacent motif (PAM), and then cleaves it (Figure 5a). Once the host has been transformed with expression cassettes for the Cas9 and sgRNA, the

cell may undergo two DNA repair pathways, NHEJ (non-homologous end joining) or HR (homologous repair) (LIU et al., 2022b). As NHEJ is conducted, base pairs near the extremities generated by the break are lost or inserted (Figure 5b). Consequently, genome regions are disrupted, causing gene knockouts, for example. Differently, for HR repair a third part is required, a donor DNA with extremities homologous to the sequences that flank the Cas-sgRNA complex cleavage site in the genome. This way, the cell uses the donor DNA as a template for repairing the double strand and sequences of interest, like an expression cassette, are integrated into the genome (Figure 5c) (STOVICEK; HOLKENBRINK; BORODINA, 2017). This has been demonstrated in bacteria, yeast and human cells shortly after the CRISPR function discovery in 2012 (JIANG et al., 2013; DICARLO et al., 2013; CONG et al., 2013; MALI et al., 2013).

The CRISPR system is useful in genetic engineering due to its ability to direct recombination in the genome. This way, unknown modifications in the DNA are unlikely to occur, which could hinder main metabolic pathways and the functioning of the cell. However, other tools like Zinc Finger Nucleases and Transcription Activator-Like Effector Nucleases (TALEN) are also able to generate target-oriented double strand breaks in the DNA. The advantages CRISPR shows over these include less strict rules for targeting the double strand break, simpler design as it does not require protein engineering, and higher scalability to target multiple regions of the genome simultaneously (JAVED et al., 2018). Therefore, CRISPR presents decreased time in genome editing experiments, high efficiency and straightforward editing of diploid microorganisms, among others (DARVISHI et al., 2018).

Figure 5: Components of the CRISPR-Cas9 system and some of its mechanisms for genome editing.



(A)CRISPR-Cas9 components: the sgRNA (composed of a tracrRNA and a specific crRNA), in conjunction with the Cas9 protein, promotes a double-strand break downstream of a PAM recognition sequence at the target genome sequence. (B) Non-Homologous End Joining: the cell machinery excises and the adds nucleotides to the ends generated by the double-strand break . (C) Homologous Repair: Given a donor DNA sequence (which may be a synthetic construct with an expression cassette, as represented by the green block), the cell may identify sequence similarity and use the donor DNA to repair the double-strand break. Source: (JAVED et al., 2018).

Nowadays, modifications to the CRISPR system described in the literature show different applications of this genetic tool (LIU et al., 2022b). For example, Cas9 proteins different from the one from *S. pyogenes*, which is the most frequently applied, may be utilized as they present different PAM sequence recognition patterns (COLLIAS; BEISEL, 2021). Also, ribonucleases from the Cas13 family can be employed to cleave RNA instead of DNA (ABUDAYYEH et al., 2016). Moreover, the Cas protein may be modified as to lose its endonuclease activity. In CRISPRi (CRISPR inactivation), in which the binding of a dead Cas9 (dCas9) to the DNA, without cleavage, impairs the transcription machinery operation, decreasing the expression of a target gene (QI et al., 2013). As a last example, a dCas9 fused with transcription activator or repressor effectors may be used to fine-tune target gene activity (GILBERT et al., 2014).

Due to *Y. lipolytica* potential as a industrial biotechnology chassis for market-relevant molecules, TALEN tools have been developed. However, the far more numerous application attempts of CRISPR to this host have been successfully achieved by a number of works, which are described in some review papers (DARVISHI et al., 2018; MADZAK, 2021b). A

relevant fact is that *Y. lipolytica* relies mostly on NHEJ for double-strand breaks in DNA. Among them, disruption of the *ku70* or *ku80* loci, required for NHEJ, is often carried out for increased HR mediated genome editing (SOONG et al., 2019; XIE, 2017). In this host, CRISPR has been employed for multiplex gene editing (BAE et al., 2020; GAO et al., 2016) and CRISPR interference (SCHWARTZ et al., 2017), for example. The evaluation of sgRNA design and integration sites regarding the application of this tool have also been researched (BAISYA et al., 2022; BORSENBERGER et al., 2018; SCHWARTZ et al., 2019).

An important example of CRISPR use in *Y. lipolytica* is the creation of the EasyCloneYALI kit. This genome editing toolbox provides five vectors for marker-free integration of DNA constructs in stable and non-coding regions and gene knockouts using CRISPR. Other eleven vectors bear selective markers and are available for integration in the *Y. lipolytica* genome (HOLKENBRINK et al., 2018). The kit is compatible with the USER cloning (NEB) technique for facilitated construction of DNA parts. It has been successfully validated by other works aiming at heterologous protein expression in this yeast (ARNESEN et al., 2020; MARELLA et al., 2020).

2.4.2 Bioprocesses with *Yarrowia lipolytica*

There is a predominance for fed-batch operation in *Y. lipolytica* processes. Possibility to control nutrient uptake and to decouple microbial growth from target molecule synthesis are relevant practices in this dimorphic yeast cultures (VANDERMIES; FICKERS, 2019). Despite this and the risks of contamination and of strain mutation, there is a sole example of industrial continuous bioprocess for the synthesis of single-cell oil in the 1970's (VANDERMIES; FICKERS, 2019; WOLF, 1996). Other industrial applications of this biocatalyst have been performed in fed-batch modes.

It is advocated that lipid synthesis benefits from a two-stage fermentation operation (XIE, 2017). Based in the oleaginous yeast lipid metabolism, it is established first a growth phase followed by a lipogenic phase, in which a nitrogen-limiting condition is enforced in the culture to increase lipid synthesis. As an illustration, this was a promising approach to increase lipid synthesis from volatile fatty acids (PEREIRA et al., 2022).

The influence of culture medium composition, especially carbon source, is also described for *Y. lipolytica* fermentation. One of the most common alternative substrates is glycerol. In one study, it was observed that glycerol was the favored carbon source in a mixture with glucose and its uptake was associated with higher citrate co-metabolite

production (PAPANIKOLAOU et al., 2003). Another example evaluated the use of a mixture of acetate, propionate, and butyrate in cultures based on volatile fatty acids as carbon source (PEREIRA et al., 2022). The addition of co-substrates, in this case, allowed simultaneous growth and lipid accumulation, although no difference in lipid yield was observed. In another example, *Y lipolytica* growth on saturated fatty acids from animal fat was conducted for single-cell oil production (PAPANIKOLAOU, I. CHEVALOT, M. KOMAI, 2002).

The influence of oxygen transfer and pH on *Y. lipolytica* cultures have also been described in the scientific literature. While sufficient oxygen transfer is required for cell growth as *Y. lipolytica* is strictly aerobic, higher aeration and agitation is also directly correlated with product formation, as shown for lipid production in (PEREIRA et al., 2022). However, in another study, 5 to 15% dissolved oxygen experiments showed higher lipid accumulation than the higher aerated (60 to 70 %) ones (PAPANIKOLAOU, I. CHEVALOT, M. KOMAI, 2002). This could be related to upregulation of the MVA pathway, as it is described that, for *S. cerevisiae*, oxygen deprivation increases expression of enzymes related to sterol synthesis (KLUG; DAUM, 2014). Such cellular mechanism might exist in *Y. lipolytica*, but no reports that describe this have been found. Finally, *Y. lipolytica* strains thrive in acidic pH, growing in pH of three to seven (WOLF, 1996). Indeed, it was found that pH 6 sustained highest lipid accumulation without compromising growth (PAPANIKOLAOU, I. CHEVALOT, M. KOMAI, 2002).

A modeling approach to bioprocesses can be applied as a means of investigating the effects of culture conditions on titers, rates and yields before committing to the costly empirical, wet-lab experiments. For *Y. lipolytica*, different modeling methods have been employed in lipid and citric acid production. In the first, growth and lipid metabolism was analyzed with a model considering the culture with industrial fats carbon source (PAPANIKOLAOU; AGGELIS, 2003). In another example, C/N ratio and nitrogen source were determined and successfully used for scaling-up the citric acid biosynthesis to 2 L bioreactors (GIACOMOBONO et al., 2022).

2.4.3 *Yarrowia lipolytica* for terpene production

Terpene synthesis in yeast was thought to occur exclusively by the MVA pathway. However, in the last few years there have been uncovered indications of the presence of the MEP pathway in fungi (SOLIMAN; TSAO; RAIZADA, 2011). Moreover, a recent thorough research endeavor provided three pieces of evidence that the MEP compound is present in *Y.*

lipolytica in nitrogen-limiting conditions and that its carbons are directed to a fraction of the ergosterol produced by the yeast (DISSOOK et al., 2021). Even though concomitant expression of MEP and MVA in *S. cerevisiae* is not predicted to increase IPP yield (GRUCHATTKA et al., 2013), the impact of this discovery in the terpene biosynthesis research field is noteworthy nonetheless. In spite of this, as the enzymes, mechanisms of activation and metabolic flux of the MEP pathway in *Y. lipolytica* are not yet ascertained, only the MVA pathway is considered for terpene heterologous production in this host.

Based on the high carbon flux through acetyl-CoA, *Y. lipolytica* lipid metabolism may be explored for terpene biosynthesis. Different strategies may be employed to divert acetyl-CoA from lipid biosynthesis to the MVA in this host. There is a plethora of publications describing the production of different terpene molecules in *Y. lipolytica*, and the field has been extensively reviewed (ARNESEN; BORODINA, 2022; LI et al., 2021b; LU et al., 2021; MA et al., 2019; ZHANG et al., 2021). This host has been modified for the biosynthesis of α-bisabolol (LI et al., 2022), *trans*-nerolidol (LIU et al., 2022a), astaxanthin (ZHU et al., 2022), zeaxanthin (XIE; CHEN; XIONG, 2021), the asiatic, madecassic, and arjunolic acid triterpenoids (ARNESEN et al., 2022b), amyrin (KONG et al., 2022), abscisic acid (ARNESEN et al., 2022a), amorphadiene (MARSAFARI; XU, 2020), oleanolic acid (LI et al., 2020), ginsenoside compound K (LI et al., 2019), betulinic acid (JIN et al., 2019), α- and β-farnesene (LIU et al., 2022c), β-ionone (LU et al., 2020) and α-humulene (GUO et al., 2022a). This variety of molecules, some of them the best with the highest reported titers in yeast yet, shows that the potential of *Y. lipolytica* for terpene synthesis is well acknowledged among the scientific environment. Some examples are further discussed below.

In (MATTHÄUS et al., 2014), the tetraterpenoid lycopene heterologous biosynthesis was explored, reaching 16 mg g^{-1} DCW^{-1}. It was observed a positive association between nitrogen limitation and specific lycopene content. Differently, oxygen limitation did not increase lycopene production. It is also reported that oleic acid used as carbon source, when compared to glucose or glycerol, increased both lipid body size and culture growth, but not lycopene accumulation. Furthermore, control of pH at 3.5, as opposed to 5.5, decreased citrate and isocitrate formation, which in turn led to a longer stationary growth phase and to an approximately 20 % higher lycopene synthesis (MATTHÄUS et al., 2014). For β-carotene production, (GAO et al., 2017) reported a production approximately 70 % higher when using a C/N ratio of 3:1.5 (wt / wt) compared to 2:3. Taking together the metabolic engineering of the tetraterpenoid pathway, medium optimization and fed-batch operation, 4 g L^{-1} β-carotene

was achieved. It is worth noting that genetically modified *Y. lipolytica* has been already employed for industrial production and commercialization of carotenoids in a process developed by Microbia (USA) and bought by DSM in 2010 (The Netherlands) (MADZAK, 2021b).

In relation to monoterpenes, α-pinene, linalool and limonene are the main compounds synthesized in *Y. lipolytica*. For α-pinene biosynthesis, cooking oil and lignocellulosic wastes were used as carbon sources to achieve more than 30 mg L^{-1} of the product, the highest titer in a yeast so far (WEI et al., 2021). Metabolic engineering and bioprocess optimization of *Y. lipolytica* for linalool synthesis was conducted (CAO et al., 2017). While citrate as the sole carbon source resulted in high productivity but low biomass formation, 10 g L^{-1} citrate plus 10 g L^{-1} glucose, glycerol or fructose improved cell growth without negatively affecting linalool specific production. The addition of at least 4 g L^{-1} pyruvate to the medium was also beneficial to linalool biosynthesis. Overall, combining overexpression of *HMG* and *IDI* MVA pathway genes with addition of pyruvic acid to the culture medium, flask production of linalool reached a concentration of 6.96 mg L^{-1}.

As for limonene, the same authors presented the proof-of-concept with the synthesis of 23.56 mg L^{-1} of this monoterpene in *Y. lipolytica* (CAO et al., 2016). Notwithstanding, this concentration was surpassed in yet another work by the same group (CHENG et al., 2019). The introduction of a second copy of the limonene synthase gene under the hp4d promoter control was carried out and displayed higher results compared to single copy. Eight different carbon sources were tested, and glycerol showed the highest specific production, 1.76 times higher than that of glucose. Interestingly, it was found that initial 20 g L^{-1} glycerol or 50 g L^{-1} of glucose resulted in the highest limonene productions, but also that higher concentrations of glycerol decreased limonene titers obtained in 72 h flask fermentation. Supplementation with 4 g L^{-1} citrate further increased the production. In fed-batch operation, 144 h fermentation using glycerol with addition of citrate in 48 h reached the final titer of 165.3 mg L^{-1} (CHENG et al., 2019). In other studies, metabolic engineering for xylitol uptake, use of cooking oil as alternative carbon source and directed evolution were also strategies applied for limonene synthesis in *Y. lipolytica* (LI et al., 2021a; PANG et al., 2019; YAO et al., 2020).

Recently, Arnesen and collaborators engineered *Y. lipolytica* strains aiming at increased production of each of the five terpene classes. Compared to the base strain with unmodified MVA pathway, the yeast modified for monoterpene synthesis produced almost 100 times more limonene (from 0.36 ± 0.04 mg L^{-1} to 35.9 ± 1.1 mg L^{-1}) (ARNESEN et al.,

2020). For genome editing, the parent strain employed expresses the *CAS9* gene, integrated in the *ku70* locus for impaired NHEJ repair (MARELLA et al., 2020). It was observed that *Y. lipolytica* tends to employ NHEJ repair when dealing with DSB (XIE, 2017). Therefore, *⊿ku70* disruption was applied as a way to increase the relative efficiency of HR-based genome editing in this host, given that the ku70 protein is related to NHEJ repair. This way, chance of integration of any donor DNA provided into the genome is higher than in the wild type strain. Mediated by this established CRISPR system, other seven genome modifications were conducted to yield ST9202, the monoterpene chassis strain (Figure 4). These alterations intend to increase the cytosolic pool of acetyl-CoA (*ACL* overexpression and expression of *SeACS*, the *Salmonella enterica* acetyl-CoA synthetase gene), to intensify the carbon flux through the MVA pathway (*HMG*, *ERG 12* and *IDI* overexpression), to boost GPP availability (*ERG 20*$^{F88W-N119W}$ expression) and to downregulate the squalene synthase gene (altering the pSQS1 promoter by the weak pERG11 promoter), which drains MVA products to ergosterol synthesis but cannot be deleted without compromising cell growth (ARNESEN et al., 2020). Therefore, the ST9202 host represents a starting metabolic chassis, with an initial round of metabolic engineering already performed. The only requirement for high-level monoterpene biosynthesis is the expression of the desired monoterpene synthase. Other works have employed the base strains for terpene production (ARNESEN et al., 2022a, 2022b), although use of the ST9202 strain has not been reported yet.

2.4.4 Comparison of hosts for geraniol biosynthesis

Compared to *E. coli* and *S. cerevisiae*, the most frequently used hosts for geraniol synthesis, main disadvantages of *Y. lipolytica* comprise fewer genetic manipulation tools, dimorphism between yeast and pseudohyphae and less described physiology, genetics and metabolism (ABDEL-MAWGOUD et al., 2018; VIEIRA GOMES et al., 2018). However, in recent years these hurdles have been surpassed by the increase in research publications in this yeast. Genome editing tools have been adapted (DARVISHI et al., 2018), with the successful development of the CRISPR system as an example (HOLKENBRINK et al., 2018). As for filament formation, it is avoidable during fermentation by control of C/N ratio, pH and osmotic and oxidative stresses (SOONG et al., 2019). Furthermore, some works have identified *Y. lipolytica* genes responsible for filament formation. Accordingly, their knockout led to cells in yeast morphology and to positive effects on β-carotene and lycopene production (LIU et al., 2021a; LUPISH et al., 2022). Lastly, *Y. lipolytica* possesses the ability to use alternative carbon sources not naturally metabolized by *E. coli* and *S. cerevisiae*. For example,

different terpenes, including limonene, have been biosynthesized in *Y. lipolytica* from waste cooking oil (GUO et al., 2022a; LIU et al., 2022c; PANG et al., 2019). Other cheap carbon sources, namely industrial by-products like alkanes and glycerol, may also be investigated for terpene synthesis (LI et al., 2021b).

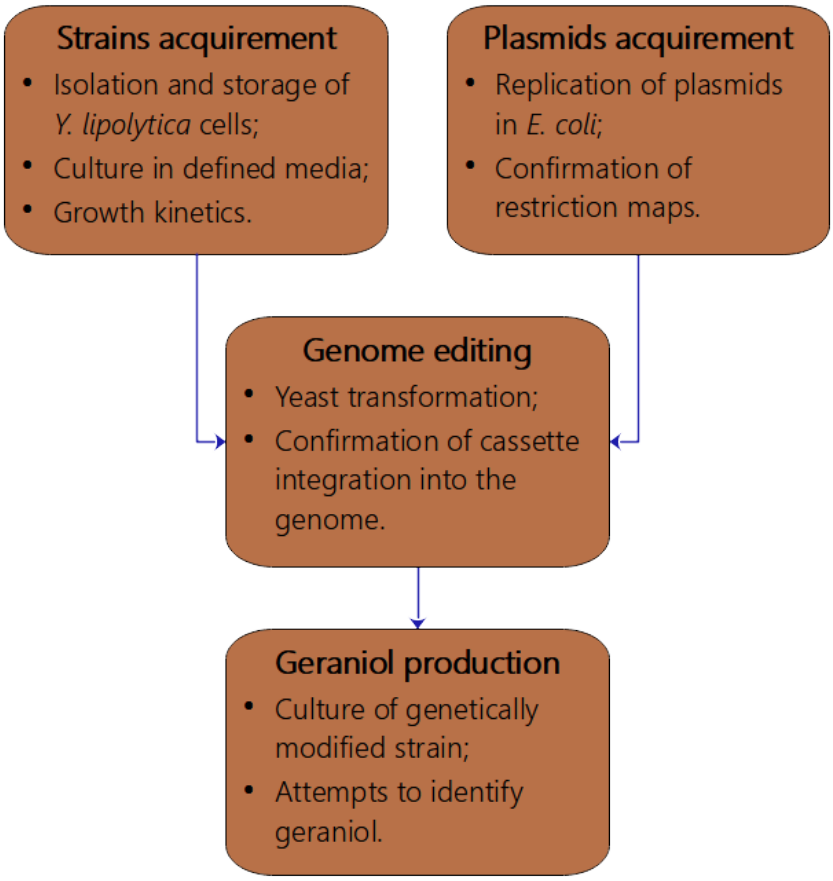
Regarding *E. coli* specifically, *Y. lipolytica* and other yeasts might display an advantage as bacterium growth is susceptible to geraniol impairment at lower concentrations (SUPPAKUL et al., 2003), although the concentration of complete inhibition is lower for yeasts, as seen in Section 2.3.4 (page 30). Moreover, yeasts are preferable hosts for further metabolic engineering aiming at MIA biosynthesis. MIA pathway enzymes require cytochrome P450 expression (WILSON; ROBERTS, 2012), which expression is impaired in bacteria (BROWN et al., 2015; MOSER; PICHLER, 2019). Finally, *Y. lipolytica* intrinsic ability to grow in different carbon sources, not naturally observed in *E. coli*, offers cost and operation versatility for industrial bioprocesses.

As for *S. cerevisiae*, the baker's yeast displays higher carbon flux through the MVA pathway than the oleaginous yeast (WRIESSNEGGER; PICHLER, 2013). Moreover, MIA *de novo* biosynthesis in the baker's yeast has been attempted in the last years, with some noteworthy works (BROWN et al., 2015; MISA et al., 2022; ZHANG et al., 2022). However, as an oleaginous yeast, *Y. lipolytica* high cytosolic acetyl-CoA pool may boost the flux through the pathway, as it was theoretically shown in models that acetyl-CoA location is a terpenoid production bottleneck in *S. cerevisiae* (GRUCHATTKA et al., 2013). Acetyl-CoA pools in *Y. lipolytica* have been manipulated for the metabolic engineering of different products, as the lipid metabolism in this yeast has been vastly explored (MILLER; ALPER, 2019; TAI; STEPHANOPOULOS, 2013). As an example, an approach of increasing lipid catabolism while decreasing fatty acid synthesis, for example, led to higher amorphadiene production (MARSAFARI; XU, 2020). Moreover, alcoholic fermentative metabolism, which is lacking in *Y. lipolytica*, decreases the carbon flux through the MVA pathway in the baker's yeast, due to unmet ATP requirements (GRUCHATTKA et al., 2013). Lastly, it can be hypothesized that the oleaginous yeast may show higher resistance to monoterpene toxicity than *S. cerevisiae*, given its known tolerance to other organic compounds and solvents (MILLER; ALPER, 2019; WRIESSNEGGER; PICHLER, 2013).

3 JUSTIFICATION

The main objective of this work was to develop a *Y. lipolytica* strain capable of synthesizing geraniol. This task was divided into the following specific objectives: (1) Acquirement of the ST9202 platform strain for monoterpene biosynthesis and the control strain ST6512, initial assessment of their growth physiology and evaluation of their growth in defined media to acquire basic know-how regarding *Y. lipolytica* manipulation; (2) Acquisition and replication of plasmids required for *Y. lipolytica* genome editing; (3) Transformation of the ST9202 strain (from step 1) with the *crGES* (geraniol synthase gene from the plant *C. roseus*) expression cassette (from step 2) and with the sgRNA plasmid required for genome editing through CRISPR, followed by verification of the integration by PCR to generate the geraniol-producing strain; (4) Attempts to identify geraniol in the culture medium and proof of geraniol biosynthesis by the genetically-modified strain obtained (from step 3). These objectives are described in the schematic in Figure 6.

Figure 6: Work objectives and main activities flowchart.



Source: Author.

Y. lipolytica high acetyl-CoA pool can be directed to the intrinsic MVA pathway, which would lead to elevated geraniol production through the heterologous expression of GES. In this context, the platform ST9202 strain represents an advantageous starting point for this molecule biosynthesis compared to the wild type strain. This is not only due to the expression of the Cas9 protein which simplifies genome editing steps, but also due to the metabolic engineering modifications it bears aiming at monoterpene production (ARNESEN et al., 2020), as described in section 2.4.3 (page 39). Furthermore, control of *GES* expression by hp4d promoter is expected to partially restrict the generation of the toxic monoterpene product to the start of the stationary growth phase, after biomass accumulation, without the need of induction (MADZAK; TRÉTON; BLANCHIN-ROLAND, 2000). Thus, given the metabolic and physiologic traits and the records of successful heterologous production of other terpenes in *Y. lipolytica*, it is regarded as a promising host for the biosynthesis of geraniol.

The main application of the line of research started with this work is the development of a bioprocess for geraniol synthesis. This would not only be potentially more sustainable and cheaper than current means of geraniol production but would also enable the synthesis of relevant but structurally complex MIA pharmaceuticals. Accordingly, the development of new biotechnological paths for biomolecules synthesis is a promising field of research both to public and to private institutions. Therefore, this work is also relevant by pushing the knowledge frontier regarding to *Y. lipolytica* application and heterologous protein expression in this host in general and by establishing the foundations for metabolic engineering for terpene biosynthesis in the Bioprocess Engineering Laboratory (BELa-USP).

4 MATERIALS AND METHODS

4.1 *Yarrowia lipolytica* culture and manipulation

4.1.1 Culture media and analytical methods

The complex medium used was YP (10 g L^{-1} yeast extract, 20 g L^{-1} peptone and, if necessary, 15 g L^{-1} agar). Growth was also attempted on two different defined media, the Glucose minimal medium (which will be called Matthäus defined medium or MDM) and the Minimal medium with thiamine (which will be called Barth defined medium or BDM). One liter of the MDM, according to (MATTHÄUS et al., 2014), is composed of the four following solutions, mixed in aseptic conditions: 100 mL of carbon source solution with 500 g L^{-1} glucose or glycerol, 898 mL of main solution with 1 g L^{-1} KH_2PO_4, 0.16 g L^{-1} $K_2HPO_4 \cdot 3H_2O$, 3 g L^{-1} $(NH_4)_2SO_4$, 0.7 g L^{-1} $MgSO_4 \cdot 7H_2O$, 0.5 g L^{-1} NaCl and 0.4 g L^{-1} $Ca(NO_3)_2 \cdot 4H_2O$, 1 mL of trace elements solution with 0.5 g L^{-1} H_3BO_3, 0.04 g L^{-1} $CuSO_4 \cdot 5H_2O$, 0.1 g L^{-1} KI, 0.4 g L^{-1} $MnSO_4 \cdot 4H_2O$, 0.2 g L^{-1} $Na_2MoO_4 \cdot 2H_2O$, 0.4 g L^{-1} $ZnSO_4 \cdot 7H_2O$ and 6 g L^{-1} $FeCl_3 \cdot 6H_2O$ and 1 mL of thiamine solution with 0.3 g L^{-1} thiamine hydrochloride. Adaptations to this medium done in this work include use of 200 g L^{-1} glucose solution and substitution of the following reagents: 0.16 g L^{-1} $K_2HPO_4 \cdot 3H_2O$ to 0.122 g L^{-1} K_2HPO_4 and 0.4 g L^{-1} $MnSO_4 \cdot 4H_2O$ to 0.303 g L^{-1} $MnSO_4 \cdot 1H_2O$. The glucose solution and the main solution were sterilized by autoclave at 121º C for 30 min and the trace element and thiamine solutions were filtered through 0.22 µm PVDF filters. The other medium employed was the BDM, from (WOLF, 1996), which is composed of the following solutions: 100 mL of carbon source solution with 200 g L^{-1} of glucose, 898 mL of main solution with 5 g L^{-1} $(NH_4)H_2PO_4$, 2.5 g L^{-1} KH_2PO_4 and 1 g L^{-1} $MgSO_4 \cdot 7H_2O$, 1 mL of trace elements solution with 0.4 g L^{-1} $ZnSO_4 \cdot 7H_2O$, 0.4 g L^{-1} $MnSO_4 \cdot 4H_2O$, 0.1 g L^{-1} $CoCl_2$, 0.1 g L^{-1} $CuSO_4 \cdot 5H_2O$, 0.2 g L^{-1} Na_2MoO_4, 20 g L^{-1} $Ca(NO_3)_2 \cdot 4H_2O$, 2 g L^{-1} $FeCl_3 \cdot 6H_2O$, 0.5 g L^{-1} H_3BO_3 and 0.1 g L^{-1} KI and 1 mL of thiamine solution with 0.3 g L^{-1} thiamine hydrochloride. Adaptations to this medium done in this work include addition of 15 g L^{-1} of disodium EDTA to the trace elements solution and substitution of the following reagents: 5 g L^{-1} $(NH_4)H_2PO_4$ to 3 g L^{-1} $(NH_4)_2SO_4$, 0.4 g L^{-1} $MnSO_4 \cdot 4H_2O$ to 0.31 g L^{-1} $MnSO_4 \cdot 1H_2O$, 0.1 g L^{-1} $CoCl_2$ to 0.183 g L^{-1} $CoCl_2 \cdot 6H_2O$ and 0.2 g L^{-1} Na_2MoO_4 to 0.235 g L^{-1} $Na_2MoO_4 \cdot 2H_2O$.

For YP medium, MDM and BDM, glucose (D for 20 g L^{-1}, 4D for 80 g L^{-1}) was used as carbon source. This solution were sterilized by autoclave at 121º C for 30 min separately from the media and added under aseptic conditions. For defined media, thiamine solution was filtered through 0.22 μm PVDF filters. Both autoclaved and filtered trace elements solutions were tested. Compositions of media used in this work for *Y. lipolytica* growth are shown in Table 3.

Table 3: Media used for *Y. lipolytica* growth and their compositions (g L^{-1}).

	YPD medium	Modified MDM-D	Modified BDM-D
$Ca(NO_3)_2 \cdot 4H_2O$	-	$4 \cdot 10^{-1}$	$2 \cdot 10^{-2}$
$CoCl_2 \cdot 6H_2O$	-	-	$1.83 \cdot 10^{-4}$
$CuSO_4 \cdot 5H_2O$	-	$4 \cdot 10^{-5}$	$1 \cdot 10^{-4}$
EDTA	-	-	$1.5 \cdot 10^{-2}$
$FeCl_3 \cdot 6H_2O$	-	$6 \cdot 10^{-3}$	$2 \cdot 10^{-3}$
Glucose	20	20	20
H_3BO_3	-	$5 \cdot 10^{-4}$	$5 \cdot 10^{-4}$
K_2HPO_4	-	$1.22 \cdot 10^{-1}$	-
KH_2PO_4	-	1	2.5
KI	-	$1 \cdot 10^{-4}$	$1 \cdot 10^{-4}$
$MgSO_4 \cdot 7H_2O$	-	$7 \cdot 10^{-1}$	1
$MnSO_4 \cdot H_2O$	-	$3.03 \cdot 10^{-4}$	$3.1 \cdot 10^{-4}$
$Na_2MoO_4 \cdot 2H_2O$	-	$2 \cdot 10^{-4}$	$2.35 \cdot 10^{-4}$
NaCl	-	$5 \cdot 10^{-1}$	-
$(NH_4)_2SO_4$	-	3	3
Peptone	20	-	-
Thiamine	-	$3 \cdot 10^{-4}$	$3 \cdot 10^{-4}$
Yeast Extract	10	-	-
$ZnSO_4 \cdot 7H_2O$	-	$4 \cdot 10^{-4}$	$4 \cdot 10^{-4}$

For OD_{600} measurement with the Shimadzu (Japan) UV-2600 spectrophotometer, samples were diluted so as to return results within the 0.1 to 0.4 range. The OD_{600} measured was multiplied by the dilution factor to find the cultures optical density values. For estimation of the number of cells per mL of culture, non-viable cells were stained with 0,1 M final concentration of methylene blue and viable cells were counted in the Neubauer chamber. Observation and counting of cells were made with the Zeiss (Germany) AXIO Scope.A1 microscope.

In the HPLC analysis for glucose and citrate/isocitrate (which exhibit the same retention time, thus are indistinguishable by this method), the HPX-87H Aminex column from Bio-Rad (USA) was employed. Samples were run with a 0.5 M H_2SO_4 mobile phase at a 0.6 mL min^{-1} injection rate, 60º C oven temperature, with 15 min between injections. Glucose was

detected with a RID whereas citrate and isocitrate detection was tested with a UV detector at the 210 nm wavelength, as in (MAKRI; FAKAS; AGGELIS, 2010). Concentration of different compunds were determined using standard-grade reagents from Sigma Aldrich (USA).

Geraniol quantification was conducted in the GCMS-QP2020 equipment from Shimadzu (Japan), with SH-Rtx-5ms column (0.25 µm DF, 0.25 mm ID and 30 m length) from Shimadzu (Japan). The method applied was based in (ZHAO et al., 2017), with modifications to the column temperature program: 60º C for 2 min, heating at 5º C min^{-1} to 150º C, 10 min at 150º C, heating at 20º C min^{-1} to 280º C and 5 min at 280º C. Geraniol and limonene mass spectra were identified by the software and geraniol retention time and concentration was determined with the geraniol standard acquired from Sigma Aldrich (USA). The calibration curve for geraniol concentration estimation was composed of five solutions of geraniol standard in dodecane, at concentrations of 1.778, 0.889, 0.356, 0.178 and 0.089 g L^{-1}.

4.1.2 *Yarrowia lipolytica* strains

Y. lipoytica strains used in this work are shown in Table 4. Strains ST6512 and ST9202 are based in the W29 strain and were provided by PhD Prof. Irina Borodina research group. Upon receipt, strains were grown in 2 mL YPD overnight and then inoculated in YPD-agar plates, incubated for at least 48 h. Colonies were streaked a second time to ensure there was no contamination. Colonies obtained were then grown in 30 mL of YPD in 125 mL flasks at 30 ºC and 200 rpm for approximately 24 h before being frozen in stock aliquots as described below. Strains ST9250 and BYa3105 were obtained by transformation of the ST9202 strain, as described below in section 4.3 (page 55).

Table 4: *Y. lipolytica* strains used in this work and their characteristics.

Strain	Base strain	Genotype characteristics	Reference
W29	-	Wild type isolate, MatA	(GAILLARDIN; CHAROY; HESLOT, 1973)
ST6512	W29	MatA ku70Δ::PrTEF1→*SpCas9*-TTef12::PrGPD→*DsdA*-TLip2	(MARELLA et al., 2020)
ST9202	ST6512	ST6512 background IntC_2-*HMG1*←PrGPD-PrTefInt→*ERG12* IntC_3-SeACS←PrGPD-PrTefInt→*YlACL1* pERG11::pSQS1 IntD_1-*IDI1*←PrGPD-PrTefInt→ $ERG20^{F88W-N119W}$	(ARNESEN et al., 2020)
ST9250	ST9202	ST9202 background IntE_1-PrTefInt→*PfLS*	(ARNESEN et al., 2020)
BYa3105	ST9202	ST9202 background IntE_1-hp4d→*CrGES*-TLip2	This work

Yeast frozen stocks were prepared as follows. Cells grown in YPD for approximately 24 h were harvested by centrifugation at 3000 rpm for 10 min at a Hitachi (Japan) Himac CT 15RE centrifuge. The supernatant was discarded and the pellet was washed in sterile water. Centrifugation and washing were repeated once. Then, sterile glycerol was added to the water suspension of cells to the final concentration of 20 % (v/v). Aliquots of 1 mL were dispensed in cryogenic tubes, which were stored in a – 80º C ultrafreezer.

4.1.3 Culture and growth kinetics

Growth of *Y. lipolytica* on both defined media (MDM-D and BDM-D) was first tested in 125 mL erlenmeyer flasks, with 20 mL of medium. Inocula from ST6512 and ST9202 frozen stocks were conducted. Every test was accompanied with a negative control with no inoculum and a positive control in YPD medium. Incubation was done at 30º C and 200 rpm for at least 24 h. Growth on defined media was compared to growth in YPD by OD_{600} and cell appearance in the microscope, conducted as described in section 4.1.1 (page 46).

First assessment of *Y. lipolytica* growth physiology was conducted in YPD medium, at 30º C and 200 rpm. This is called culture experiment “A” throughout this work. As inoculum cultures, ST6512 and ST9202 cell stocks were directly used to inoculate 25 mL (in 250 mL flasks) cultures in YPD and were incubated for approximately 36 h at 30º C and 200 rpm. The

cells were washed once in sterile medium and then these inoculum cultures were used to inoculate 50 mL of medium in 500 mL shake flasks at an initial OD_{600} of approximately 0.2. Cultures were incubated at 30º C and 200 rpm. The OD_{600} was measured at 0, 3, 6, 9, 12, 24, 28, 48 and 54 h. Aliquots were sampled at 0, 6, 12, 24, 48 and 54 h for HPLC analysis.

Partial reproduction of experiment A was conducted, with the following modifications, in culture experiment "B". Inoculum cultures were prepared in 20 mL of YPD medium (in 125 mL flasks) and incubated at 28º C and 250 rpm for approximately 18 h. The cells were washed once in sterile medium and then these inoculum cultures were used to inoculate 100 mL of medium in 500 mL baffled shake flasks, at an initial OD_{600} of approximately 0.3. Cultures were incubated at 28º C and 250 rpm. The OD_{600} was measured at 0, 3, 4, 6, 9, 12, 23.5, 28 and 31 h. Aliquots were sampled at 0, 6, 12, 23.5 and 28 h for HPLC analysis.

The culture experiment "C" was conducted in BDM-D and was divided in two parts. For both of them, inoculum cultures of the W29 (which was then available as control strain) and ST9202 strains were prepared as for experiment "B", but in BDM-D. Cells were washed once and used to inoculate 100 mL of BDM-D in erlenmeyer flasks with baffles at an initial OD_{600} of approximately 0.3. Flasks were incubated at 28º C and 250 rpm. For the first part, named "CI", the inoculum cultures were incubated for 20 h, the OD_{600} was measured at 0, 3, 6, 9, 12, 24, 27 and 30 h and aliquots were sampled at 0, 6, 12, 24 and 30 h for HPLC analysis. As for the second part, "CII", the inoculum cultures were incubated for 23h, the OD_{600} was measured at 0, 12, 15, 18, 21 and 24 h and aliquots were sampled at 0, 12, 18, 21 and 24 h for HPLC analysis. Experiment "CII" was required as cell growth phases were not fully observed in "CI".

In all three experiments, inocula were made in triplicates and negative controls without cell inocula were conducted to readily discover any accidental contamination related to the media manipulation. No more than 10 % of the initial culture volumes was sampled throughout the experiments. The OD_{600} data of cultures were transformed with the natural logarithm and plotted against time of experiment: $\ln(OD_{600})$ at the y axis and time (in h) at the x axis. Then, it was determined linear regression models fit to at least four of the points within the identified exponential growth phases. The angular coefficients of the equation obtained from the regression line was reported as the maximum specific growth rates (μ_{Max}, in h^{-1}) of the cultures. As for the HPLC data obtained, glucose concentration data was plotted against time of sampling: glucose concentration (in g L^{-1}) at the y axis against time (in h) at the x axis. Dispersion measures such as mean and standard deviation estimates were conducted in the

software LibreOffice Calc 7.2. Differences between μ_{Max} means of treatments were estimated with the t-test with Welch's correction for unequal variances, as the sampling size (n = 3) was small. The statistical test was conducted in the software GraphPad Prism 8.

An overview of *Y. lipolytica* growth experiments described in this section is shown in Table 5 below.

Table 5: Different experiments of *Y. lipolytica* culture conducted in this work, their main conditions and objectives.

Experiment	Strains	Medium	Volume (mL)	Shake flask	Temperature (º C)	Agitation (rpm)	Objective
MDM-D first screening	ST6512 and ST9202	MDM-D and a control in YPD	20	Common	30	200	Test the growth in the Matthäus Defined Medium
BDM-D first screening	ST6512 and ST9202	BDM-D and a control in YPD	20	Common	30	200	Test the growth in the Barth Defined Medium
A	ST6512 and ST9202	YPD	50	Common	30	200	Analyze growth kinetics in YPD medium
B	ST6512 and ST9202	YPD	100	Baffled	28	250	Analyze growth kinetics in YPD medium
CI	W29 and ST9202	BDM-D	100	Baffled	28	250	Analyze growth kinetics in BDM-D medium
CII	W29 and ST9202	BDM-D	100	Baffled	28	250	Analyze growth kinetics in BDM-D medium

4.1.4 Geraniol production

Assessment of geraniol biosynthesis by genetically modified BYa3105 strain was conducted based in the methods of (ARNESEN et al., 2020; CAO et al., 2016), with changes described below. Cells from BYa3105, ST9250 and ST9202 were streaked in YPD-agar plates from frozen stocks and incubated at 30º C for at least 24 h. Isolated colonies were harvested and used for inoculum cultures in 125 mL flasks with 20 mL of YP4D (YP medium with 80 g L^{-1} glucose). After approximately 18 h of incubation at 30º C and 250 rpm, inoculum cultures were used to inoculate 20 mL of YP4D in 125 mL flasks at an initial OD_{600} of 0.1. Prior to inoculation and incubation at 30º C and 250 rpm, 2 mL of dodecane overlay, as terpene extractant, was added to these cultures. Duplicates were conducted for each strain, and a

negative control without inoculum was performed to verify possible contamination due to manipulation or from the dodecane, which was not sterilized.

After incubation for approximately 72 h, cultures were transferred to 50 mL tubes and centrifuged at 12000 g for 5 min. Then, the dodecane upper phase was harvested and transferred to vials, which were stored at approximately -4º C until analysis. Geraniol identification and quantification was performed in GC-MS with these dodecane samples, as described above in section 4.1.1 (page 46). Geraniol concentrations in the 2 mL of dodecane overlays analyzed in GC-MS were divided by 10 to generate the estimation of geraniol produced in the 20 mL of culture medium, assuming a complete separation of geraniol to this dodecane organic phase.

4.2 Molecular Biology practices and methods

4.2.1 Manipulation of DNA

Agarose gels for DNA visualization were prepared with TAE buffer (40 mM Tris-Acetate and 1 mM EDTA) and stained with SYBR Safe from Sigma Aldrich (USA). Agarose was acquired from Thermo Fisher Scientific (USA) and its brands. The PowerPac Power Supply from Bio-Rad (USA) was used to provide the voltage difference for DNA migration and agarose gels were visualized with the ImageQuant LAS 4000 Mini equipment from GE Healthcare Life Sciences (USA). Length of DNA fragments in agarose gels were estimated by comparison with fragments from GeneRuler 1 kb DNA Ladder or 1 kb Plus DNA Ladder from Thermo Fisher Scientific (USA), shown in Appendices Figure 1.

Treatment of DNA with restriction endonucleases was conducted using enzymes from NEB (USA) or Thermo Fisher Scientific (USA), according to the instruction manuals from the manufacturers. Correct cleavage was verified in agarose gels, prepared as described above. If required, the digestion reactions were purified using the PureLink PCR Micro Kit from Thermo Fisher Scientific (USA).

ReadyMix *Taq* PCR Reaction Mix with $MgCl_2$, Platinum *Taq* DNA Polymerase, dNTP solutions and $MgCl_2$ solution were acquired from Thermo Fisher Scientific (USA) and its brands. For PCR, reagent concentrations used were based on the manufacturer instruction manuals and reactions were incubated in Veriti Thermal Cycler from Thermo Fisher Scientific

(USA). DNA oligos sequences are described in Table 6. All primers in this work were acquired from Exxtend (Brazil).

To amplify regions of the *Y. lipolytica* genome, colony PCR was conducted. For this, cell lisates were used to provide the DNA template. Cell lysis was carried out with suspension of cells in 20 µL of a 20 mM NaOH solution and heating at 70 ºC for 15 min. Cells used were harvested from colonies grown in YPD-agar for approximately 24 h. Alternatively to colony PCR, cells were grown in 50 mL tubes with 5 mL of YPD medium at 30 ºC and 250 rpm for approximately 18 h and used for genomic DNA extraction using the YeaStar Genomic DNA kit from Zymo Research (USA). Then, the genomic DNA solutions obtained were used as template in PCR. Reactions and thermocycler program were set as recommended by manufacturers. Amplification was verified in agarose gels.

Table 6: Primers (DNA oligonucleotides) used in this work.

Primer name and orientation	Sequence (5'→3')	Annealing site	Annealing temp. used	Source
PR-14442 (Forward)	agttgtgaccaagacaaatg	Upstream of IntE_1 upper homology arm	52 ºC	(HOLKENBRINK et al., 2018)
PR-14398 (Reverse)	gttagaagcaattggagaag	3' region of the IntE_1 lower homology arm	52 ºC	(HOLKENBRINK et al., 2018)
PR-InDFor (Forward)	gtcaatgagctggtatagac	TLip2 Terminator	54 ºC	This work
PR-InDRev (Reverse)	caaggtggtctttacaatgc	Downstream of IntE_1 lower homology arm	54 ºC	This work

4.2.2 Plasmid replication

Plasmids employed in this work are described in Table 7, and their maps and sequences are shown in the Appendices Figures 10 to 13. The pCfB6633, pCfB6677 and pCfB8828 plasmids were kindly provided by Prof. PhD Irina Borodina from the Technical University of Denmark (ARNESEN et al., 2020; HOLKENBRINK et al., 2018). The pCfB6633 plasmid allows the expression of a sgRNA which directs a Cas9-mediated double-strand DNA break to the IntE_1 genomic region. It also bears the gene for Neourseothricin resistance for yeast selection. As for the pCfB6677 plasmid, it has two sequences homologous to parts of IntE_1 for integration at this locus, in addition to terminators and a region for gene and promoter cloning. Therefore, this vector acts as a skeleton for assembly of expression cassettes ready for integration into the IntE_1 site. Finally, the pCfB6677 vector had been

cloned with a limonene synthase expression cassette from *Perilla frutescens* (*pfLS*) to yield the pCfB8828 plasmid (ARNESEN et al., 2020). As its transformation had been already performed and used to biosynthesize limonene in *Y. lipolytica*, this plasmid may be used as a positive control for transformation steps.

Table 7: Plasmids used in this work.

Plasmid	Relevant features	Source
pCfB6633	Nat resistance, IntE_1 sgRNA	(HOLKENBRINK et al., 2018)
pCfB6677	IntE_1 homology arms, cloning site	(HOLKENBRINK et al., 2018)
pCfB8828	IntE_1 homology arms, *pfLS* expression cassette	(ARNESEN et al., 2020)
pUC57_crGES	IntE_1 homology arms, *crGES* expression cassette	This work

The pUC57_crGES plasmid was obtained from GenScript (USA), through its representative company FastBio (Brazil). This plasmid contains the *C. roseus* geraniol synthase gene expression cassette (*crGES*) flanked by sequences homologous to the IntE_1 region in the genome of *Y. lipolytica* (Figure 7). The *crGES* gene sequence (accession number: KF561459) was optimized for *Y. lipolytica* and truncated at the 126 base position. This removes the first 42 amino acids from the encoded protein, which represent the signal peptide for protein mobilization to plant cell plastids. It has been described in the literature that this signal peptide decreases the activity of terpene synthase enzymes in the yeast cytosol (JIANG et al., 2017; SIMKIN et al., 2013). Moreover, among four different possible locations for truncation tested in (JIANG et al., 2017), the removal of the first 42 amino acids achieved the highest geraniol production in *S. cerevisiae*. The hp4d promoter sequence, including the used Kovak consensus sequence, was provided by Prof. PhD Catherine Madzak (MADZAK; GAILLARDIN; BECKERICH, 2004; MADZAK; TRÉTON; BLANCHIN-ROLAND, 2000). The Tlip2 terminator sequence was chosen from (ARNESEN et al., 2020), as it is the terminator used for *pfLS* expression in pCfB8828.

Figure 7: Representation of the *crGES* expression cassette (4085 bp) and its features in the pUC57_crGES plasmid.

IntE_1 UP (498 bp): Upstream IntE_1 homology arm. hp4d (542 bp): *Y. lipolytica* synthetic promoter. *crGES* (1644 bp): Codon-optimized and truncated geraniol synthase gene from *C. roseus*. Tlip2 (937 bp): *Y. lipolytica* endogenous terminator. IntE_1 DOWN (464 bp): Downstream IntE_1 homology arm. Source: Author. Image from SnapGene.

Plasmids used in this work were all shuttle vectors with bacterial origin of replication and ampicillin resistance gene. For plasmid replication, calcium chloride transformation protocol of *E. coli* strain DH5α was conducted as in (SAMBROOK; RUSSELL, 2001), with minor changes. The frozen competent cell solutions (50 µL), acquired as described in the protocol, were incubated in ice for 30 min after thawed and being mixed with plasmid solution (up to 10 µL). Heat-shock was conducted at 42º C for 2 min followed by 5 min in ice. For cell recovery, 270 µL of LB medium (0.5 % yeast extract, 1 % peptone and 1 % NaCl) was added to the tube, which was then incubated for 1 h at 37º C and 225 rpm. Cells were plated in LB-agar medium (1.5 % agar added to LB medium as described above) with 100 mg L^{-1} ampicillin. Petri dishes were incubated for approximately 12 h at 37º C. Negative controls were conducted with cells that received water instead of plasmid DNA solutions at the transformation protocol.

Isolated transformant colonies were selected and grown overnight in 5 mL LB medium with 100 mg L^{-1} ampicillin in 50 mL tubes. The required amount of cells, verified by OD_{600} measurements, was used for miniprep using the Promega (USA) PureYield Plasmid Miniprep Kit. The remainder of cell cultures were frozen at – 80º C after addition of 20 % (v/v) of glycerol. Purified plasmid solutions were quantified and applied in agarose gels to verify plasmid integrity. These plasmid DNA solutions were treated with specific endonucleases to confirm the locations of restriction sites, compared to the sequences obtained for each plasmid.

4.3 *Yarrowia lipolytica* transformation

The ST9202 strain was used to establish a calibration curve between OD_{600} and number of cells per mL of culture. A inoculum culture prepared as in experiments “B” and

"C" described in section 4.1.3 (page 49) was used to inoculate 100 mL of YPD in baffled erlenmeyer flasks at the initial OD_{600} of 0.3 and 0.7, both in triplicates. Cultures were incubated at 28º C and 250 rpm. The OD_{600} was measured and the viable cells were counted in the Neubauer chamber (as described in section 4.1.1, page 46) at 0, 3, 6, 9, 12 and 25.5 h. The OD_{600} was plotted against the number of cells per mL obtained and a linear regression model was fit to the data (Appendices Figure 2). The linear regression equation obtained was used estimate, from OD_{600} measurements, approximately $5 \cdot 10^7$ cells, number required for the transformation protocol.

The pUC57_crGES and the pCfB8828 plasmids were treated with *Hin*dIII and *Not*I restriction enzymes, respectively, to yield the fragments with *crGES* and *pfLS* expression cassettes flanked by the homology arms, which are sequences with similarity to the IntE_1 locus in the genome. Cleavage was verified and reactions were purified as described in section 4.2.2 (page 52). The DNA solutions obtained were separately transformed into the ST9202 strain along with the pCfB6633 plasmid.

The *Y. lipolytica* transformation protocol conducted was based in (CHEN; BECKERICH; GAILLARDIN, 1997; HOLKENBRINK et al., 2018), with minor changes as described below.

The PEG stock solution of the transformation mix prepared was more concentrated than the described in the protocol, which would yield a transformation mix with final volume of 85 µL, not 100µL. This was devised to account for the extra 15 to 20 µL volume from the DNA solutions used, as opposed to only 1 µL described in the protocol. Therefore, the final concentration of components would be the established in the protocol. Another modification was the use of PEG-3350 instead of PEG-4000 as advised in (CHEN; BECKERICH; GAILLARDIN, 1997), due to reagent unavailability in the lab. All reagents for the transformation mix were purchased from Sigma Aldrich (USA). DNA sodium salt from salmon testes was dissolved in Milli-Q water and then boiled at 99 ºC for 10 min followed by 100 ºC for 10 min to form the ssDNA (single-stranded DNA from salmon testes) solution. Stock solutions were mixed into the transformation mix before use as described in Table 8.

Table 8: *Yarrowia lipolytica* transformation mix composition.

Stock solutions	Concentration of stock solutions used	Volume of stock solutions added in transformation mix used (μL)	Approximate final concentration in the mix after addition of DNA solutions
PEG-3350	62 % (w/v)	72.5	44.95 % (w/v)
Lithium acetate	2 M	5	0.1 M
ssDNA	10 g L^{-1}	2.5	0.25 g L^{-1}
DTT	2 M	5	0.1 M

Harvest, washing, incubation with DNA and transformation mix and recovery of *Y. lipolytica* cells were conducted as described in (HOLKENBRINK et al., 2018). After recovery, cell suspensions were inoculated in YPD-agar medium with 250 mg L^{-1} nourseothricin using Drigalski spatulas. Prior to inoculation, plates had been left open inside the biological safety cabin for 1 h to dry. According to (CHEN; BECKERICH; GAILLARDIN, 1997), this increases transformation efficiency. Inoculated plates were incubated at 30 ºC for approximately 62 h.

The ST9250 and BYa3105 colonies obtained were harvested and used to inoculate 50 mL tubes with 5 mL of YPD medium with 250 mg L^{-1} nourseothricin, which were then incubated at 30 ºC and 250 rpm for approximately 18 h. The appropriate amount of cells was harvested for genomic DNA extraction as described in section 4.2.2 (page 52) and the remainder of cultures were used in frozen glycerol stocks of transformant cells, as described in section 4.1.2 (page 48).

5 RESULTS AND DISCUSSION

5.1 *Yarrowia lipolytica* morphology and physiology

5.1.1 Isolation and colony morphology

The *Y. lipolytica* strains used in this section are W29, ST6512 and ST9202. Strain ST6512 is based on the wild type W29 strain (MARELLA et al., 2020). It was modified with the integration of a codon optimized *Cas9* expression cassette from *S. pyogenes* in the *ku70* locus. Finally, to generate strain ST9202, *Cas9* expression allowed for other seven genomic modifications applied in strain ST6512, that aimed to increase monoterpene biosynthesis, as described in Section 2.4.3 (page 39). These modifications result in the increase of the cytosolic acetyl-CoA pool, in the intensified carbon flux through the MVA pathway, boost in GPP availability and in the downregulation of the squalene synthase gene. Thus, the ST9202 cell line is considered a monoterpene production chassis (ARNESEN et al., 2020).

Cells from strains ST6512 and ST9202 were kindly provided by Prof. Irina Borodina from the Technical University of Denmark (ARNESEN et al., 2020; MARELLA et al., 2020). Upon acquisition, these strains were inoculated in YPD-agar medium plates and were re-streaked once before being frozen as glycerol (20 % v/v) stocks for later use. This was performed to confirm that no bacterial contamination had arisen during transportation and manipulation of cells. As seen by the petri dishes (Appendices Figure 3), no contamination was present. Furthermore, the morphology of colonies observed was similar to the reported for *Y. lipolytica* cultures in the literature (ABDEL-MAWGOUD et al., 2018; LUPISH et al., 2022). Indeed, colonies had the appearance of a homogeneous and voluminous white mass with well-defined borders after 48 h growth (Appendices Figure 3 A to D). After being stored at 4º C for three days, the colonies show wrinkled borders with an apparently bigger agglomerate of cells (Appendices Figure 3 E & F).

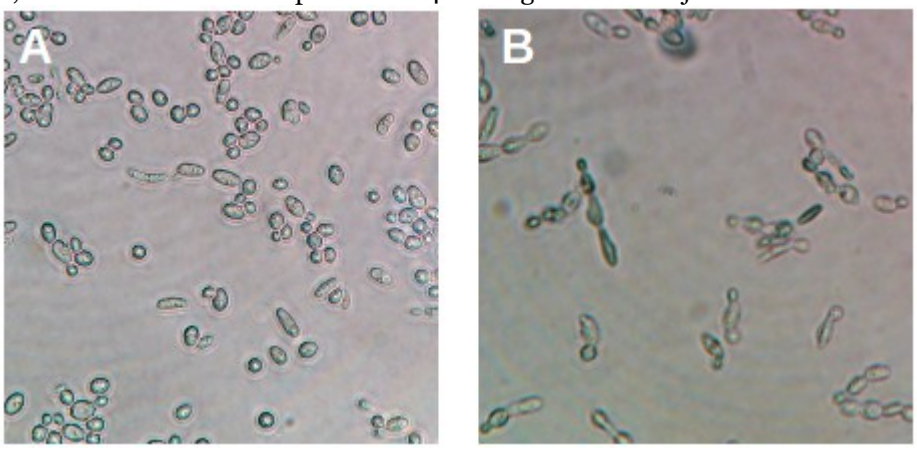
5.1.2 Growth in defined media

Complex media, such as YPD, are known to be composed of a plethora of compounds. This, on one hand, allows optimal growth of a range of different microorganisms. On the other hand, its complex composition hinders the precise quantification of kinetic parameters of microbial growth. For example, *Y. lipolytica* is known to use unidentified carbon and

energy sources in yeast extract present in YPD to grow solely in this compound (OLIVEIRA, 2014). The original paper describing the engineering of the ST9202 strain for limonene production used YPD medium. However, to obtain quantitative physiological parameters on biomass and product formation and on carbon source consumption, the use of a defined medium for *Y. lipolytica* growth is required. This was a specific objective of this work that, although not completely relevant to the development of a genetically-modified strain for geraniol biosynthesis, is related to the establishment of reproducible protocols for growth in the BELa-USP. Therefore, it was important for the reasearch and for educational purposes. As described in Section 3.1.1 (page 46), the MDM, from (MATTHÄUS et al., 2014), and the BDM, from (WOLF, 1996), are the two defined media used for *Y. lipolytica* growth in the literature that were tested in this work.

For MDM, sterilization of the trace elements solution by autoclave generated a yellow precipitate in the solution. Attempts to adjust the solution pH also resulted in precipitation, notwithstanding alteration of reagent addition order to the solution. Therefore, a filtered trace elements solution was utilized. Test of the MDM defined medium was conducted as described in Section 3.1.4 (page 49), in flasks without baffles. Data on final OD_{600} and cell morphology was acquired 40 h after inoculum. While in YPD the ST9202 strain reached an OD of approximately 60, the growth in MDM was limited to approximately 10. As to cell morphology, it was observed by light microscopy that cells grown in MDM, differently from those grown in YPD, were more elongated and prone to form aggregates, as seen in Figure 8. Similarly, compared to growth in YPD, the ST6512 strain in MDM also showed smaller final OD and pseudo-hyphal morphology (results not shown).

Figure 8: ST9202 cells grown in YPD and MDM, at 30º C and 200 rpm for 40 h, observed in a microscope under a 40 x magnification objective lens.

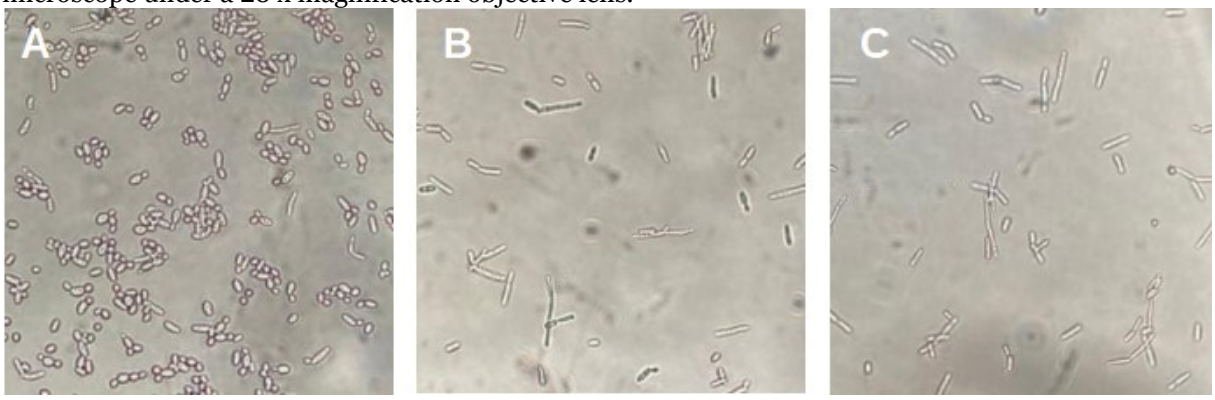


(A) Non-stained culture in YPD complex medium. (B) Non-stained culture in MDM defined medium. Source: Author.

In the original work that reported the MDM composition, neither the final OD_{600} nor the cell morphology reported were noticeably different between YPD and MDM. Moreover, no information regarding specific methods for MDM preparation is given (MATTHÄUS et al., 2014). In the literature, pseudo-hyphae formation is related to nutrient depletion (ABDEL-MAWGOUD et al., 2018; LUPISH et al., 2022). Therefore, the observed growth in MDM could have been an early stage of filament formation, although that was not seen in the original MDM work and would not explain the low OD_{600}. Nevertheless, the hypothesis that MDM lacked an unknown essential nutrient for *Y. lipolytica* growth was formed.

Further research into the literature revealed the BDM medium, which has $CoCl_2$ in its composition, while MDM does not. To test the hypothesis that impaired growth in MDM was due to lack of Co^{2+} ions, the BDM medium was also evaluated. However, preparation of the trace elements solution of this medium was also troublesome. The issue of precipitate formation was observed even before sterilization. The addition of the $FeCl_3 \cdot 6H_2O$ salt destabilized the solution regardless of order of addition. After some unfruitful attempts, the method for preparation of the Verduyn medium (VERDUYN et al., 1992) trace elements solution was adapted to the BDM: 15 g L^{-1} of EDTA was added to deionized water, followed by the other reagents in the specific order they appear in Section 3.1.1 (page 46). Two tests were run in flasks without baffles: one with a filtered trace elements solution and another with a trace elements solution sterilized by autoclave. As results, growth in BDM after 30 h of incubation, regardless of the trace elements solution used, yielded an OD_{600} of approximately 2 versus the value of 30 obtained in YPD. Moreover, the cell morphology was also different between the two media (Figure 9). Once more, compared to growth in YPD, the ST6512 strain in BDM also showed smaller final OD and pseudo-hyphal morphology (results not shown).

Figure 9: ST9202 cells grown in YPD and BDM, at 30º C and 200 rpm for 30 h, observed in a microscope under a 20 x magnification objective lens.



(A) Non-stained culture in YPD complex medium. (B) Non-stained culture in BDM with filtered trace elements solution. (C) Non-stained culture in BDM with autoclaved trace elements solution. Source: Author.

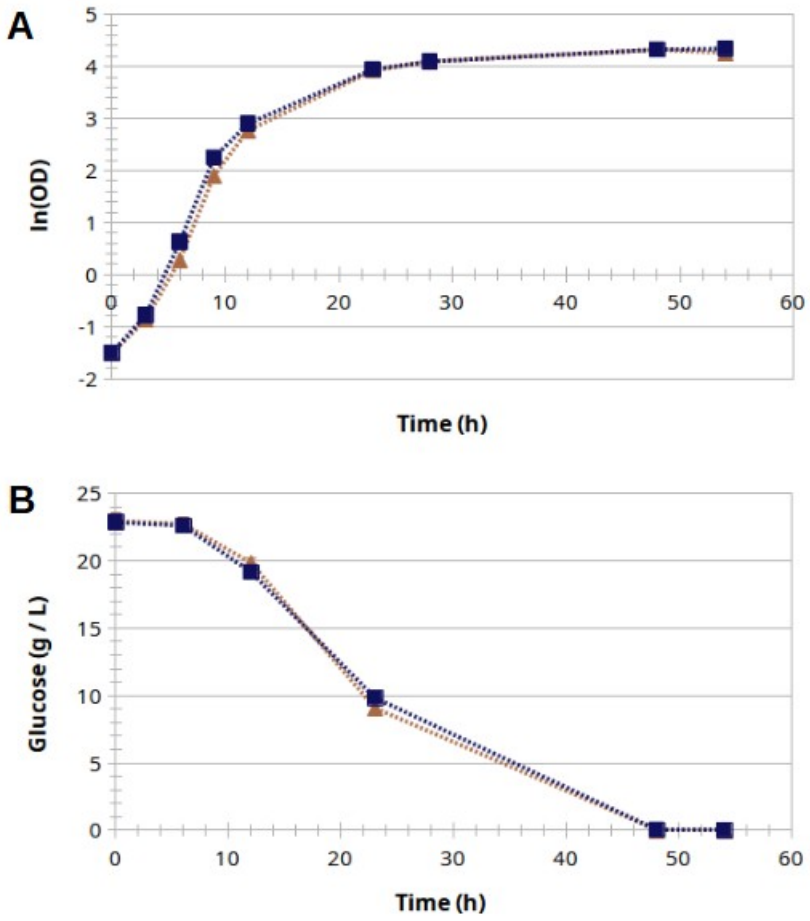
Growth of *Y. lipolytica* in defined media is described in other works (BACCIOTTI, 2016; OLIVEIRA, 2014). In these, BDM was observed to be the best medium regarding μ_{Max} among other defined media, although final OD_{600} was at most half of the achieved in YPD cultures. Atypical cell morphology was not reported (BACCIOTTI, 2016). In the literature, it is reported that *Y. lipolytica* cells with pseudo-hyphal morphology present downregulation of the ATP-citrate lyase gene, among others. This results in impaired acetyl-CoA formation and, therefore, in limited carbon flux to the fatty acid biosynthesis and mevalonate pathways (CELIŃSKA, 2022). Taken together, this points out that the atypical morphology is a visible symptom of a cell condition which is non-optimal to geraniol biosynthesis, as acetyl-CoA availability is decreased. For this reason, growth of the recombinant *Y. lipolytica* proposed in this work should avoid the stressful culture conditions related to pseudo-hyphal morphology incidence.

It has been reviewed that oxygen deprivation is associated with the pseudo-hyphal cell morphology (SOONG et al., 2019). Thus, lack of oxygen due to the use of flasks without baffles was hypothesized to be the limiting factor which resulted in the impaired growth and atypical morphology observed. However, lack or low availability of other nutrient in the defined media tested might also be the reason for the pseudo-hyphal morphology, and further experiments would be necessary to investigate this. Nevertheless, the BDM was chosen as the defined medium to be further used in this work, as it has been employed in previous works (AZAMBUJA, 2016; BACCIOTTI, 2016; OLIVEIRA, 2014) and presents Co ions in its composition.

5.1.3 Growth kinetics

Growth of *Y. lipolytica* in 200 μL of culture volume, in microplates, was atempted. However, due to low oxygen transfer to the culture media in the conditions applied, the method resulted in severely impaired growth. No kinetic parameters could be estimated from the data acquired. Therefore, assessment of *Y. lipolytica* growth physiology was initially performed in experiment "A". For this set of experiments, strains ST6512 and ST9202 were cultured in triplicate in shake-flasks without baffles, in the complex YPD medium, at 30º C and 200 rpm, as described in Section 4.1.3 (page 49). The ST6512 strain was used due to unavailability of the W29 strain at the time of the experiments. Results are shown in Figure 10 and in Table 9.

Figure 10: Growth curves of ST6512 and ST9202 strains obtained with data from growth kinetics experiment "A" in YPD medium at 30º C and 200 rpm and common flasks.



(—■—) ST6512 strain (the Cas9 expressing strain used for construction of the ST9202 strain). (—▲—) ST9202 strain (the monoterpene production chassis strain). (A) Microbial growth curve (in $\ln(OD_{600})$) versus time (in h). Deviation from means is not shown. Growth curve with OD_{600} and standard deviation of measurements are shown in Appendices Figure 4. (B) Curve of glucose consumption (in g L^{-1}) versus time (in h). Error bars represent the standard deviation (experiments made in triplicate). Source: Author.

Table 9: Specific growth rates of *Y. lipolytica* strains obtained in YPD medium.

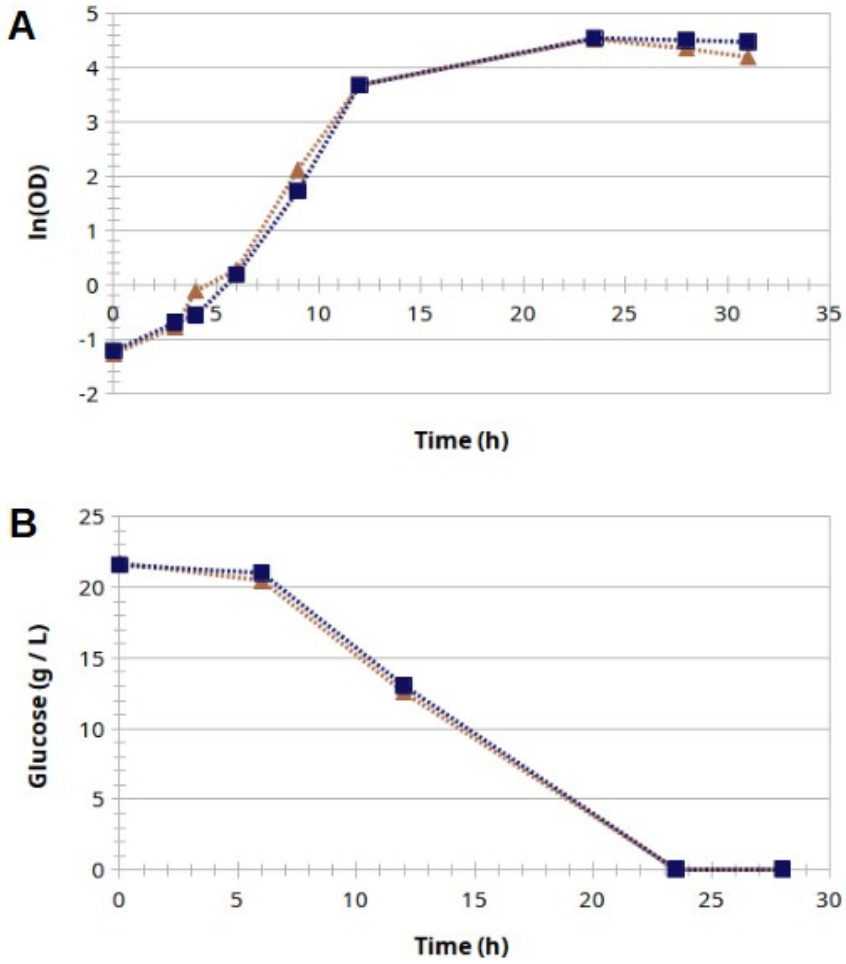
Experiment	Strain	μ_{Max} (h^{-1}) ± Standard deviation (n = 3)	Period of exponential growth	Statistically different?
A	ST6512	0.39 ± 0.01	0 to 12 h	Yes. p = 0.048, α = 0.05
	ST9202	0.38 ± 0.00		
B	ST6512	0.50 ± 0.03	3 to 12 h	No. p = 0.83, α = 0.05
	ST9202	0.49 ± 0.06		

Linear regression models are shown in Appendices Figures 5 and 7.

Glucose quantification via HPLC revealed carbon source depletion concomitant with the stationary growth phase, observed at 48 h onwards (Figure 10). The data from experiment "A" point to a statistical difference ($p = 0.048$, $\alpha = 0.05$) between the calculated μ_{Max} values of strains ST6512 and ST9202 (Table 9) (calculation of μ_{Max} and statistical tests conducted as described in section 4.1.3, page 49). Although not reported in the original work in which the ST9202 strain was developed (ARNESEN et al., 2020), it could be possible that the seven genetic modifications it bears would impose a metabolic burden upon its growth. However, despite the statistical difference, the μ_{Max} average values are close to each other, and the growth profiles of both strains are very similar. Thus, it is not plausible to assume that the statistical difference is a direct signal of distinct cell physiology phenomena among the strains. As seen by the results from experiment "B" discussed below, the statistical difference found is better explained by OD_{600} measurement variations in biological replicates.

In experiment "B", growth was also conducted in YPD medium, but in baffled flasks to increase oxygen availability (transfer rate). Temperature and rotation were also modified to 28º C and 250 rpm. Based on similarly conducted experiments found in the literature (BACCIOTTI, 2016), it was hypothesized that these measures would increase the observed μ_{Max} and decrease the experiment total time. Results are shown in Figure 11 and in Table 9.

Figure 11: Growth curves of ST6512 and ST9202 strains obtained with data from growth kinetics experiment "B" in YPD medium at 28º C and 250 rpm and baffled flasks.



(—■—) ST6512 strain (the Cas9 expressing strain used for construction of the ST9202 strain). (—▲—) ST9202 strain (the monoterpene production chassis). (A) Microbial growth curve (in ln(OD_{600})) versus time (in h). Deviation from means is not shown. Growth curve with OD_{600} and standard deviation of measurements are shown in Appendices Figure 6. (B) Curve of glucose consumption (in g L^{-1}) versus time (in h). Error bars represent the standard deviation (experiments made in triplicate). Source: Author.

As observed (Figure 11), maximum OD and depletion of glucose were observed earlier than cultures of experiment "A", at 24 h, instead of 48 h. Also, the altered conditions of lower temperature and increased oxygen transfer using baffled flasks and higher agitation resulted in μ_{Max} values almost 30 % higher in experiment "B" than in experiment "A" (Table 9). However, differently from the previous assay, no statistical difference was observed between the μ_{Max} of the strains tested. Accordingly, the quality of the data in experiment "B" is questionable, with measures varying greatly between replicates (as seen in Appendices Figure

6). Therefore, once again the statistical difference, or lack thereof, may be due to the dispersion of measurements, not due to a biological process. It would be necessary to retry the experiment and verify the data obtained. Nevertheless, the calculated μ_{Max} in experiment "B" was also found in the literature for the W29 strain in experiments with similar settings, which range from 0.49 h^{-1} to 0.51 h^{-1} (BACCIOTTI, 2016; OLIVEIRA, 2014).

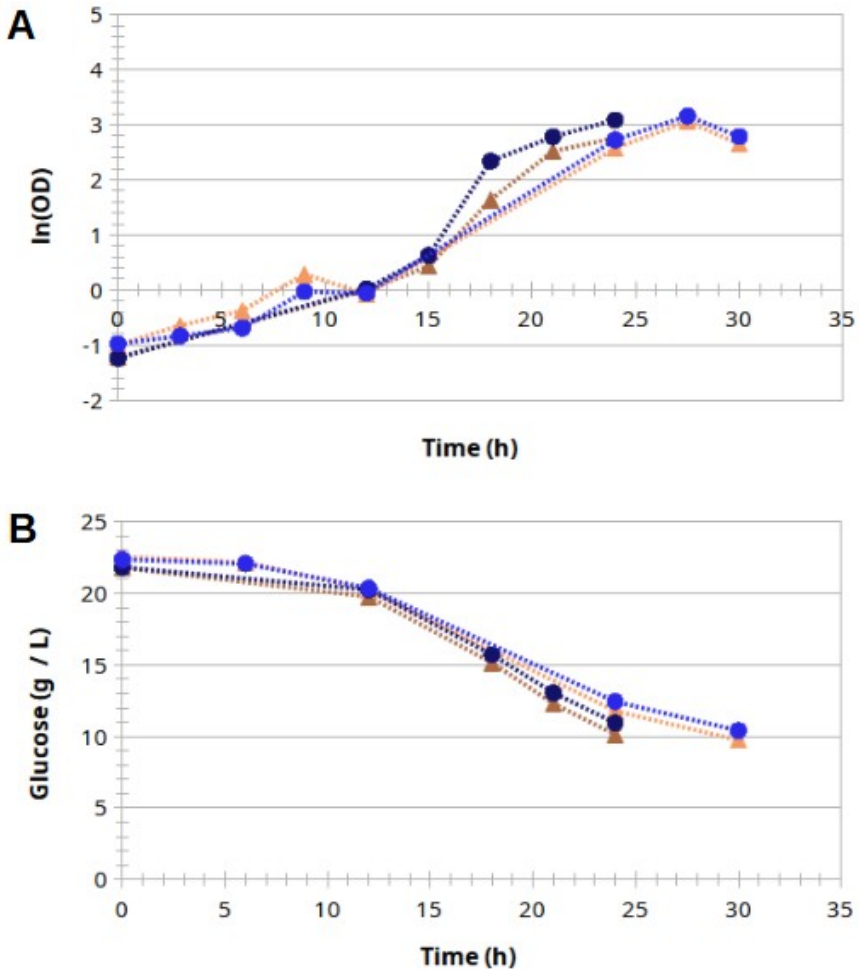
Long deceleration phases after the exponential growth phases, when compared to typical *S. cerevisiae* growth in complex medium, were also observed, especially in experiment "A". In Figure 10, the growth profile of cells in experiment "A" show the deceleration phase from 24 to 28 h. Although the sampling intervals conducted herein were not efficient at capturing this growth period (which might be also present in the growth conditions of experiment "B"), its presence might be related to a mild oxygen limitation. An 8h deceleration phase was also observed in similar growth conditions for the W29 strain in (BACCIOTTI, 2016). In that work, low oxygen transfer to the culture medium is pointed out as the main hypothesis to the end of the exponential growth phase without complete consumption of glucose. Therefore, the oxygen transfer rate to the medium at the conditions tested might be insufficient to provide enough oxygen as cells grow and especially towards the end of the cultivation, when the cell density is high. Nevertheless, no other signs of oxygen limitation (e.g. pseudo-hyphal morphology) were observed in experiments "A" and "B".

No citrate and isocitrate were quantified in traceable amounts in any of the experiments conducted. While these are the main by-products obtained during *Y. lipolytica* culture, citrate and isocitrate are usually produced in smaller amounts when grown in glucose, compared to glycerol or fatty acids as carbon sources (PAPANIKOLAOU et al., 2003). Another hypothesis is that the high C/N ratio observed in cultures during citrate production phases was not achieved due to low carbon source initial titer (MAKRI; FAKAS; AGGELIS, 2010; PAPANIKOLAOU et al., 2003). Indeed, the growth seemed to be limited at the stationary phase by carbon source availability, as seen in Figure 10 B and 11 B.

Finally, experiment "CI" was devised to observe the μ_{Max} and the final OD_{600} of the *Y. lipolytica* strains grown in BDM-D in baffled flasks. It would also indicate whether increased oxygen transfer would avoid the pseudo-hyphal morphology. However, it was observed that culture in this medium displayed greatly reduced growth rate as compared to complex media, and thus it seemed that exponential growth phase was not achieved during the time span in which samples were analyzed in experiment "CI". Thus, experiment "CII", with the same conditions as "CI" and observations from 12 to 24 h of growth, was also conducted. The W29

strain was used instead of the ST6512 strain so as to compare the ST9202 chassis strain with the unmodified wild type strain. Results are shown in Figure 12.

Figure 12: Growth curves of W29 and ST9202 strains obtained with data from growth kinetics experiment "C" in BDM-D medium at 28º C and 250 rpm and baffled flasks.

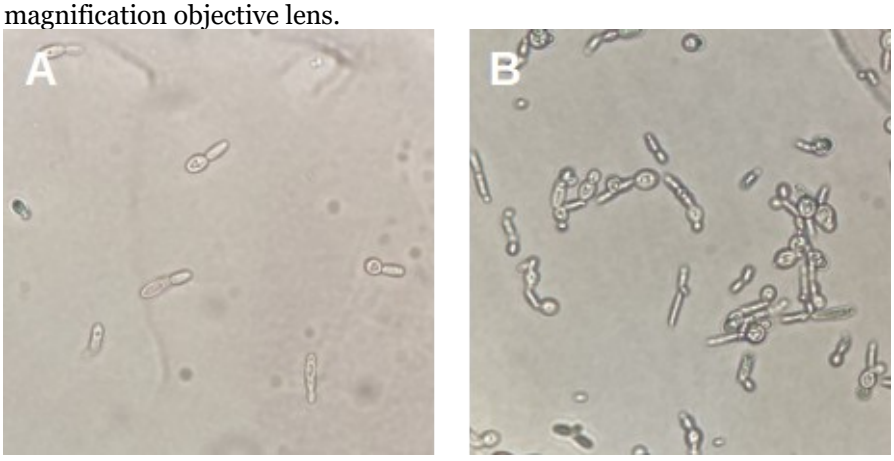


(—●—) W29 strain (the wild type strain) in "CI". (—●—) W29 strain (the wild type strain) in "CII". (—▲—) ST9202 strain (the monoterpene production chassis) in "CI". (—▲—) ST9202 strain (the monoterpene production chassis) in "CII". (A) Microbial growth curve (in $\ln(OD_{600})$) versus time (in h). Deviation from means is not shown. Growth curve with OD_{600} and standard deviation of measurements are shown in Appendices Figure 8. (B) Curve of glucose consumption (in g L^{-1}) versus time (in h). Error bars represent the standard deviation (experiments made in triplicate). Source: Author.

Although a reduction in the number of cells was observed at 30 h, glucose was not depleted from the medium. This might be due to an error on the OD_{600} measurements at 30 h. Another possibility is that growth was limited by a nutrient different from the carbon source. In (BACCIOTTI, 2016), growth of the W29 strain in defined medium also reached the

stationary phase without the complete consumption of the carbon source used, glucose. The main hypothesis given was that the culture suffered from slight oxygen deprivation. As for the experiments "C" in this work, it would be plausible to conduct another experiment with OD_{600} measurements for longer periods. In regard to cell morphology, the cells presented only a slight improvement to yeast-like form throughout the experiment (Figure 13, observations after 6 and 30 h of incubation), compared to the initial tests (Figure 9, observation after 30 h of incubation). Cells from the W29 strain showed similar morphology. Thus, oxygen may be relevant to the cell morphology but it is not the only factor that influences it during growth in BDM. This is further evidenced by experiment "A", in which cells grown in YPD displayed yeast morphology despite the use of flasks without baffles. It is possible that the BDM lacks an essential nutrient to this oleaginous yeast growth. Compared to the Verduyn defined medium that contains seven vitamins for *S. cerevisiae* growth, BDM only uses thiamine (VERDUYN et al., 1992; WOLF, 1996), but literature reports that *Y. lipolytica* biosynthesizes all but this one vitamin among the necessary for its metabolism (ABDEL-MAWGOUD et al., 2018; OLIVEIRA, 2014). Less certain are the inorganic (trace elements) requirements for *Y. lipolytica* growth, as a relatively recent review on *Y. lipolytica* stated that there has not been a thorough investigation of the exact trace elements required for this yeast species (ABDEL-MAWGOUD et al., 2018).

Figure 13: ST9202 cells grown in BDM at 28º C and 250 rpm in baffled flasks (Experiment "CI") observed in a microscope under a 40 x magnification objective lens.



(A) Non-stained culture after 6 h of incubation. (B) Non-stained culture after 30 h of incubation. Source: Author.

The growth in BDM-D reached less than half of the final OD_{600} achieved in YPD (experiment "B", Figure 11). Moreover, it was observed that, accounting for both experiments "CI" and "CII", exponential growth of cells happened between 0 and 24 h. The μ_{Max} calculated

for the growth in BDM-D (Table 10) were approximately 35 % of the μ_{Max} in YPD. However, as data from two different experiments were used, this is only an estimate and thus, no statistical test was conducted. The literature reports higher μ_{Max} values for the W29 strain grown in BDM, ranging from 0.27 (with 2.5 g L^{-1} glucose as carbon source) to 0.30 h^{-1} (with 20 g L^{-1} glucose), but still below the ones obtained in YPD (BACCIOTTI, 2016; OLIVEIRA, 2014). The prolonged exponential growth phase might be due to a slow rate of uptake of nutrients from the medium, which is possibly related to the problems in media element trace solution components stabilization. While it could be expected that growth in the defined media (which requires the *de novo* synthesis of vitamins) would not be as robust in as in the rich complex media (which is composed with ready-to-use vitamins), this could not justify the presence of cells with atypical morphology. Thus, this decreased μ_{Max} compared to that in YPD, especially when allied to the pseudo-hyphal morphology observed, points to the possibility that BDM lacks one or more unknown essential nutrients and to the necessity of medium optimization studies for *Y. lipolytica*.

Table 10: Specific growth rates of *Y. lipolytica* strains obtained in BDM.

Experiments	Strain	μ_{Max} (h^{-1}) ± Standard deviation (n = 3)	Period of exponential growth
CI and CII	W29	0.19 ± 0.00	0 to 24 h
	ST9202	0.16 ± 0.01	

Linear regression models shown in Appendices Figure 9.

It is worth mentioning that the replicates of cultures both in complex media and in defined media presented an unexpected behavior, with higher standard deviations on each observation, especially in the experiments that used baffled flasks (as seen in the Appendices Figures 4, 6 and 8. In some cases, a replicate that showed higher growth than the other in one observation would then display smaller in the next observation, or vice-versa. In these cases, measuring the OD_{600} again often did not point to mistakes in the measurement practice. Nevertheless, this possible irregularity of biomass growth between measurements was not observed in the profile of glucose consumption measured by HPLC data. This would indicate that the OD measurements were not a precise method to assess biomass formation by *Y. lipolytica*. A hypothesis for these scattered measurements is the presence of stains in the cuvettes used to measure the OD_{600}, possibly caused by oil produced by the yeast. Moreover, dilution of samples measured in the spectrophotometer may also be a source of error, as seen in (AZAMBUJA, 2016), although the equipment used requires dilution to keep measurements

inside the reliable range. Further experiments would be needed, with use of another method for biomass estimation, such as dry mass, not used in this work due to the culture volumes employed and the restriction to sample, at most, 10 % of these volumes.

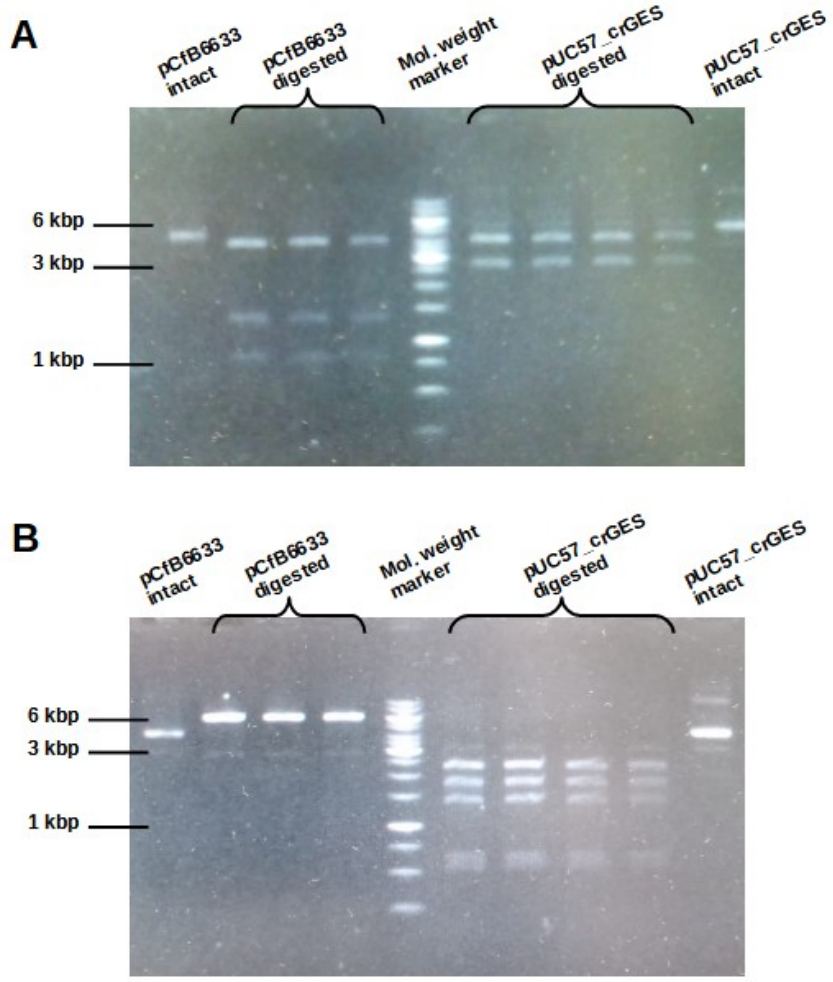
Finally, it is worth mentioning that these experiments conducted here were conducted as means to get a first understanding of *Y. lipolytica* growth. Further experiments are required to fully investigate the growth kinetics of this yeast species, and this work offers a starting point. From here, new investigation may analyze the effect of geraniol in *Y. lipolytica* growth, for example. However, as the work progressed and the plasmid parts became available in the laboratory, it was possible to direct our efforts to molecular biology and the metabolic engineering.

5.2 Replication of plasmids

The plasmids attained and manipulated are shown in Section 4.2.2 (page 53, Table 7), and their maps and sequences are provided in the Appendices Figures 10 to 13. Replication of plasmids in *E. coli* strains is a method to preserve the recombinant DNA in a safe and lasting way. The bacterial cells were transformed with the plasmid DNA as described in section 4.2.2 (page 53), and the resulting colonies were frozen in glycerol stocks. To ascertain that the DNA received were the correct plasmids reported, a miniprep extraction was conducted, followed by restriction enzyme assays. Then, in agarose gels, the band patterns obtained were compared to the expected plasmids restriction maps.

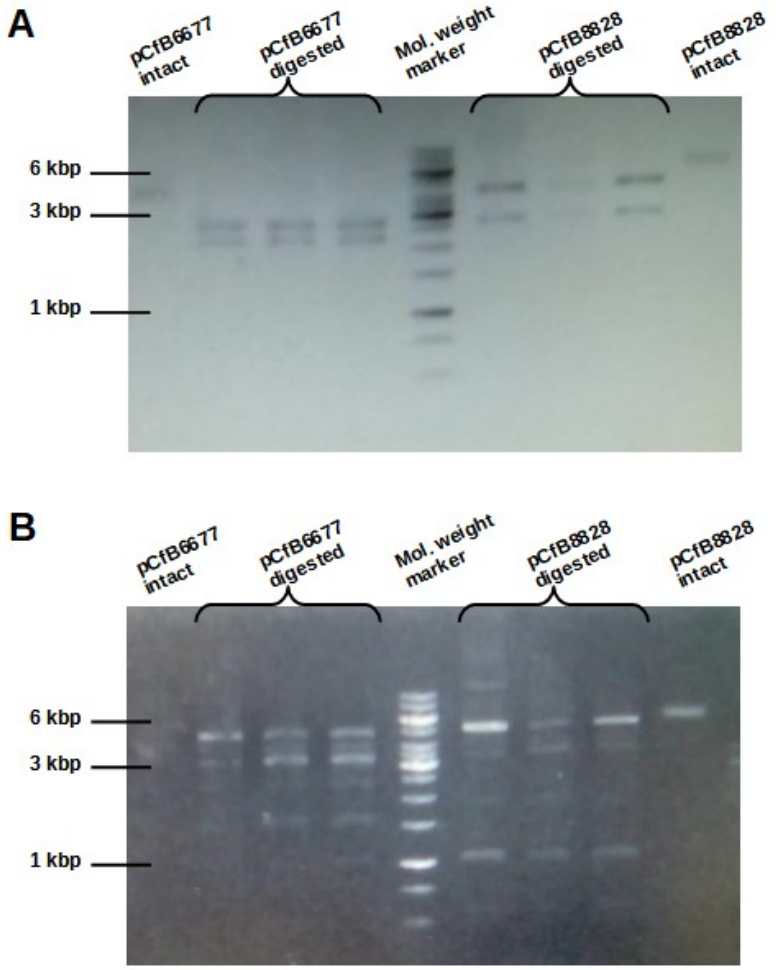
Early attempts of *E. coli* transformation and plasmid preparations proceeded without complications: colonies resistant to ampicillin were obtained, no colonies were observed in negative controls and the miniprep yielded pure DNA solutions, although with low concentrations (results not shown). However, the observed gel band patterns after endonuclease treatment were different than expected (results not shown). While some bands of expected sizes were visible, other bands, of unknown origin, were also observable. After new attempts ruled out the possibility of over digestion and reached the same results, it was raised the hypothesis that the bacteria strain used in the transformation already carried an unknown plasmid with another selection mark. This would explain why the ampicillin selection seemed to work, with no growth of negative controls but transformant colony formation. By repeating the transformations with a new *E. coli* strain, the obtained band patterns after plasmid extraction and treatment with restriction endonucleases were obtained as expected (Figures 14, 15 & 16).

Figure 14: Images of agarose gels with endonuclease reactions of the pCfB6633 and pUC57_crGES plasmids.



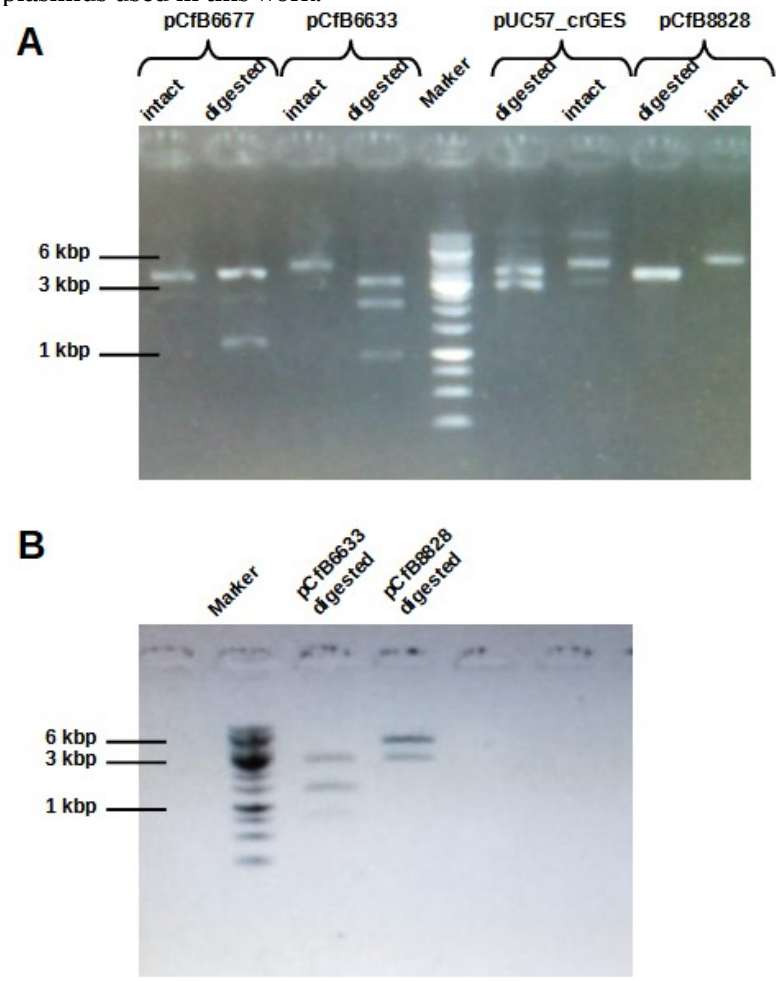
Molecular weight marker: GeneRuler 1 kb DNA Ladder, as shown in Appendices Figure 1 A. Intact plasmids are shown as digestion negative controls. (A) Treatment of plasmids with the *Hin*dIII enzyme. pCfB6633 expected band pattern: 4152 + 1354 + 819 bp. pUC57_crGES expected band pattern: 4091 + 2710 bp. (B) Treatment of plasmids with the *Pvu*II enzyme. pCfB6633 expected band pattern: 6325 bp. pUC57_crGES expected band pattern: 2364 + 1834 + 1408 + 636 + 559 bp. Source: Author.

Figure 15: Images of agarose gels with endonuclease reactions of the pCfB6677 and pCfB8828 plasmids.



Molecular weight marker: GeneRuler 1 kb DNA Ladder, as shown in Appendices Figure 1 A. Intact plasmids are shown as digestion negative controls. (A) Treatment of plasmids with the *Not*I enzyme. pCfB6677 expected band pattern: 2824 + 2301 bp. pCfB8828 expected band pattern: 4625 + 2824 bp. (B) Treatment of plasmids with the *Pvu*II enzyme. pCfB6677 expected band pattern: 5125 bp. pCfB8828 expected band pattern: 5315 + 1048 + 582 + 504 bp. Source: Author.

Figure 16: Images of agarose gels with endonuclease reactions of the plasmids used in this work.



Molecular weight marker: GeneRuler 1 kb DNA Ladder, as shown in Appendices Figure 1 A. Intact plasmids are shown as digestion negative controls. (A) Treatment of plasmids with the *Eco*RV enzyme. pCfB6677 expected band pattern: 3894 + 1231 bp. pCfB6633 expected band pattern: 4636 + 962 + 727 bp. pUC57_crGES expected band pattern: 3886 + 2915 bp. pCfB8828 expected band pattern: 3894 + 3555 bp. (B) Treatment of plasmids with the *Xba*I enzyme. pCfB6633 expected band pattern: 2742 + 1415 + 1396 + 772 bp. pCfB8828 expected band pattern: 4707 + 2742 bp. Source: Author.

In Figure 14A, the pUC57_crGES plasmid was not completely cleaved, as seen by the bands identical to the undigested plasmid above the expected bands. The same was observed for treatment of this plasmid with the *Eco*RV restriction enzyme (Figure 16A). Nonetheless, the expected bands are visible. Treatment with the *Pvu*II restriction endonuclease (Figure 15B) displayed breaks at non-specific sites and bands different from those expected, attributed to excessive incubation time during digestion. However, reactions with two other enzymes, *Eco*RV and *Xba*I (Figure 16), resulted in the expected patterns. Finally, the pCfB6633 treatment with *Eco*RV (Figure 16A) yielded bands of approximately 3.1, 2.1 and 0.9 kbp,

which is different from the expected pattern. However, the sum of these lengths is close to the plasmid total length of 6.3 kbp. Furthermore, considering that the 0.9 kbp fragment seen was indeed expected and that treatment with other three endonucleases (*Hin*dIII, *Pvu*II and *Xba*I) yielded the anticipated results, the possibility that the plasmid was not the pCfB6633 was discarded. It was hypothesized that the position of the third *Eco*RV restriction site was misplaced in the restriction map used. Thus, the plasmids obtained were confirmed as pCfB6633, pCfB6677, pCfB8828 and pUC57_crGES. With the correct replication and confirmation of the plasmids, it is now possible to move to the steps of transformation of *Y. lipolytica* in the coming months.

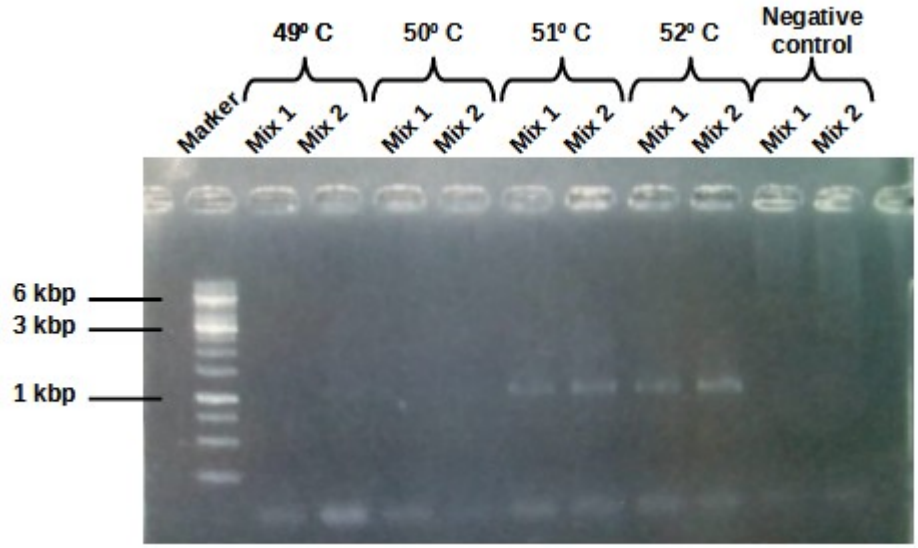
5.3 Geraniol synthase gene integration

5.3.1 Transformant colonies screening protocol

Aiming to establish a reproducible PCR protocol to be used for screening of *Y. lipolytica* transformants, different tests were performed. For this, the PR-14442 and PR-14398 primers and the ST9202 strain were employed as described in Section 4.2 (page 52). Thus, the reactions conducted should yield the 1066 bp amplicon fragment, which includes the IntE_1 region for genomic integration of constructs (HOLKENBRINK et al., 2018).

The first protocol attempted was based on colony PCR, as recommended in (HOLKENBRINK et al., 2018). Conditions investigated (results not shown) included three different DNA taq polymerase enzymes among the available in the laboratory, four annealing temperatures and six cell lysis procedures. These procedures included suspension in water or NaOH 20mM, preceded or followed by heating in a microwave or in a water bath. The use of Platinum *Taq* Polymerase, 51º C and 52º C annealing temperature and cell lysis with NaOH and heating (suspension of a fraction of a colony in 20 µL of NaOH 20 mM followed by incubation at 70º C for 15 min) resulted in correct amplification, with a visible band in the verification agarose gel at the position of the 1 kbp band of the DNA ladder used (Figure 17).

Figure 17: Image of agarose gel with colony PCR reactions of the IntE_1 locus from the *Y. lipolytica* ST9202 strain genomic DNA acquired with the NaOH and heating cell lysis protocol.



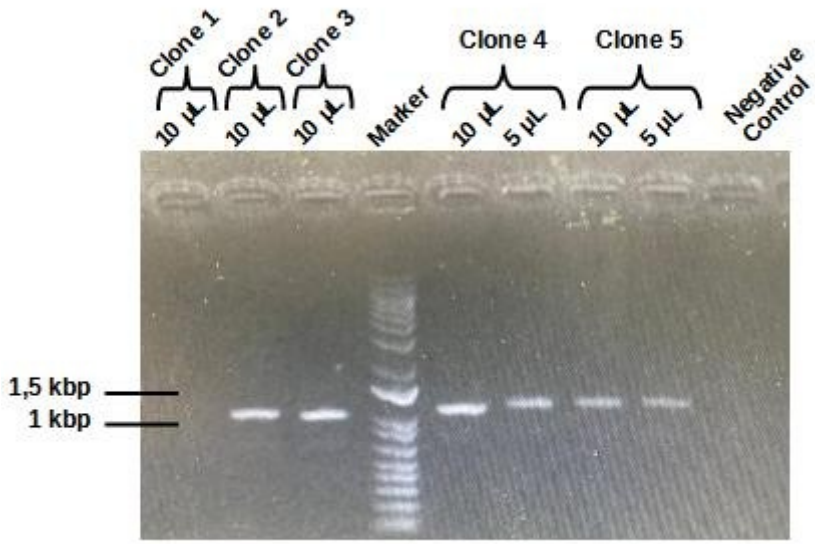
Expected DNA fragment size: 1066 bp. "Mix 1" and "Mix 2" refer to different Platinum *Taq* Polymerase reagents tested. Temperatures from 49º C to 52º C were the annealing temperatures tested. No cell lysis solution was added in negative controls, which were incubated at an annealing temperature of 51º C. Molecular weight marker: GeneRuler 1 kb DNA Ladder, as shown in Appendices Figure 1 A. Source: Author.

After this successful set of conditions was found, details in the technique were evaluated to further increase the success rate of the colony PCR method. Since the enzyme, reagent concentration in the reaction and annealing temperature were easy to control and reproduce, false negative results were probably related to the variation introduced by the cell lysis method. Quantity of cells collected, "age" of the colonies used and whether the cell lysate solution is centrifuged or not prior to addition to the PCR reaction are conditions that may alter the availability of template DNA for the PCR reaction. However, the tests in which these conditions were altered (results not shown) were not successful at increasing the proportion of correct results (approximately 20 %) over false negatives. This means that, out of five colonies, only one is lysed in a way that allows amplification of the region of interest when the NaOH and heating lysis protocol is used. This was probably due to the high quantity of cell debris and metabolites present in the reactions, as PCR is a method known to be highly sensitive to contaminants.

Thus, genomic DNA extraction with a commercial kit was attempted as described in section 4.2.1 (page 52), with the hypothesis that it would increase the proportion of correct results in the PCR. After ST9202 genomic DNA extraction, solutions obtained were used for DNA quantification. Extraction yields measured were extremely low, below the equipment

range of detection. Nonetheless, 20 µL PCR conducted with 5 or 10 µL of these solutions displayed amplification at the expected size, as seen in Figure 18. Annealing temperature of primers and reagents used were the same of those used in the colony PCR approach. The success rate over false negatives was 80 %. Therefore, extraction of genomic DNA followed by PCR was established as the protocol for screening of genetically modified strains after transformation.

Figure 18: Image of agarose gel with PCR reactions of the IntE_1 locus from the genomic DNA of *Y. lipolytica* ST9202 strain.



Expected DNA fragment size: 1066 bp. Position of wells in gel was irregular. "Clones" refer to different genomic DNA extraction solutions used in PCR. Volumes of 5 or 10 µL of genomic DNA solutions were used in each PCR reaction. Water instead of genomic DNA solution was added in negative control reaction. Molecular weight marker: 1 kb DNA Ladder, as shown in Appendices Figure 1 B. Source: Author.

5.3.2 *Yarrowia lipolytica* transformation

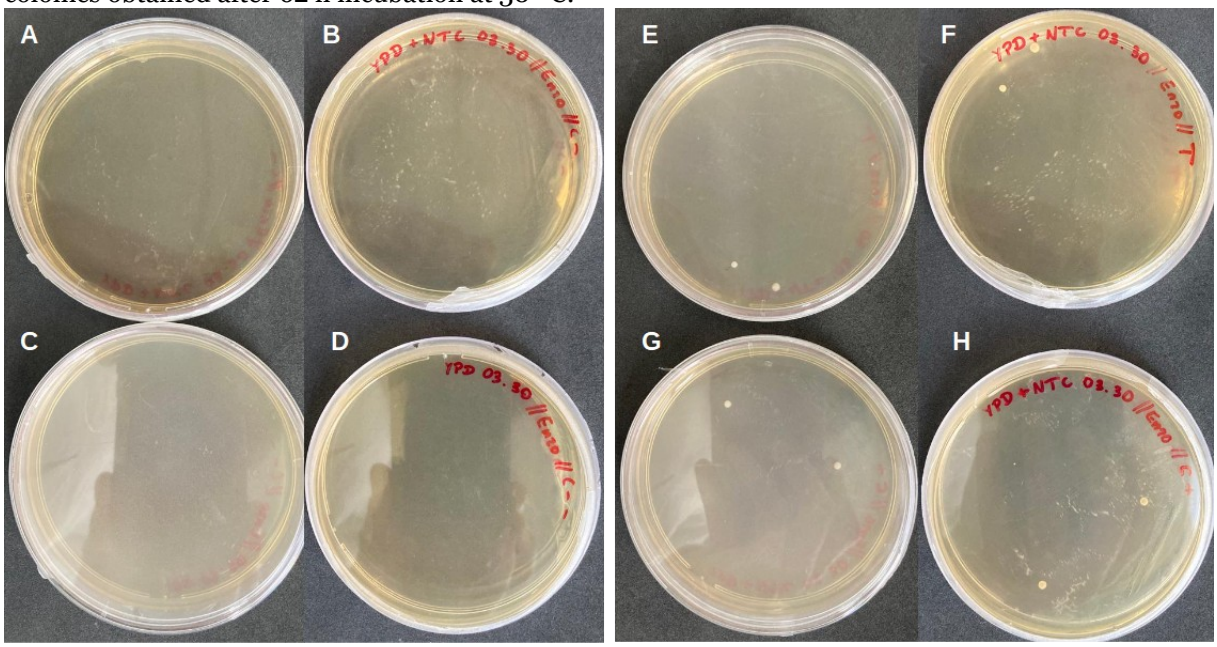
Y. lipolytica ST9202 strain was used for transformation with the *crGES* expression cassette. The *GES* gene from *C. roseus* was chosen among different plant sources due to its use in various works (BROWN et al., 2015; MISHRA et al., 2020; WANG et al., 2021). Moreover, in (JIANG et al., 2017), *crGES* was the *GES* which resulted in highest geraniol production in *S. cerevisiae* among nine genes tested. Transformation with the *pfLS* expression cassette from the pCfB8828 plasmid was performed as a positive control, given that such transformation was successfully conducted by (ARNESEN et al., 2020). Along with the expression cassette that would act as donor DNA (either *crGES* or *pfLS*), transformation required the pCfB6633 plasmid for sgRNA expression. This sgRNA directs the Cas9-

mediated cleavage of the genomic DNA at the IntE_1 locus, in the phenomena described in Section 2.4.1.1 (page 35). During transformation of cells, the DSB must be repaired by HR, as NHEJ repair is impaired due to deletion the *ku70* locus. The donor DNA (either *crGES* or *pfLS* expression casseettes) acts as template for HE, which results in the integration of the expression cassete in IntE_1. This genomic region was evaluated as an ideal integration *locus* as it does not contain open reading frames, but it is located in the proximity of at least five genes with high expression during exponential growth phase and in conditions of limited nitrogen (HOLKENBRINK et al., 2018). Furthermore, in this work, expression cassette integration at this *locus* was chosen due to its already reported use for *pfLS* cassette integration and expression (ARNESEN et al., 2020). Therefore, the tools for its application, mainly the pCfB6633 plasmid with the sgRNA and the proof of concept of its efficiency, are readily available.

The transformation and integration strategy employed in this work, based in (HOLKENBRINK et al., 2018), is considered marker-free as no selection mark is integrated into the genome. Nevertheless, two types of "selection" of transformants take place: first, cells not transformed with the pCfB6633 plasmid do not grow in the nouseothricin selective media; second, cells with the pCfB6633 plasmid but without the donor DNA do not replicate because its NHEJ repair pathway is compromised. Thus, as the selection marker is present only in the episomal pCfB6633 plasmid, not in the donor DNA, marker-free genetically modified cells are obtained as antifungal resistance is lost when cells are grown in non-selective media.

The transformation protocol consists in rinsing the appropriate number of cells with a transformation mix, as described in section 4.3 (page 55). After 62 h, transformant colonies were observed in plates with selective YPD medium (Figure 19). Four colonies of each transformation reaction (ST9250 and BYa3105 strains) were observed, with no colonies in the negative controls without cells or with cells that did not receive the pCfB6633 plasmid. However, only cells from two of those colonies from each reaction (coincidentally, the bigger colonies) were able to grow after inoculation in selective YPD medium for cell freezing and genomic DNA extraction. These colonies were used to confirm the integration, as described in the next section.

Figure 19: Petri dishes with YPD medium and nourseothricin antifungic of transformant *Y. lipolytica* colonies obtained after 62 h incubation at 30º C.



(A) C-: Frontal vision of plate with nourseothricin selective medium, inoculated with cells transformed without the pCfB6633 plasmid; (B) C-: Back vision of plate with nourseothricin selective medium, inoculated with cells transformed without the pCfB6633 plasmid; (C) C- -: Frontal vision of plate with non-selective medium, kept open during transformation to account for manipulation contamination; (D) C- -: Back vision of plate with non-selective medium, kept open during transformation to account for manipulation contamination; (E) T: Frontal vision of plate with nourseothricin selective medium, inoculated with cells transformed with the pCfB6633 plasmid and the *crGES* donor DNA (treatment, BYa3105, the genetically-modified ST9202 with the *crGES* cassette integrated into its genome); (F) T: Back vision of plate with nourseothricin selective medium, inoculated with cells transformed with the pCfB6633 plasmid and the *crGES* donor DNA (treatment, BYa3105, the genetically-modified ST9202 with the *crGES* cassette integrated into its genome); (G) C+: Frontal vision of plate with nourseothricin selective medium, inoculated with cells transformed with the pCfB6633 plasmid and the *pfLS* donor DNA (positive control, ST9250, the genetically-modified ST9202 with the *pfLS* cassette integrated into its genome); (H) C+: Back vision of plate with nourseothricin selective medium, inoculated with cells transformed with the pCfB6633 plasmid and the *pfLS* donor DNA (positive control, ST9250, the genetically-modified ST9202 with the *pfLS* cassette integrated into its genome).

5.3.3 Confirmation of integration

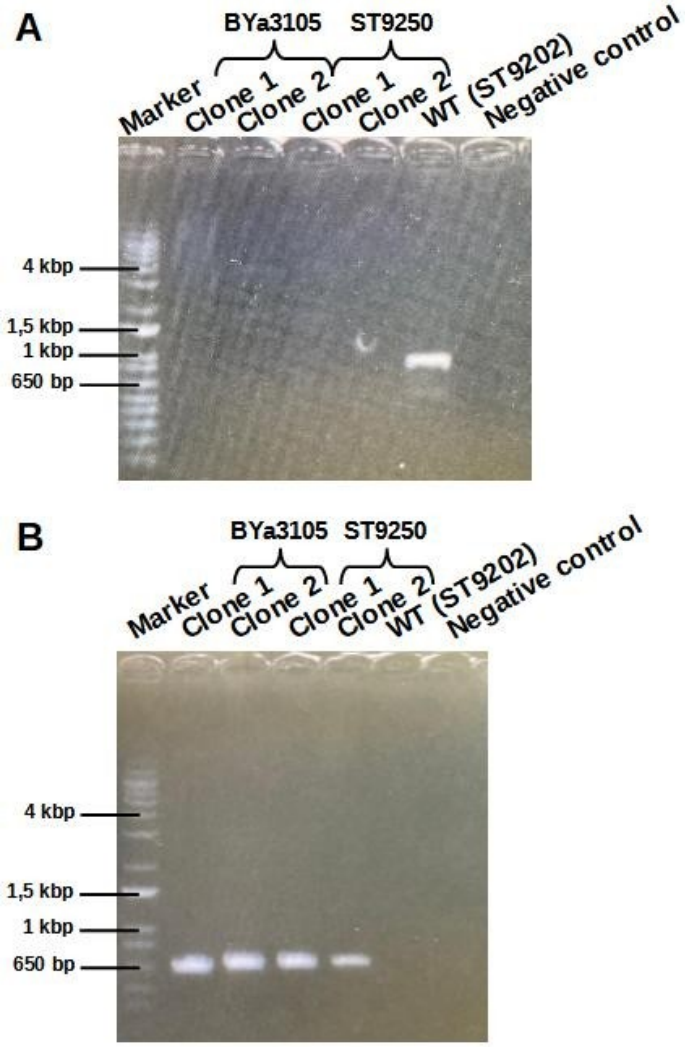
PCR with genomic DNA of the two clones of ST9250 and the two clones of BYa3105 as templates was conducted as described in sections 4.2.1 (page 52) and 5.3.1 (page 74) to verify whether integration of expression cassettes in the IntE_1 locus was successful. Two different primer pairs were used, and the sizes of the expected fragments are described in Table 11.

Table 11: Primer pairs used to verify integration in the IntE_1 locus and the sizes of fragments amplified in the genomic DNA from each strain used.

Primer pair	Amplified region in the IntE_1 locus	Amplicon size (bp)	
PR-14442 & PR-14398	Whole cassette and 54 bp of the genome upstream.	WT genome	1066
		BYa3105	4139
		ST9250	4643
PR-InDFor & PR-InDRev	38 bp of Tlip2 terminator, whole IntE_1 Down sequence and 164 bp downstream.	WT genome	No amplification
		BYa3105	666
		ST9202	677

Results for diagnostic PCR are displayed in Figure 20. Amplification with the PR-InDRev and PR-InDFor pair was observed for both constructions, but not in the negative control (Figure 20 B). However, it was not possible to verify which expression cassette, whether the one with the *crGES* or the *pfLS* gene, integrated in each strain: the 11 bp difference between the amplified fragments could not be resolved in the agarose gel. This hurdle was expected and PCR with the PR-14442 and PR-14398 pair was done to surmount it. However, amplification was not observed for any of the constructions (Figure 20 A). The hypothesis for this is that contaminants in the extracted DNA used as templates and that the long size of the expected fragments impaired proper *Taq* polymerase function. Nevertheless, results with the first primer pair indicated that the integration of donor DNA occurred at the IntE_1 locus.

Figure 20: Images of agarose gels with PCR reactions of the IntE_1 locus from the genomic DNA of *Y. lipolytica* BYa3105, ST9250 and ST9202 for confirmation of transformation.



"Clones" refer to different genomic DNA extraction solutions used in PCR. "WT" refers to the wild type, unmodified IntE_1 region. Negative controls were made with water instead of genomic DNA template. Molecular weight marker: 1 kb DNA Ladder, as shown in Appendices Figure 1 B. (A) PCR with PR-14442 and PR-14398 primer pair. BYa3105 expected amplicon size: 4139 bp; ST9250 expected amplicon size: 4643 bp; ST9202 expected amplicon size: 1066 bp. (B) PCR with PR-InDFor and PR-InDRev primer pair. BYa3105 expected amplicon size: 666 bp; ST9250 expected amplicon size: 677 bp; No amplification for ST9202 was expected. Source: Author.

Therefore, despite the low transformation efficiency observed, the transformation of obtained clones was confirmed. To differentiate the integration of the *crGES* and *pfLS* cassettes through this PCR approach, it would be necessary to design a new set of primers. Meanwhile, geraniol production assays were conducted, as described in the next section.

5.4 Evaluation of geraniol production

Production of geraniol by the *crGES* bearing BYa3105 strain, in YP4D medium with 10 % dodecane added at culture start, was performed as described in section 4.1.4 (page 51). Nourseothricin was not added to media in plates in which cells were grown, to allow cells to lose the pCfB6633 plasmid with the selection mark. As the expression cassettes were integrated into the yeast genome, losing this sgRNA plasmid would not impair monoterpene synthesis in transformed cells. Dodecane used for liquid-liquid extraction of geraniol was not autoclaved due to fire hazard. Although identification of geraniol production would probably be possible directly from the culture media, without any extraction phases, dodecane extraction is relevant not only to avoid losing the highly volatile monoterpene to the atmosphere but also to alleviate the monoterpene toxicity in the culture medium. This was described in Section 2.3.4 (page 30).

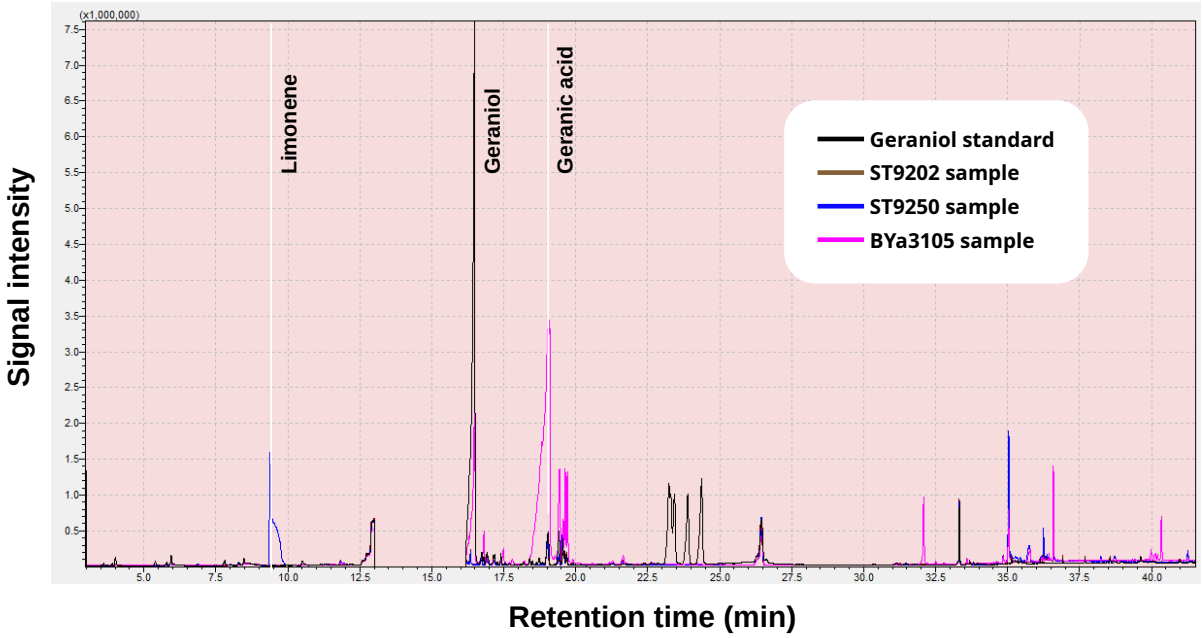
Cultivation of BYa3105, ST9250 and ST9202 strains was performed in complex YP4D (YP with 80 g L^{-1} glucose) medium. Dodecane was added in the medium prior to inoculation, for terpene extraction. The negative control without cell inoculum presented no growth and thus indicated that no contamination was present in the dodecane added. However, some differences were observed among the strains. Firstly, the dodecane overlay of BYa3105, like the one from the negative control, had no color, while the organic solvent harvested from ST9250 and ST9202 cultures displayed a green to yellow color (Appendices Figure 14). This was received as a negative sign of geraniol production by the BYa3105 strain, as the monoterpene is reported to display a light-yellow color (NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION, 2004). Secondly, the cell pellet from BYa3105 was smaller and presented a stronger beige color than those from ST9250 and ST9202 (results not shown). Less cells could be related to growth impairment due to geraniol presence, but this would also be expected in the limonene producing ST9250 culture. Dodecane samples were then stored at -20º C until GC-MS analysis.

After GC-MS identification method establishment using geraniol standard in dodecane solutions, dodecane samples from cultures were analyzed in GC-MS, as described in section 4.1.1 (page 46). Geraniol was identified in the BYa3105 samples, while limonene was identified in the ST9250 samples (Fig. 21). This was achieved by crossing the mass spectra of chromatogram peaks with the GC-MS software database (shown in the Appendices Figures 15 and 16). At the retention time of 16.5 min, mass spectra peaks of dodecane samples from

BYa3105 and the geraniol standard were similar to one another and matched the m/z patterns expected for geraniol in the database, but were clearly different from those of ST9202 and ST9250. This indicates that only these two samples had geraniol. Similarly, at the limonene retention time of 9.5 min, only the dodecane sample of the ST9250 strain displayed the mass spectra peaks identified as limonene in the software database. No monoterpenes were found in the ST9202 or negative control samples.

Thus, despite the first impression by the color of samples, it was proven that the *crGES* gene integrated into BYa3105 was being expressed and that the geraniol synthase protein executed its expected function. This is the first known report, to date, of geraniol biosynthesis by *Y. lipolytica*. Moreover, the mass spectra of geraniol related compounds, like geranial, nerol, neral and geranic acid, were also identified by the software in the BYa3105. Their presence might be due to natural oxidation of geraniol (MĄCZKA; WIŃSKA; GRABARCZYK, 2020) or due to unspecific changes to geraniol by intrinsic enzymes. Therefore, another strategy for increasing geraniol titer is to target its metabolism. While this has been established in *S. cerevisiae* (BROWN et al., 2015), studies that characterize metabolizing enzymes in *Y. lipolytica* are still lacking.

Figure 21: Chromatogram from gas chromatography with dodecane samples from different strains obtained in the geraniol production assay.



Black line: geraniol standard solution in dodecane, with geraniol retention time (16.5 min) in detail; Brown line: ST9202 dodecane overlay sample; Blue line: ST9250 dodecane overlay sample, with limonene retention time (9.5 min) in detail; Pink line: BYa3105 dodecane overlay sample, with geraniol retention time (16.5 min) and geranic acid retention time (19 min) in detail.

A geraniol standard calibration curve was used to quantify the geraniol in the dodecane phases. It was found that the BYa3105 strain produced 35 ± 10 mg L^{-1} of geraniol in the culture medium (that is, based in the aqueous volume). Assuming that all glucose added was consumed, this represents a yield of 0.08 % C-mol geraniol C-mol^{-1} glucose, which is 0.14 % of the MVA pathway maximum theoretical yield as seen in Section 2.3.1 (page 25). In this context, it is reported that quantification of geraniol might be underestimated when using dodecane as extractant. In (CHACÓN et al., 2019), only 66 % of geraniol added to culture medium was recovered by the dodecane overlay. There are also reports of certain *Y. lipolytica* strains which may consume dodecane as carbon source (CHENG et al., 2019; OLIVEIRA, 2014). The possibility of altered *Y. lipolytica* growth due to dodecane use should be considered in future experiments. Screening for other extractants or extraction methods could be advantageous to the proposed process of geraniol biosynthesis with *Y. lipolytica*.

The obtained titer of geraniol is close to that of limonene in (ARNESEN et al., 2020), in which the ST9250 strain achieved 35.9 mg L^{-1} of limonene, also from YP with 80 g L^{-1} of glucose, as done in this work. For industrial applications and scale up endeavors, further research with the BYa3105 strain must focus on additional genetic and bioprocess engineering targets, mainly directed to increase geraniol titers and yield. For example, fed-batch cultures might be employed to explore the hp4d growth-dependent promoter after the deceleration phase of growth, in which low carbon source concentrations are observed, towards the end of cultivation. While it is reported that hp4d activity is higher during stationary growth phase (MADZAK; TRÉTON; BLANCHIN-ROLAND, 2000), this is also the phase in which glucose is expected to have depleted from the culture medium in the geraniol production assay conducted. Thus, a low glucose feed to the culture would supply carbon for geraniol synthesis during the stationary phase. In this context, the activity of the hp4d promoter might be investigated with quantitative PCR. Furthermore, fed-batch culture might be explore to increase acetyl-CoA availability through control of nitrogen sources. As discussed in section 2.4 (page 32), under nitrogen limiting conditions *Y. lipolytica* increases its acetyl-CoA pool for citrate biosynthesis (POORINMOHAMMAD; KERKHOVEN, 2021). This is another trait that can be explored in fed-batch operation to increase geraniol synthesis, given that pseudo-hyphal morphology is successfully avoided.

6 CONCLUSIONS

The work described here shows the first exploratory investigations and results required for the development of a bioprocess using a *Y. lipolytica* strain engineered for geraniol biosynthesis. To achieve this main objective, basic methods for culture of the base strain, manipulation of nucleic acid molecules and yeast transformation were conducted.

The strains W29, ST6512 and ST9202 were acquired and properly stored for future assays. Although two defined media were tested, only growth in YPD complex medium resulted in no interference on the *Y. lipolytica* cell morphology. It is thus required to test different media reagents and search the literature for other defined media. Nevertheless, YPD and BDM media were used in assessment of the growth physiology of *Y. lipolytica*. Further growth culture experiments of *Y. lipolytica* may use the conclusions form this work to form new hypotheses and plan experiment conditions.

While the main objective of this work was to generate the genetically modified *Y. lipolytica* strain capable of *de novo* geraniol synthesis, the specific objectives in section 3 (page 44) and the experiments described in the above sections also represent the first steps of a research line in development inside the Bioprocess Engineering Laboratory (BELa) at the Polytechnic School of the University of São Paulo. Therefore, it was also of interest to reproduce growth culture protocols and analyze the growth kinetics of *Y. lipolytica* here, aims which proved themselves more troublesome than first expected.

Aiming at the establishment of a suitable protocol for screening of transformants, colony PCR was conducted in *Y. lipolytica*. After screening of PCR conditions and final tests to increase amplification rates, the protocol employed could achieve only 20 % of amplification, indicating high incidence of false negatives. Given these results, a different approach was devised and a genomic DNA extraction kit was acquired. DNA extraction was capable of increasing the availability and purity of DNA template in the PCR, which resulted in 80 % correct amplification over false negatives.

Another preparation step for the transformation of *Y. lipolytica* cells is the manipulation and replication of plasmids. Here, the acquired plasmids were successfully cloned into *E. coli* for replication and storage. The plasmids were confirmed using their unique digestion maps.

Transformation of the ST9202 monoterpene chassis strain was conducted and integration of the expression cassette into the yeast genome was confirmed by genomic DNA PCR. Then, the genetically modified BYa3105 strain was cultivated and the dodecane overlay was used for extraction and later identification of the produced geraniol. Therefore, the project's proof of concept, which is the geraniol producing *Y. lipolytica* BYa3105 strain, was achieved (35 mg geraniol L^{-1} in complex medium).

The experiments described in this work set the basis for the development of a sustainable bioprocess for geraniol biosynthesis. Further experiments in the project require to assess the geraniol yield of the obtained BYa3105 strain, to increase the viability of geraniol production via *Y. lipolytica*. For this, different approaches might be explored, for instance defined media optimization, oxygen transfer enhancement, geraniol synthase cassette engineering, pathway compartmentalization and geraniol extraction method improvement. Moreover, it is necessary to assess the current economic, environmental and societal viability of the bioprocess for geraniol biosynthesis.

REFERENCES

ABDEL-MAWGOUD, A. M. et al. Metabolic engineering in the host *Yarrowia lipolytica*. **Metabolic Engineering**, v. 50, p. 192–208, nov. 2018.

ABUDAYYEH, O. O. et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. **Science**, v. 353, n. 6299, p. aaf5573, 5 aug. 2016.

ALIBABA.COM. **Geraniol Manufacturers, Suppliers and Exporters on Alibaba.com**. Disponível em: <https://www.alibaba.com/trade/search?fsb=y&IndexArea=product_en&CatId=&tab=&SearchText=geraniol&viewtype=>. Acesso em: 13 jun. 2022.

AMARADIO, M. N. et al. Pareto optimal metabolic engineering for the growth-coupled overproduction of sustainable chemicals. **Biotechnology and Bioengineering**, p. bit.28103, 21 apr. 2022.

ARNESEN, J. A. et al. *Yarrowia lipolytica* Strains Engineered for the Production of Terpenoids. **Frontiers in Bioengineering and Biotechnology**, v. 8, p. 945, 14 aug. 2020.

ARNESEN, J. A. et al. Production of abscisic acid in the oleaginous yeast *Yarrowia lipolytica*. **FEMS Yeast Research**, p. foac015, 11 mar. 2022a.

ARNESEN, J. A. et al. Engineering of *Yarrowia lipolytica* for the production of plant triterpenoids: Asiatic, madecassic, and arjunolic acids. **Metabolic Engineering Communications**, v. 14, p. e00197, jun. 2022b.

ARNESEN, J. A.; BORODINA, I. Engineering of Yarrowia lipolytica for terpenoid production. **Metabolic Engineering Communications**, v. 15, p. e00213, dec. 2022.

ATEF, M. M. et al. Mechanistic Insights into Ameliorating Effect of Geraniol on d-Galactose Induced Memory Impairment in Rats. **Neurochemical Research**, v. 47, n. 6, p. 1664–1678, jun. 2022.

AZAMBUJA, S. P. H. **Fisiologia e capacidade de acúmulo de lipídeos de diferentes linhagens de *Yarrowia lipolytica* e *Rhodosporidium toruloides* em meio contendo glicerol**. Mestra em Engenharia de Alimentos—Campinas: Universidade Estadual de Campinas, 29 mar. 2016.

BACCIOTTI, F. **Fisiologia e formação de partículas lipídicas durante o crescimento da levedura *Yarrowia lipolytica* IMUFRJ 50682.** Mestrado em Engenharia Química—São Paulo: Universidade de São Paulo, 8 jul. 2016.

BAE, S. et al. Multiplex Gene Disruption by Targeted Base Editing of *Yarrowia lipolytica* Genome Using Cytidine Deaminase Combined with the CRISPR/Cas9 System. **Biotechnology Journal**, v. 15, n. 1, p. 1900238, jan. 2020.

BAGHBAN, R. et al. Yeast Expression Systems: Overview and Recent Advances. **Molecular Biotechnology**, v. 61, n. 5, p. 365–384, may 2019.

BAISYA, D. et al. Genome-wide functional screens enable the prediction of high activity CRISPR-Cas9 and -Cas12a guides in *Yarrowia lipolytica*. **Nature Communications**, v. 13, n. 1, p. 922, dec. 2022.

BAKKALI, F. et al. Biological effects of essential oils – A review. **Food and Chemical Toxicology**, v. 46, n. 2, p. 446–475, feb. 2008.

BANERJEE, A.; SHARKEY, T. D. Methylerythritol 4-phosphate (MEP) pathway metabolic regulation. **Nat. Prod. Rep.**, v. 31, n. 8, p. 1043–1055, 2014.

BARTH, G.; GAILLARDIN, C. Physiology and genetics of the dimorphic fungus *Yarrowia lipolytica*. **FEMS Microbiology Reviews**, p. 19, 1997.

BECKER, T. et al. Biotechnology. Em: WILEY-VCH VERLAG GMBH & CO. KGAA (Ed.). **Ullmann's Encyclopedia of Industrial Chemistry**. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, 2007. p. a04_107.pub2.

BEOPOULOS, A.; CHARDOT, T.; NICAUD, J.-M. *Yarrowia lipolytica*: A model and a tool to understand the mechanisms implicated in lipid accumulation. **Biochimie**, v. 91, n. 6, p. 692–696, jun. 2009.

BERGMAN, A. et al. Functional expression and evaluation of heterologous phosphoketolases in *Saccharomyces cerevisiae*. **AMB Express**, v. 6, n. 1, p. 115, dec. 2016.

BHATTAMISRA, S. K. et al. Protective activity of geraniol against acetic acid and *Helicobacter pylori*-induced gastric ulcers in rats. **Journal of Traditional and Complementary Medicine**, v. 9, n. 3, p. 206–214, jul. 2019.

BORSENBERGER, V. et al. Multiple Parameters Drive the Efficiency of CRISPR/Cas9-Induced Gene Modifications in *Yarrowia lipolytica*. **Journal of Molecular Biology**, v. 430, n. 21, p. 4293–4306, oct. 2018.

BREDEWEG, E. L. et al. A molecular genetic toolbox for *Yarrowia lipolytica*. **Biotechnology for Biofuels**, v. 10, n. 1, p. 2, dec. 2017.

BROWN, S. et al. De novo production of the plant-derived alkaloid strictosidine in yeast. **Proceedings of the National Academy of Sciences**, v. 112, n. 11, p. 3205–3210, 17 mar. 2015.

BURSE, A.; BOLAND, W. Deciphering the route to cyclic monoterpenes in *Chrysomelina* leaf beetles: source of new biocatalysts for industrial application? **Zeitschrift für Naturforschung C**, v. 72, n. 9–10, p. 417–427, 26 sep. 2017.

CAO, X. et al. Metabolic engineering of oleaginous yeast *Yarrowia lipolytica* for limonene overproduction. **Biotechnology for Biofuels**, v. 9, n. 1, p. 214, dec. 2016.

CAO, X. et al. Enhancing Linalool Production by Engineering Oleaginous Yeast *Yarrowia lipolytica*. **Bioresource Technology**, v. 245, p. 17, 2017.

CARRAU, F. M. et al. De novo synthesis of monoterpenes by *Saccharomyces cerevisiae* wine yeasts. **FEMS Microbiology Letters**, p. 9, 2005.

CELIŃSKA, E. “Fight-flight-or-freeze” – how *Yarrowia lipolytica* responds to stress at molecular level? **Applied Microbiology and Biotechnology**, v. 106, n. 9–10, p. 3369–3395, may 2022.

CHACÓN, M. G. et al. Esterification of geraniol as a strategy for increasing product titre and specificity in engineered *Escherichia coli*. **Microbial Cell Factories**, v. 18, n. 1, p. 105, dec. 2019.

CHANG, M. C. Y.; KEASLING, J. D. Production of isoprenoid pharmaceuticals by engineered microbes. **Nature Chemical Biology**, v. 2, n. 12, p. 674–681, dec. 2006.

CHEN, D.-C.; BECKERICH, J.-M.; GAILLARDIN, C. One-step transformation of the dimorphic yeast *Yarrowia lipolytica*. **Applied Microbiology and Biotechnology**, v. 48, n. 2, p. 232–235, 25 aug. 1997.

CHEN, W.; VILJOEN, A. M. Geraniol — A review of a commercially important fragrance material. **South African Journal of Botany**, v. 76, n. 4, p. 643–651, oct. 2010.

CHEN, Y. et al. Primary and Secondary Metabolic Effects of a Key Gene Deletion (Δ *YPL062W*) in Metabolically Engineered Terpenoid-Producing *Saccharomyces cerevisiae*. **Applied and Environmental Microbiology**, v. 85, n. 7, p. e01990-18, apr. 2019.

CHENG, B.-Q. et al. Elevating Limonene Production in Oleaginous Yeast *Yarrowia lipolytica* via Genetic Engineering of Limonene Biosynthesis Pathway and Optimization of Medium Composition. **Biotechnology and Bioprocess Engineering**, v. 24, n. 3, p. 500–506, jun. 2019.

CHOI, K. R. et al. Systems Metabolic Engineering Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering. **Trends in Biotechnology**, v. 37, n. 8, p. 817–837, aug. 2019.

CHUANG, L.-T. et al. Co-expression of heterologous desaturase genes in *Yarrowia lipolytica*. **New Biotechnology**, v. 27, n. 4, p. 277–282, sep. 2010.

CLARK IV, G. S. Geraniol. **Perfumer & Flavorist**, v. 23, p. 19–25, 1 may 1998.

Climate Change 2022: Impacts, Adaptation and Vulnerability. IPCC’s Working Group II, , feb. 2022.

CLOMBURG, J. M. et al. The isoprenoid alcohol pathway, a synthetic route for isoprenoid biosynthesis. **Proceedings of the National Academy of Sciences**, v. 116, n. 26, p. 12810–12815, 25 jun. 2019.

COLLIAS, D.; BEISEL, C. L. CRISPR technologies and the search for the PAM-free nuclease. **Nature Communications**, v. 12, n. 1, p. 555, dec. 2021.

CONG, L. et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. **Science**, v. 339, n. 6121, p. 819–823, 15 feb. 2013.

DALOZE, D.; PASTEELS, J. M. Isolation of 8-hydroxygeraniol-8-O-β-D-glucoside, a probable intermediate in biosynthesis of iridoid monoterpenes, from defensive secretions of *Plagiodera versicolora* and *Gastrophysa viridula* (Coleoptera: Chrysomelidae). **J Chem Ecol**, v. 20, p. 9, 1994.

DARVISHI, F. et al. Advances in synthetic biology of oleaginous yeast *Yarrowia lipolytica* for producing non-native chemicals. **Applied Microbiology and Biotechnology**, v. 102, n. 14, p. 5925–5938, jul. 2018.

DASGUPTA, P. **The economics of biodiversity: the Dasgupta review: full report**. Updated: 18 February 2021 ed. London: HM Treasury, 2021.

DE LUCA, V. et al. Discovery and Functional Analysis of Monoterpenoid Indole Alkaloid Pathways in Plants. Em: **Methods in Enzymology**. [s.l.] Elsevier, 2012. v. 515p. 207–229.

DEHSHEIKH, A. B. et al. Monoterpenes: Essential Oil Components with Valuable Features. **Mini-Reviews in Medicinal Chemistry**, v. 20, n. 11, p. 958–974, 17 jul. 2020.

DEMYTTENAERE, J. C. R.; DEL CARMEN HERRERA, M.; DE KIMPE, N. Biotransformation of geraniol, nerol and citral by sporulated surface cultures of *Aspergillus niger* and *Penicillium sp.* **Phytochemistry**, v. 55, n. 4, p. 363–373, oct. 2000.

DEVILLERS, H. et al. Draft Genome Sequence of *Yarrowia lipolytica* Strain A-101 Isolated from Polluted Soil in Poland. **Genome Announcements**, p. 2, 2016.

DEVILLERS, H.; NEUVÉGLISE, C. Genome Sequence of the Oleaginous Yeast *Yarrowia lipolytica* H222. **Microbiology Resource Announcements**, v. 8, n. 4, p. MRA.01547-18, e01547-18, 24 jan. 2019.

DICARLO, J. E. et al. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. **Nucleic Acids Research**, v. 41, n. 7, p. 4336–4343, 1 apr. 2013.

DISSOOK, S. et al. Stable isotope and chemical inhibition analyses suggested the existence of a non-mevalonate-like pathway in the yeast *Yarrowia lipolytica*. **Scientific Reports**, v. 11, n. 1, p. 5598, dec. 2021.

DONG, L. et al. Characterization of two geraniol synthases from *Valeriana officinalis* and *Lippia dulcis*: Similar activity but difference in subcellular localization. **Metabolic Engineering**, v. 20, p. 198–211, nov. 2013.

DUAN, Y. et al. *Aspergillus oryzae* Biosynthetic Platform for *de Novo* Iridoid Production. **Journal of Agricultural and Food Chemistry**, v. 69, n. 8, p. 2501–2511, 3 mar. 2021.

DULERMO, R. et al. Using a vector pool containing variable-strength promoters to optimize protein production in *Yarrowia lipolytica*. **Microbial Cell Factories**, v. 16, n. 1, p. 31, dec. 2017.

DUSSÉAUX, S. et al. Transforming yeast peroxisomes into microfactories for the efficient production of high-value isoprenoids. **Proceedings of the National Academy of Sciences**, v. 117, n. 50, p. 31789–31799, 15 dec. 2020.

EGGERSDORFER, M. Terpenes. Em: WILEY-VCH VERLAG GMBH & CO. KGAA (Ed.). **Ullmann's Encyclopedia of Industrial Chemistry**. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, 2000. p. a26_205.

FABRIS, M. et al. Extrachromosomal Genetic Engineering of the Marine Diatom *Phaeodactylum tricornutum* Enables the Heterologous Production of Monoterpenoids. **ACS Synthetic Biology**, v. 9, n. 3, p. 598–612, 20 mar. 2020.

FAHLBUSCH, K.-G. et al. Flavors and Fragrances. Em: WILEY-VCH VERLAG GMBH & CO. KGAA (Ed.). **Ullmann's Encyclopedia of Industrial Chemistry**. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA, 2003. p. a11_141.

FAKAS, S. Lipid biosynthesis in yeasts: A comparison of the lipid biosynthetic pathway between the model nonoleaginous yeast *Saccharomyces cerevisiae* and the model oleaginous yeast *Yarrowia lipolytica*. **Engineering in Life Sciences**, v. 17, n. 3, p. 292–302, mar. 2017.

FARIA, J. M. S. et al. Biotransformation of menthol and geraniol by hairy root cultures of *Anethum graveolens*: effect on growth and volatile components. **Biotechnology Letters**, v. 31, n. 6, p. 897–903, jun. 2009.

FATIMA, K. et al. Geraniol exerts its antiproliferative action by modulating molecular targets in lung and skin carcinoma cells. **Phytotherapy Research**, v. 35, n. 7, p. 3861–3874, jul. 2021.

FISCHER, M. J. C. et al. Metabolic engineering of monoterpene synthesis in yeast. **Biotechnology and Bioengineering**, v. 108, n. 8, p. 10, 2011.

FRIEDMAN, R. M.; BEALL, A. E. Biotechnology by Mid-Century: Assessing current capabilities. Anticipating tomorrow's leaders. p. 206, 30 jun. 2020.

GAILLARDIN, C. M.; CHAROY, V.; HESLOT, H. A study of copulation, sporulation and meiotic segregation in *Candida lipolytica*. **Archiv fur Mikrobiologie**, v. 92, n. 1, p. 69–83, 1973.

GÁLVEZ-LÓPEZ, D. et al. The metabolism and genetic regulation of lipids in the oleaginous yeast *Yarrowia lipolytica*. **Brazilian Journal of Microbiology**, v. 50, n. 1, p. 23–31, jan. 2019.

GAO, S. et al. Multiplex gene editing of the *Yarrowia lipolytica* genome using the CRISPR-Cas9 system. **Journal of Industrial Microbiology and Biotechnology**, v. 43, n. 8, p. 1085–1093, 1 aug. 2016.

GAO, S. et al. Iterative integration of multiple-copy pathway genes in *Yarrowia lipolytica* for heterologous β-carotene production. **Metabolic Engineering**, v. 41, p. 10, 2017.

GERKE, J. et al. Production of the Fragrance Geraniol in Peroxisomes of a Product-Tolerant Baker's Yeast. **Frontiers in Bioengineering and Biotechnology**, v. 8, p. 582052, 23 sep. 2020.

GIACOMOBONO, R. et al. Modelling of the Citric Acid Production from Crude Glycerol by Wild-Type *Yarrowia lipolytica* DSM 8218 Using Response Surface Methodology (RSM). **Life**, v. 12, n. 5, p. 621, 21 apr. 2022.

GILBERT, L. A. et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. **Cell**, v. 159, n. 3, p. 647–661, oct. 2014.

GROENEWALD, M. et al. *Yarrowia lipolytica* : Safety assessment of an oleaginous yeast with a great industrial potential. **Critical Reviews in Microbiology**, v. 40, n. 3, p. 187–206, aug. 2014.

GRUCHATTKA, E. et al. In silico profiling of *Escherichia coli* and *Saccharomyces cerevisiae* as terpenoid factories. **Microbial Cell Factories**, v. 12, n. 1, p. 84, dec. 2013.

GÜNDÜZ ERGÜN, B. et al. Established and Upcoming Yeast Expression Systems. Em: GASSER, B.; MATTANOVICH, D. (Eds.). **Recombinant Protein Production in Yeast**. Methods in Molecular Biology. New York, NY: Springer New York, 2019. v. 1923p. 1–74.

GUO, Q. et al. High-yield α-humulene production in *Yarrowia lipolytica* from waste cooking oil based on transcriptome analysis and metabolic engineering. **Microbial Cell Factories**, v. 21, n. 1, p. 271, 24 dec. 2022a.

GUO, Y. et al. Dissecting carbon metabolism of *Yarrowia lipolytica* type strain W29 using genome-scale metabolic modelling. **Computational and Structural Biotechnology Journal**, v. 20, p. 2503–2511, 2022b.

HAGVALL, L. et al. Contact allergy to oxidized geraniol among Swedish dermatitis patients-A multicentre study by the Swedish Contact Dermatitis Research Group. **Contact Dermatitis**, v. 79, n. 4, p. 232–238, oct. 2018.

HOLKENBRINK, C. et al. EasyCloneYALI: CRISPR/Cas9-Based Synthetic Toolbox for Engineering of the Yeast *Yarrowia lipolytica*. **Biotechnology Journal**, v. 13, n. 9, p. 1700543, sep. 2018.

HOMEM, V. et al. Gone with the flow - Assessment of personal care products in Portuguese rivers. **Chemosphere**, v. 293, p. 133552, apr. 2022.

IGNEA, C. et al. Engineering Monoterpene Production in Yeast Using a Synthetic Dominant Negative Geranyl Diphosphate Synthase. **ACS Synthetic Biology**, v. 3, n. 5, p. 298–306, 16 may 2014.

IIJIMA, Y. et al. Characterization of Geraniol Synthase from the Peltate Glands of Sweet Basil. **Plant Physiology**, v. 134, n. 1, p. 370–379, jan. 2004.

ISHINO, Y. et al. Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. **Journal of Bacteriology**, v. 169, n. 12, p. 5429–5433, dec. 1987.

JAVED, M. R. et al. CRISPR-Cas System: History and Prospects as a Genome Editing Tool in Microorganisms. **Current Microbiology**, v. 75, n. 12, p. 1675–1683, dec. 2018.

JIANG, G.-Z. et al. Manipulation of GES and ERG20 for geraniol overproduction in *Saccharomyces cerevisiae*. **Metabolic Engineering**, v. 41, p. 57–66, may 2017.

JIANG, H.; WANG, X. Biosynthesis of monoterpenoid and sesquiterpenoid as natural flavors and fragrances. **Biotechnology Advances**, v. 65, p. 108151, jul. 2023.

JIANG, W. et al. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. **Nature Biotechnology**, v. 31, n. 3, p. 233–239, mar. 2013.

JIN, C.-C. et al. Boosting the biosynthesis of betulinic acid and related triterpenoids in *Yarrowia lipolytica* via multimodular metabolic engineering. **Microbial Cell Factories**, v. 18, n. 1, p. 77, dec. 2019.

JINEK, M. et al. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. **Science**, v. 337, n. 6096, p. 816–821, 17 aug. 2012.

KIRBY, J. et al. Engineering a functional 1-deoxy-D-xylulose 5-phosphate (DXP) pathway in *Saccharomyces cerevisiae*. **Metabolic Engineering**, v. 38, p. 494–503, nov. 2016.

KLUG, L.; DAUM, G. Yeast lipid metabolism at a glance. **FEMS Yeast Research**, v. 14, n. 3, p. 369–388, may 2014.

KOLEWE, M. E.; GAURAV, V.; ROBERTS, S. C. Pharmaceutically Active Natural Product Synthesis and Supply via Plant Cell Culture Technology. **Molecular Pharmaceutics**, v. 5, n. 2, p. 243–256, 1 apr. 2008.

KONG, J. et al. Enhanced production of amyrin in *Yarrowia lipolytica* using a combinatorial protein and metabolic engineering approach. **Microbial Cell Factories**, v. 21, n. 1, p. 186, 9 sep. 2022.

KORPYS-WOŹNIAK, P. et al. Impact of overproduced heterologous protein characteristics on physiological response in *Yarrowia lipolytica* steady-state-maintained continuous cultures. **Applied Microbiology and Biotechnology**, v. 104, n. 22, p. 9785–9800, nov. 2020.

KRIVORUCHKO, A.; NIELSEN, J. Production of natural products through metabolic engineering of *Saccharomyces cerevisiae*. **Current Opinion in Biotechnology**, v. 35, p. 7–15, dec. 2015.

KUNG, S. H. et al. Approaches and Recent Developments for the Commercial Production of Semi-synthetic Artemisinin. **Frontiers in Plant Science**, v. 9, p. 87, 31 jan. 2018.

LAPCZYNSKI, A. et al. Fragrance material review on geraniol. **Food and Chemical Toxicology**, v. 46, n. 11, p. S160–S170, nov. 2008.

LEE DÍAZ, A. S. et al. Exploring the Volatiles Released from Roots of Wild and Domesticated Tomato Plants under Insect Attack. **Molecules**, v. 27, n. 5, p. 1612, 28 feb. 2022.

LEE, J. W. et al. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. **Nature Chemical Biology**, v. 8, n. 6, p. 536–546, jun. 2012.

LEI, Y. et al. Pharmacological Properties of Geraniol – A Review. **Planta Medica**, v. 85, n. 01, p. 48–55, jan. 2019.

LI, D. et al. Production of Triterpene Ginsenoside Compound K in the Non-conventional Yeast *Yarrowia lipolytica*. **Journal of Agricultural and Food Chemistry**, v. 67, n. 9, p. 2581–2588, 6 mar. 2019.

LI, D. et al. Metabolic engineering of *Yarrowia lipolytica* for heterologous oleanolic acid production. **Chemical Engineering Science**, v. 218, p. 115529, jun. 2020.

LI, J. et al. Simultaneous Improvement of Limonene Production and Tolerance in *Yarrowia lipolytica* through Tolerance Engineering and Evolutionary Engineering. **ACS Synthetic Biology**, v. 10, n. 4, p. 884–896, 16 apr. 2021a.

LI, M.; XU, S.; LU, W. Engineering *Corynebacterium glutamicum* for Geraniol Production. **Transactions of Tianjin University**, v. 27, n. 5, p. 377–384, oct. 2021.

LI, Y.-W. et al. YALIcloneNHEJ: An Efficient Modular Cloning Toolkit for NHEJ Integration of Multigene Pathway and Terpenoid Production in *Yarrowia lipolytica*. **Frontiers in Bioengineering and Biotechnology**, v. 9, p. 816980, 2 mar. 2022.

LI, Z.-J. et al. Advanced Strategies for the Synthesis of Terpenoids in *Yarrowia lipolytica*. **Journal of Agricultural and Food Chemistry**, v. 69, n. 8, p. 2367–2381, 3 mar. 2021b.

LIANG, Z. et al. Genetic engineering of yeast, filamentous fungi and bacteria for terpene production and applications in food industry. **Food Research International**, v. 147, p. 110487, sep. 2021.

LIU, F. et al. Enhancing *Trans* -Nerolidol Productivity in *Yarrowia lipolytica* by Improving Precursor Supply and Optimizing Nerolidol Synthase Activity. **Journal of Agricultural and Food Chemistry**, v. 70, n. 48, p. 15157–15165, 7 dec. 2022a.

LIU, G. et al. The CRISPR-Cas toolbox and gene editing technologies. **Molecular Cell**, v. 82, n. 2, p. 333–347, jan. 2022b.

LIU, J. et al. Overproduction of geraniol by enhanced precursor supply in *Saccharomyces cerevisiae*. **Journal of Biotechnology**, v. 168, n. 4, p. 446–451, dec. 2013.

LIU, M. et al. Morphological and Metabolic Engineering of *Yarrowia lipolytica* to Increase β-Carotene Production. **ACS Synthetic Biology**, p. acssynbio.1c00480, 11 nov. 2021a.

LIU, W. et al. Engineering *Escherichia coli* for high-yield geraniol production with biotransformation of geranyl acetate to geraniol under fed-batch culture. **Biotechnology for Biofuels**, v. 9, n. 1, p. 58, dec. 2016.

LIU, Y. et al. Monoterpenoid biosynthesis by engineered microbes. **Journal of Industrial Microbiology and Biotechnology**, v. 0, p. 11, 2021b.

LIU, Y. et al. Engineering *Yarrowia lipolytica* for the sustainable production of β-farnesene from waste oil feedstock. **Biotechnology for Biofuels and Bioproducts**, v. 15, n. 1, p. 101, 3 oct. 2022c.

LIU, Y. et al. Mechanistic insight of the potential of geraniol against Alzheimer's disease. **European Journal of Medical Research**, v. 27, n. 1, p. 93, dec. 2022d.

LOŽIENĖ, K.; VAIČIULYTĖ, V. Geraniol and Carvacrol in Essential Oil Bearing Thymus pulegioides: Distribution in Natural Habitats and Phytotoxic Effect. **Molecules**, v. 27, n. 3, p. 986, 1 feb. 2022.

LU, R. et al. Engineering *Yarrowia lipolytica* to produce advanced biofuels: Current status and perspectives. **Bioresource Technology**, v. 341, p. 125877, dec. 2021.

LU, Y. et al. A modular pathway engineering strategy for the high-level production of β-ionone in *Yarrowia lipolytica*. **Microbial Cell Factories**, v. 19, n. 1, p. 49, dec. 2020.

LUAN, F. Differential incorporation of 1-deoxy-d-xylulose into (3S)-linalool and geraniol in grape berry exocarp and mesocarp. p. 9, 2002.

LUPISH, B. et al. Genome-wide CRISPR-Cas9 screen reveals a persistent null-hyphal phenotype that maintains high carotenoid production in *Yarrowia lipolytica*. **Biotechnology and Bioengineering**, p. bit.28219, 30 aug. 2022.

LYU, Z. et al. Engineering the Autotroph *Methanococcus maripaludis* for Geraniol Production. **ACS Synthetic Biology**, v. 5, n. 7, p. 577–581, 15 jul. 2016.

MA, Y.-R. et al. Advances in the metabolic engineering of *Yarrowia lipolytica* for the production of terpenoids. **Bioresource Technology**, v. 281, p. 449–456, jun. 2019.

MĄCZKA, W.; WIŃSKA, K.; GRABARCZYK, M. One Hundred Faces of Geraniol. **Molecules**, v. 25, n. 14, p. 3303, 21 jul. 2020.

MADE-IN-CHINA.COM. **Geraniol, China Geraniol, Geraniol Manufacturers, China Geraniol catalog**. Disponível em: <https://www.made-in-china.com/productdirectory.do?word=Geraniol&file=&searchType=0&subaction=hunt&style=b&mode=and&code=0&comProvince=nolimit&order=0&isOpenCorrection=1&org=top>. Acesso em: 13 jun. 2022.

MADZAK, C. *Yarrowia lipolytica* engineering as a source of microbial cell factories. Em: **Microbial Cell Factories Engineering for Production of Biomolecules**. [s.l.] Elsevier, 2021a.

MADZAK, C. *Yarrowia lipolytica* Strains and Their Biotechnological Applications: How Natural Biodiversity and Metabolic Engineering Could Contribute to Cell Factories Improvement. **Journal of Fungi**, v. 7, n. 7, p. 548, 10 jul. 2021b.

MADZAK, C.; GAILLARDIN, C.; BECKERICH, J.-M. Heterologous protein expression and secretion in the non-conventional yeast *Yarrowia lipolytica*: a review. **Journal of Biotechnology**, v. 109, n. 1–2, p. 63–81, apr. 2004.

MADZAK, C.; TRÉTON, B.; BLANCHIN-ROLAND, S. Strong Hybrid Promoters and Integrative Expression/Secretion Vectors for Quasi-Constitutive Expression of Heterologous Proteins in the Yeast *Yarrowia lipolytica*. **Journal of Molecular Microbiology and Biotechnology**, v. 2, n. 2, p. 11, 2000.

MAGNAN, C. et al. Sequence Assembly of *Yarrowia lipolytica* Strain W29/CLIB89 Shows Transposable Element Diversity. **PLOS ONE**, v. 11, n. 9, p. e0162363, 7 sep. 2016.

MAHMOUD, S. S.; CROTEAU, R. B. Strategies for transgenic manipulation of monoterpene biosynthesis in plants. **Trends in Plant Science**, v. 7, n. 8, p. 366–373, aug. 2002.

MAKRI, A.; FAKAS, S.; AGGELIS, G. Metabolic activities of biotechnological interest in *Yarrowia lipolytica* grown on glycerol in repeated batch cultures. **Bioresource Technology**, v. 101, n. 7, p. 2351–2358, apr. 2010.

MALI, P. et al. RNA-Guided Human Genome Engineering via Cas9. **Science**, v. 339, n. 6121, p. 823–826, 15 feb. 2013.

MARELLA, E. R. et al. A single-host fermentation process for the production of flavor lactones from non-hydroxylated fatty acids. **Metabolic Engineering**, p. 10, 2020.

MARSAFARI, M.; XU, P. Debottlenecking mevalonate pathway for antimalarial drug precursor amorphadiene biosynthesis in *Yarrowia lipolytica*. **Metabolic Engineering Communications**, v. 10, p. e00121, jun. 2020.

MATTHÄUS, F. et al. Production of Lycopene in the Non-Carotenoid-Producing Yeast *Yarrowia lipolytica*. **Applied and Environmental Microbiology**, v. 80, n. 5, p. 1660–1669, 1 mar. 2014.

MAURY, J. et al. Microbial Isoprenoid Production: An Example of Green Chemistry through Metabolic Engineering. Em: NIELSEN, J. (Ed.). **Biotechnology for the Future**. Advances in Biochemical Engineering/Biotechnology. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005. v. 100p. 19–51.

MEDICHERLA, K. et al. Oral administration of geraniol ameliorates acute experimental murine colitis by inhibiting pro-inflammatory cytokines and NF-κB signaling. **Food & Function**, v. 6, n. 9, p. 2984–2995, 2015.

MIETTINEN, K. The seco-iridoid pathway from Catharanthus roseus. **Nature Communications**, p. 11, 2014.

MILLER, K. K.; ALPER, H. S. *Yarrowia lipolytica*: more than an oleaginous workhorse. **Applied Microbiology and Biotechnology**, v. 103, n. 23–24, p. 9251–9262, dec. 2019.

MISA, J. et al. Engineered Production of Strictosidine and Analogues in Yeast. **ACS Synthetic Biology**, p. acssynbio.2c00037, 16 mar. 2022.

MISHRA, S. et al. Engineering a Carotenoid-Overproducing Strain of *Azospirillum brasilense* for Heterologous Production of Geraniol and Amorphadiene. **Applied and Environmental Microbiology**, v. 86, n. 17, p. e00414-20, /aem/86/17/AEM.00414- 20.atom, 26 jun. 2020.

MITSUI CHEMICALS, INC. **Products & Materials List**. Disponível em: <https://jp.mitsuichemicals.com/en/service/search_results.htm>. Acesso em: 2 jun. 2022.

MIZIORKO, H. M. Enzymes of the mevalonate pathway of isoprenoid biosynthesis. **Archives of Biochemistry and Biophysics**, p. 13, 2011.

MOSER, S.; PICHLER, H. Identifying and engineering the ideal microbial terpenoid production host. **Applied Microbiology and Biotechnology**, v. 103, n. 14, p. 5501–5516, jul. 2019.

MUHAMMAD, A. et al. Production of plant natural products through engineered *Yarrowia lipolytica*. **Biotechnology Advances**, v. 43, p. 107555, nov. 2020.

NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. **PubChem Compound Summary for CID 637566, Geraniol**. Disponível em: <https://pubchem.ncbi.nlm.nih.gov/compound/Geraniol>. Acesso em: 21 sep. 2021.

NICAUD, J.-M. et al. Protein expression and secretion in the yeast *Yarrowia lipolytica*. **FEMS Yeast Research**, p. 9, 2002.

NIELSEN, J. Metabolic engineering. **Applied Microbiology and Biotechnology**, v. 55, n. 3, p. 263–283, 1 apr. 2001.

O'CONNOR, S. E.; MARESH, J. J. Chemistry and biology of monoterpene indole alkaloid biosynthesis. **Natural Product Reports**, v. 23, n. 4, p. 532, 2006.

OHASHI, Y.; HUANG, S.; MAEDA, I. Biosyntheses of geranic acid and citronellic acid from monoterpene alcohols by *Saccharomyces cerevisiae*. **Bioscience, Biotechnology, and Biochemistry**, v. 85, n. 6, p. 1530–1535, 25 may 2021.

OLIVEIRA, P. H. S. **Análise fisiológica e cinética do crescimento da levedura oleaginosa *Yarrowia lipolytica* IMUFRJ 50682 em diferentes fontes de carbono.** Mestrado em Engenharia Química—São Paulo: Universidade de São Paulo, 12 aug. 2014.

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

PANG, Y. et al. Engineering the oleaginous yeast *Yarrowia lipolytica* to produce limonene from waste cooking oil. **Biotechnology for Biofuels**, v. 12, n. 1, p. 241, dec. 2019.

PAPANIKOLAOU, I. CHEVALOT, M. KOMAI, S. Single cell oil production by *Yarrowia lipolytica* growing on an industrial derivative of animal fat in batch cultures. **Applied Microbiology and Biotechnology**, v. 58, n. 3, p. 308–312, 1 mar. 2002.

PAPANIKOLAOU, S. et al. Accumulation of a Cocoa-Butter-Like Lipid by *Yarrowia lipolytica* Cultivated on Agro-Industrial Residues. **Current Microbiology**, v. 46, n. 2, p. 124–130, 1 feb. 2003.

PAPANIKOLAOU, S.; AGGELIS, G. Modeling Lipid Accumulation and Degradation in *Yarrowia lipolytica* Cultivated on Industrial Fats. **Current Microbiology**, v. 46, n. 6, p. 398–402, 1 jun. 2003.

PENG, B. et al. Engineered protein degradation of farnesyl pyrophosphate synthase is an effective regulatory mechanism to increase monoterpene production in *Saccharomyces cerevisiae*. **Metabolic Engineering**, feb. 2018.

PEREIRA, A. S. et al. Bio-oil production for biodiesel industry by *Yarrowia lipolytica* from volatile fatty acids in two-stage batch culture. **Applied Microbiology and Biotechnology**, v. 106, n. 8, p. 2869–2881, apr. 2022.

PONZONI, C. et al. Biotransformation of Acyclic Monoterpenoids by *Debaryomyces sp.*, *Kluyveromyces sp.*, and *Pichia sp.* Strains of Environmental Origin. **Chemistry & Biodiversity**, v. 5, n. 3, p. 471–483, mar. 2008.

POORINMOHAMMAD, N.; KERKHOVEN, E. J. Systems-level approaches for understanding and engineering of the oleaginous cell factory *Yarrowia lipolytica*. **Biotechnology and Bioengineering**, p. bit.27859, 21 jun. 2021.

PRIEBE, X. et al. Rational selection of biphasic reaction systems for geranyl glucoside production by *Escherichia coli* whole-cell biocatalysts. **Enzyme and Microbial Technology**, v. 112, p. 79–87, may 2018.

QI, L. S. et al. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. **Cell**, v. 152, n. 5, p. 1173–1183, feb. 2013.

RASTOGI, S. C. et al. Deodorants on the European market: quantitative chemical analysis of 21 fragrances. **Contact Dermatitis**, v. 38, n. 1, p. 29–35, jan. 1998.

RASTOGI, S. C. et al. Fragrance chemicals in domestic and occupational products. **Contact Dermatitis**, v. 45, n. 1, p. 221–225, 2001.

REDDY, G. K. et al. Exploring novel bacterial terpene synthases. **PLOS ONE**, v. 15, n. 4, p. e0232220, 30 apr. 2020.

RUAN, Q. et al. Current advances of endophytes as a platform for production of anti-cancer drug camptothecin. **Food and Chemical Toxicology**, v. 151, p. 112113, may 2021.

RUZICKA, L. The isoprene rule and the biogenesis of terpenic compounds. **Experimentia**, v. 9, p. 11, 1953.

SAMBROOK, J.; RUSSELL, D. W. **Molecular Cloning - A Laboratory Manual**. 3. ed. New York: Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 2001.

SANDERSON, K. Chemistry: It's not easy being green. **Nature**, v. 469, p. 3, jan. 2011.

SCARIOT, F. J. et al. Citral and geraniol induce necrotic and apoptotic cell death on *Saccharomyces cerevisiae*. **World Journal of Microbiology and Biotechnology**, v. 37, n. 3, p. 42, mar. 2021.

SCHMIDT-DANNERT, C. The future of biologically inspired next-generation factories for chemicals. **Microbial Biotechnology**, v. 10, n. 5, p. 1164–1166, sep. 2017.

SCHWARTZ, C. et al. CRISPRi repression of nonhomologous end-joining for enhanced genome engineering via homologous recombination in *Yarrowia lipolytica*. **Biotechnology and Bioengineering**, v. 114, n. 12, p. 2896–2906, dec. 2017.

SCHWARTZ, C. et al. Validating genome-wide CRISPR-Cas9 function improves screening in the oleaginous yeast *Yarrowia lipolytica*. **Metabolic Engineering**, v. 55, p. 102–110, sep. 2019.

SELL, C. S. **A Fragrant Introduction to Terpenoid Chemistry:** Cambridge: Royal Society of Chemistry, 2007.

SHARMEEN, J. B. et al. Essential Oils as Natural Sources of Fragrance Compounds for Cosmetics and Cosmeceuticals. **Molecules**, v. 26, n. 3, p. 666, 27 jan. 2021.

SILVESTRE, J. F. et al. Sensitization to fragrances in Spain: A 5-year multicentre study (2011-2015). **Contact Dermatitis**, v. 80, n. 2, p. 94–100, feb. 2019.

SIMKIN, A. J. et al. Characterization of the plastidial geraniol synthase from Madagascar periwinkle which initiates the monoterpenoid branch of the alkaloid pathway in internal phloem associated parenchyma. **Phytochemistry**, v. 85, p. 36–43, jan. 2013.

SOLIMAN, S. S. M.; TSAO, R.; RAIZADA, M. N. Chemical Inhibitors Suggest Endophytic Fungal Paclitaxel Is Derived from Both Mevalonate and Non-mevalonate-like Pathways. **Journal of Natural Products**, v. 74, n. 12, p. 2497–2504, 27 dec. 2011.

SOONG, Y. V. et al. Cellular and metabolic engineering of oleaginous yeast *Yarrowia lipolytica* for bioconversion of hydrophobic substrates into high-value products. **Engineering in Life Sciences**, v. 19, n. 6, p. 423–443, jun. 2019.

STEFFEN, W. et al. Planetary boundaries: Guiding human development on a changing planet. **Science**, v. 347, n. 6223, p. 1259855, 13 feb. 2015.

STOVICEK, V.; HOLKENBRINK, C.; BORODINA, I. CRISPR/Cas system for yeast genome engineering: advances and applications. **FEMS Yeast Research**, v. 17, n. 5, 1 aug. 2017.

SUN, C.; THEODOROPOULOS, C.; SCRUTTON, N. S. Techno-economic assessment of microbial limonene production. **Bioresource Technology**, v. 300, p. 122666, mar. 2020.

SUPPAKUL, P. et al. Antimicrobial Properties of Basil and Its Possible Application in Food Packaging. **Journal of Agricultural and Food Chemistry**, v. 51, p. 11, 2003.

TAI, M.; STEPHANOPOULOS, G. Engineering the push and pull of lipid biosynthesis in oleaginous yeast *Yarrowia lipolytica* for biofuel production. **Metabolic Engineering**, v. 15, p. 1–9, jan. 2013.

TASHIRO, M. et al. Directed evolution and expression tuning of geraniol synthase for efficient geraniol production in *Escherichia coli*. **The Journal of General and Applied Microbiology**, v. 63, n. 5, p. 287–295, 2017.

VANDERMIES, M.; FICKERS, P. Bioreactor-Scale Strategies for the Production of Recombinant Protein in the Yeast *Yarrowia lipolytica*. **Microorganisms**, v. 7, n. 2, p. 40, 30 jan. 2019.

VERDUYN, C. et al. Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. **Yeast**, v. 8, n. 7, p. 501–517, jul. 1992.

VIEIRA GOMES, A. et al. Comparison of Yeasts as Hosts for Recombinant Protein Production. **Microorganisms**, v. 6, n. 2, p. 38, 29 apr. 2018.

WANG, X. et al. Engineering *Escherichia coli* for production of geraniol by systematic synthetic biology approaches and laboratory-evolved fusion tags. **Metabolic Engineering**, v. 66, p. 60–67, jul. 2021.

WANY, A. et al. Extraction and characterization of essential oil components based on geraniol and citronellol from Java citronella (*Cymbopogon winterianus* Jowitt). **Plant Growth Regulation**, v. 73, n. 2, p. 133–145, jun. 2014.

WEI, L.-J. et al. Biosynthesis of α-Pinene by Genetically Engineered *Yarrowia lipolytica* from Low-Cost Renewable Feedstocks. **Journal of Agricultural and Food Chemistry**, v. 69, n. 1, p. 275–285, 13 jan. 2021.

WILSON, S. A.; ROBERTS, S. C. Recent advances towards development and commercialization of plant cell culture processes for the synthesis of biomolecules: Development and commercialization of plant cell culture. **Plant Biotechnology Journal**, v. 10, n. 3, p. 249–268, apr. 2012.

WOLF, K. **Nonconventional Yeasts in Biotechnology**. Berlin, Heidelberg: Springer Berlin Heidelberg, 1996.

WRIESSNEGGER, T.; PICHLER, H. Yeast metabolic engineering – Targeting sterol metabolism and terpenoid formation. **Progress in Lipid Research**, v. 52, n. 3, p. 277–293, jul. 2013.

WU, W.; MARAVELIAS, C. T. Synthesis and techno-economic assessment of microbial-based processes for terpenes production. **Biotechnology for Biofuels**, v. 11, n. 1, p. 294, dec. 2018.

XIE, D. Integrating Cellular and Bioprocess Engineering in the Non-Conventional Yeast *Yarrowia lipolytica* for Biodiesel Production: A Review. **Frontiers in Bioengineering and Biotechnology**, v. 5, p. 65, 17 oct. 2017.

XIE, Y.; CHEN, S.; XIONG, X. Metabolic Engineering of Non-carotenoid-Producing Yeast *Yarrowia lipolytica* for the Biosynthesis of Zeaxanthin. **Frontiers in Microbiology**, v. 12, p. 699235, 7 oct. 2021.

YANG, L. et al. Response of Plant Secondary Metabolites to Environmental Factors. **Molecules**, v. 23, n. 4, p. 762, 27 mar. 2018.

YAO, F. et al. Engineering oleaginous yeast *Yarrowia lipolytica* for enhanced limonene production from xylose and lignocellulosic hydrolysate. **FEMS Yeast Research**, v. 20, n. 6, p. foaa046, 26 sep. 2020.

YEE, D. A. et al. Engineered mitochondrial production of monoterpenes in *Saccharomyces cerevisiae*. **Metabolic Engineering**, v. 55, p. 76–84, sep. 2019.

YUAN, S.-F.; ALPER, H. S. Metabolic engineering of microbial cell factories for production of nutraceuticals. **Microbial Cell Factories**, v. 18, n. 1, p. 46, dec. 2019.

ZADEK, S. et al. Aligning Markets with Biodiversity. p. 50, jun. 2021.

ZEBEC, Z. et al. Towards synthesis of monoterpenes and derivatives using synthetic biology. **Current Opinion in Chemical Biology**, p. 7, 2016.

ZHANG, G. et al. Metabolic engineering of *Yarrowia lipolytica* for terpenoids production: advances and perspectives. **Critical Reviews in Biotechnology**, p. 1–16, 29 jul. 2021.

ZHANG, J. et al. A microbial supply chain for production of the anti-cancer drug vinblastine. **Nature**, v. 609, n. 7926, p. 341–347, 8 sep. 2022.

ZHAO, C. et al. Tuning Geraniol Biosynthesis via a Novel Decane-Responsive Promoter in *Candida glycerinogenes*. **ACS Synthetic Biology**, v. 11, n. 5, p. 1835–1844, 20 may 2022.

ZHAO, J. et al. Improving monoterpene geraniol production through geranyl diphosphate synthesis regulation in *Saccharomyces cerevisiae*. **Applied Microbiology and Biotechnology**, v. 100, n. 10, p. 4561–4571, may 2016.

ZHAO, J. et al. Dynamic control of ERG20 expression combined with minimized endogenous downstream metabolism contributes to the improvement of geraniol production in *Saccharomyces cerevisiae*. **Microbial Cell Factories**, v. 16, n. 1, p. 17, dec. 2017.

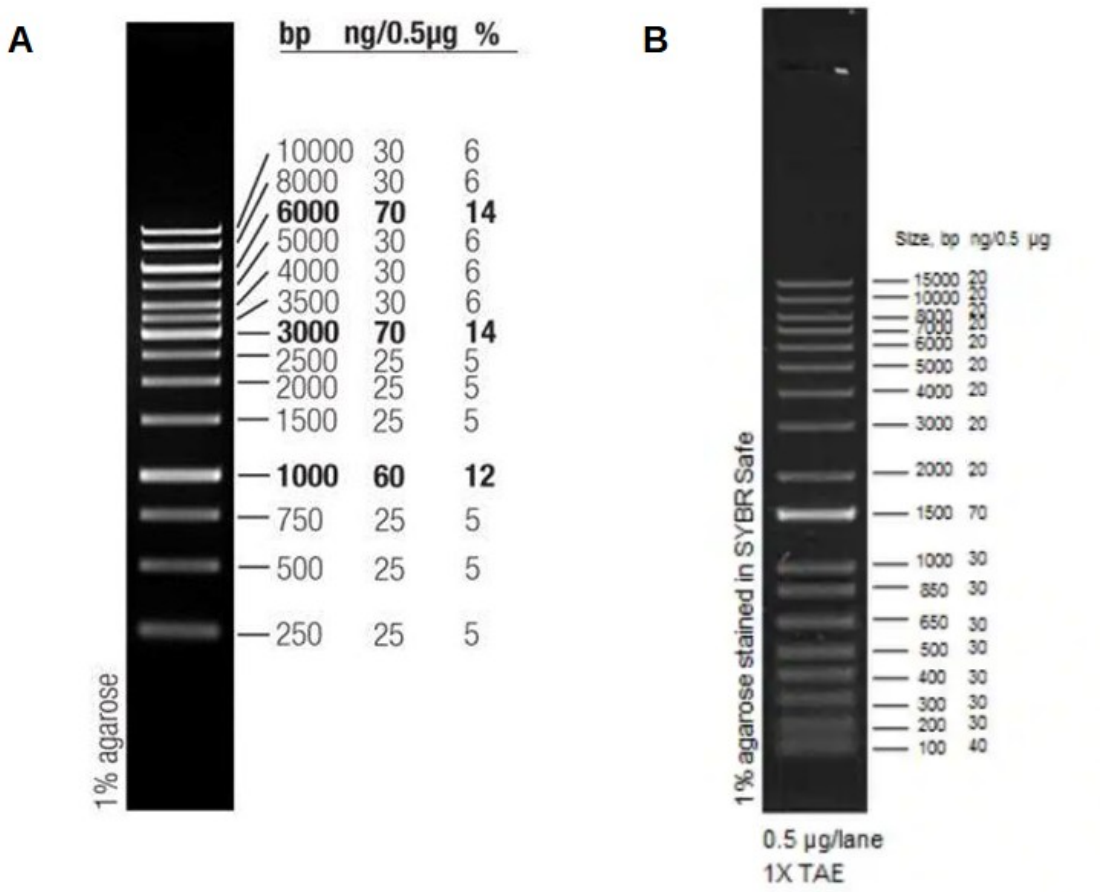
ZHOU, J. et al. Engineering *Escherichia coli* for selective geraniol production with minimized endogenous dehydrogenation. **Journal of Biotechnology**, p. 9, 2014.

ZHU, H.-Z. et al. Production of High Levels of 3 *S* ,3′ *S* -Astaxanthin in *Yarrowia lipolytica* via Iterative Metabolic Engineering. **Journal of Agricultural and Food Chemistry**, v. 70, n. 8, p. 2673–2683, 2 mar. 2022.

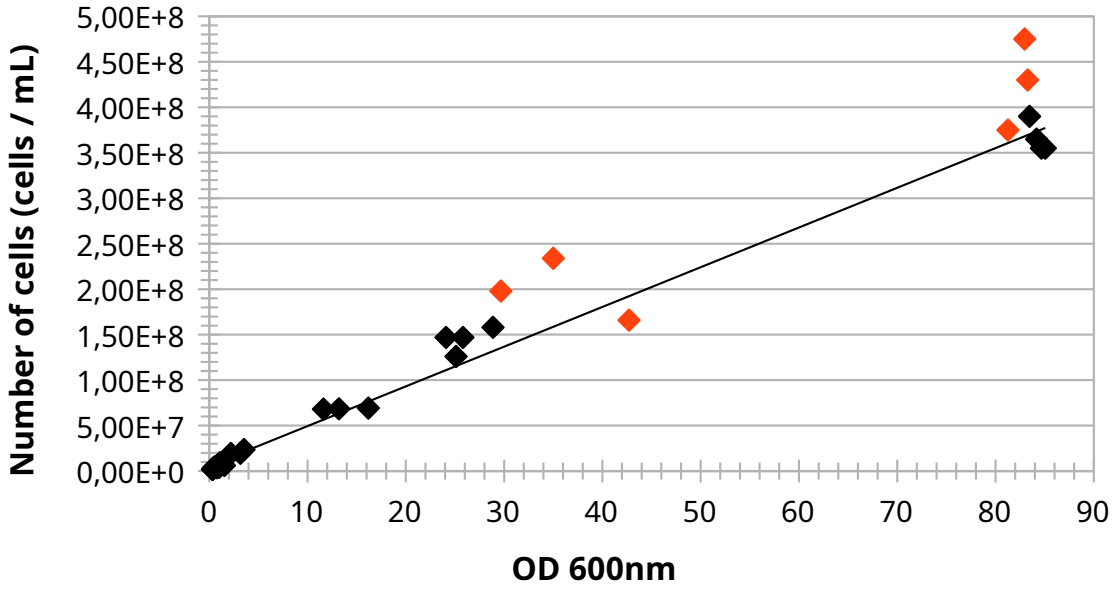
ZHU, K. et al. Metabolic engineering of microbes for monoterpenoid production. **Biotechnology Advances**, v.

53, p. 107837, dec. 2021.

APPENDICES

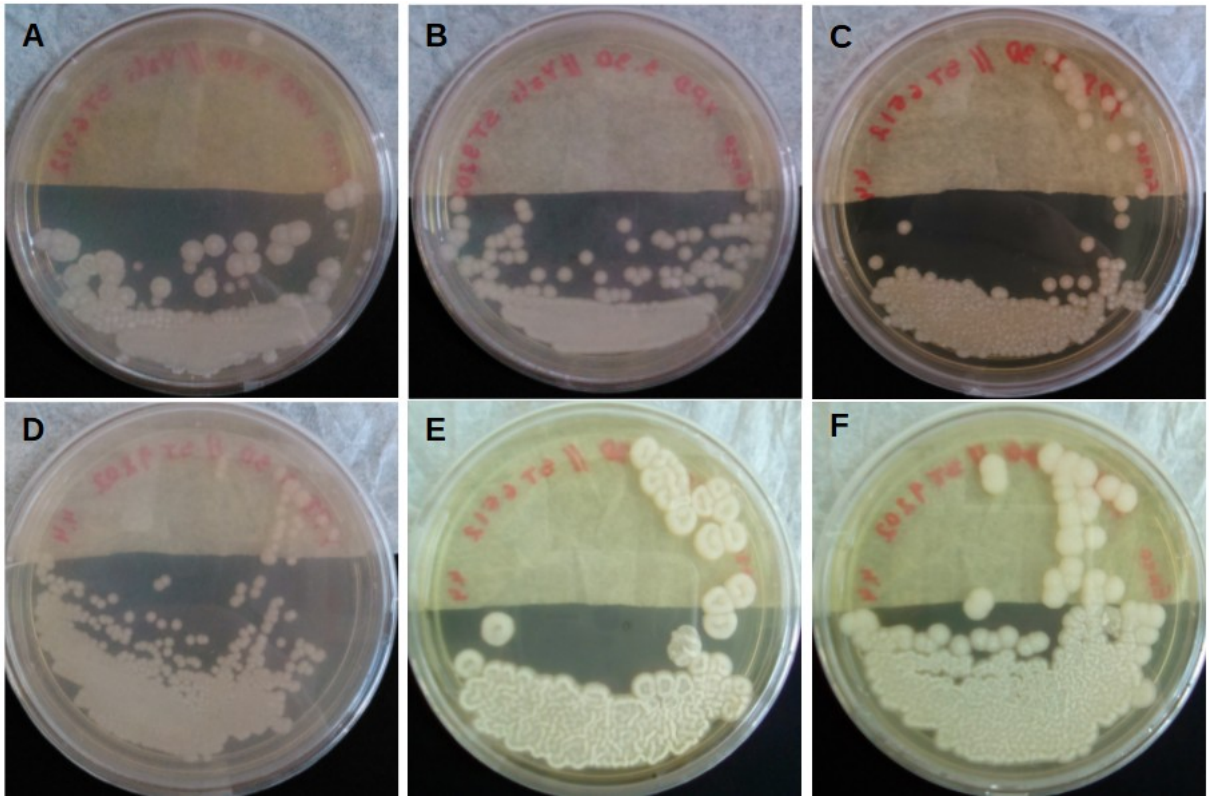


App.Figure 1: Patterns of DNA bands of ladders used in this work. (A) GeneRuler 1 kb DNA Ladder from Thermo Scientific. (B) 1 kb Plus DNA Ladder from Thermo Scientific. Source: Thermo Scientific website.

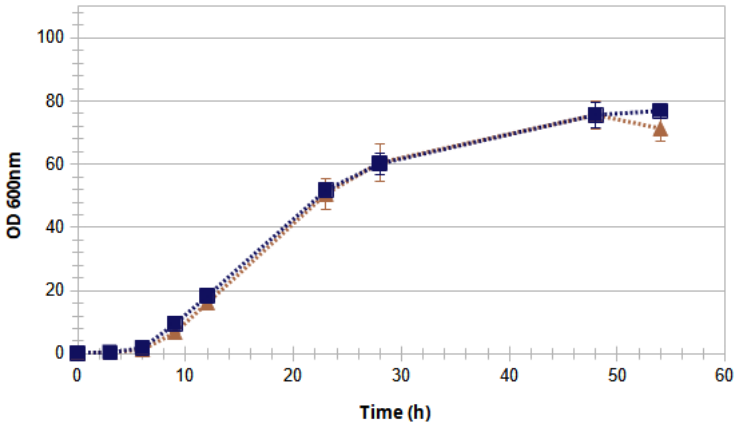


f(x) = 4367119,51809717 x + 5716402,89914586
R^2 = 0,989554546719518

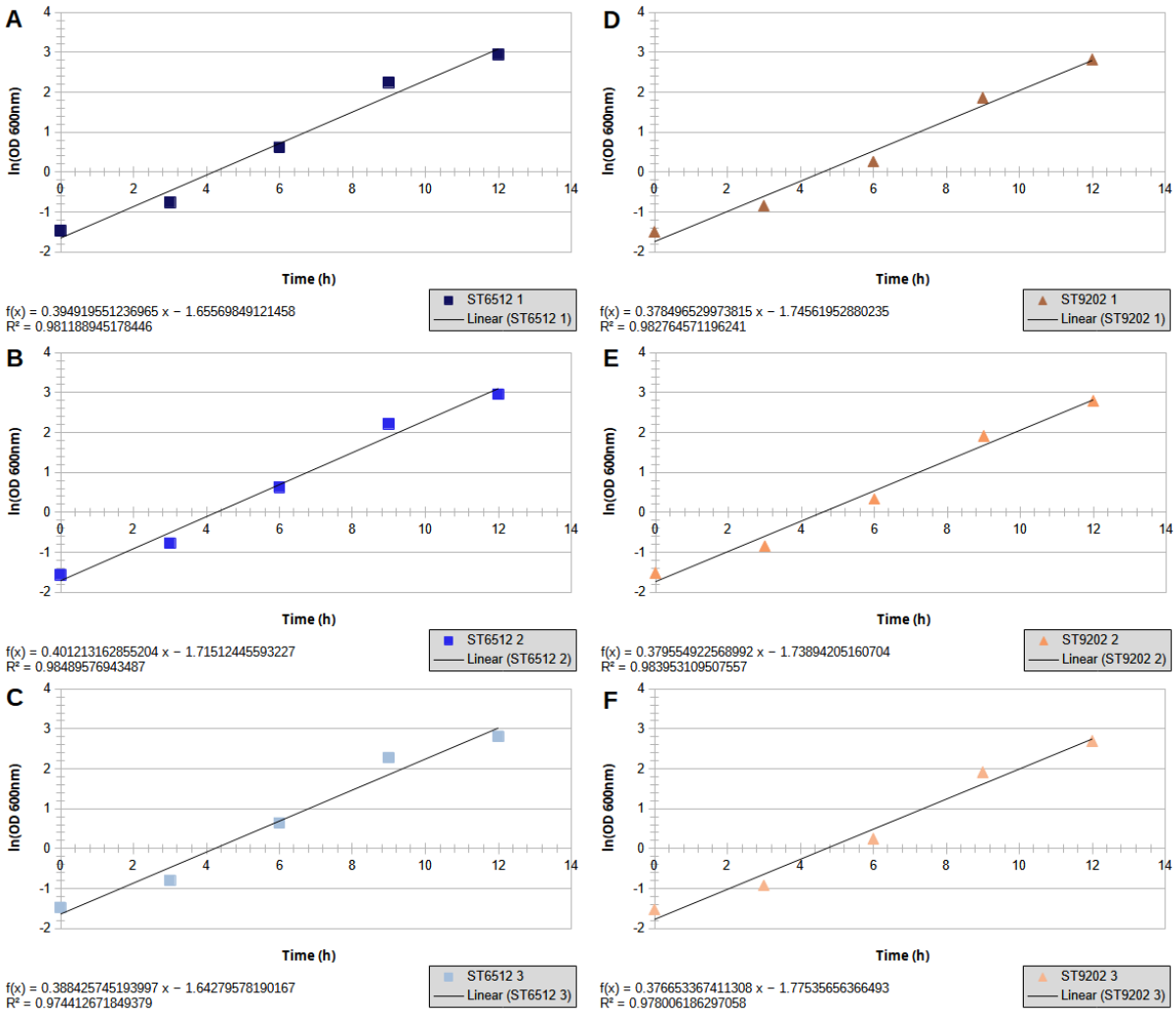
App.Figure 2: Linear regression between OD_{600} and number of ST9202 (the monoterpene production chassis strain) cells per mL of culture in YPD. Gap between OD_{600} 30 and 80 was due to sampling times, but was kept to obtain the linear regression equation. Red dots were determined as outliers and were not used for the fit of the linear regression model to achieve $R^2 > 0.965$.



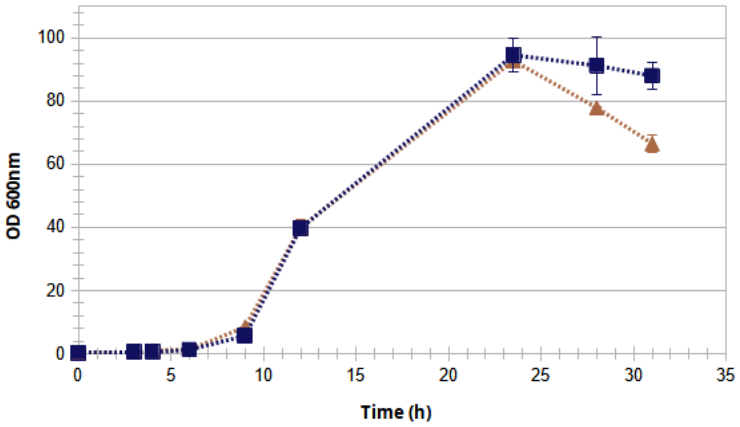
App.Figure 3: Petri dishes with YPD-agar medium obtained during isolation of pure cultures of *Y. lipolytica* ST6512 (the Cas9 expressing strain used for construction of the ST9202 strain) and ST9202 (the monoterpene production chassis strain) strains after acquisistion. Plates were incubated at 28º C for 48 h. Negative controls without inocula were performed (not shown), and no colonies were observed. (A) First *Y. lipolytica* ST6512 strain plate, after 48 h incubation. (B) First ST9202 strain plate, after 48 h incubation. (C) Second (re-streak) ST6512 strain plate, after 48 h incubation. (D) Second (re-streak) ST9202 strain plate, after 48 h incubation. (E & F) Second plates of ST6512 and ST9202 strains, respectively (same as C & D), after 48 h incubation and 72 h of storage in a refrigerator. Source: Author.



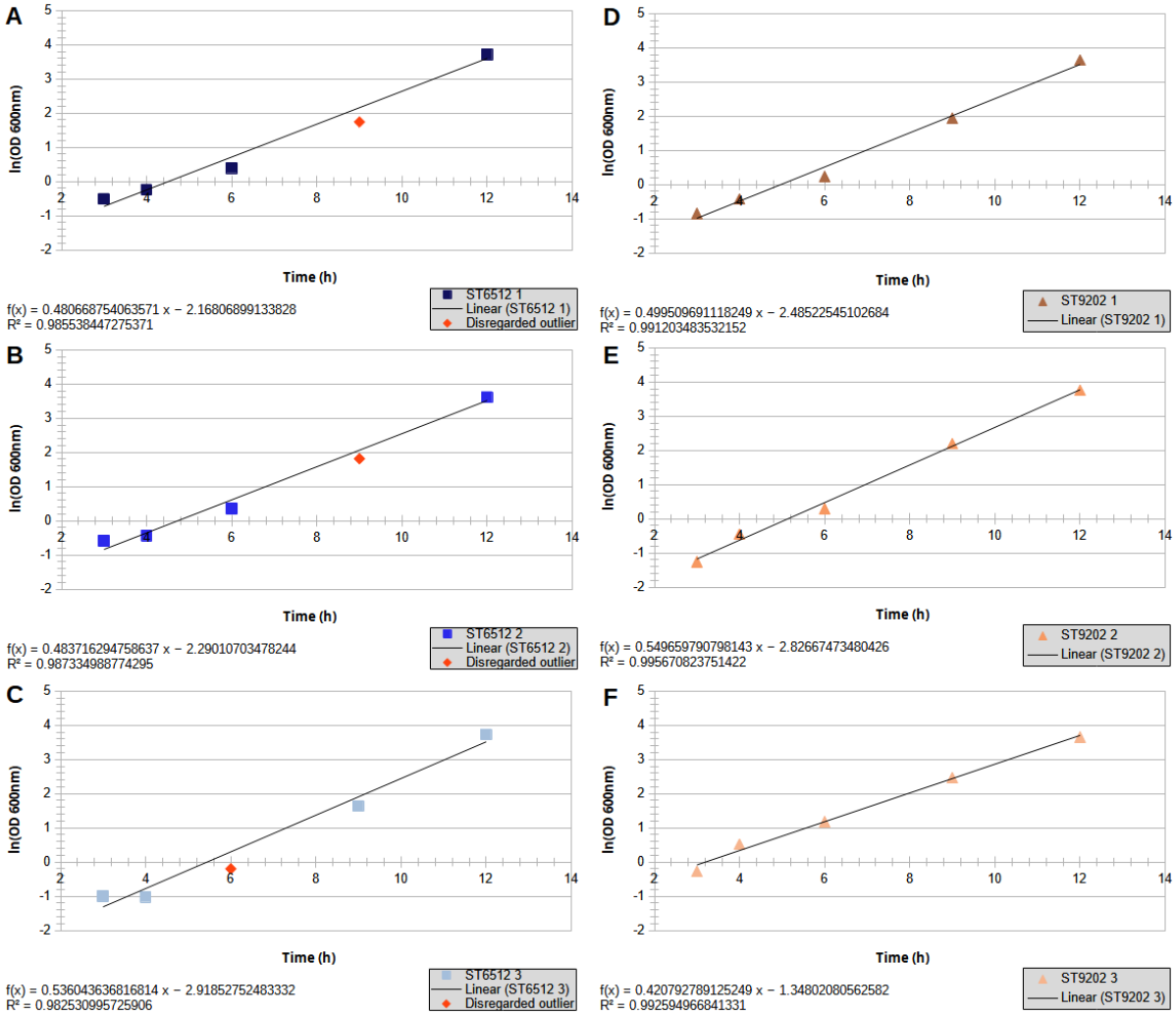
App.Figure 4: Growth curves of ST6512 (the Cas9 expressing strain used for construction of the ST9202 strain) and ST9202 (the monoterpene production chassis strain) strains obtained with data from growth kinetics experiment "A" in YPD medium and common flasks, at 30º C and 200 rpm. (—■—) ST6512 strain. (—▲—) ST9202 strain. Graph in OD_{600} measurements versus time (in h). Error bars represent standard deviation of triplicates.



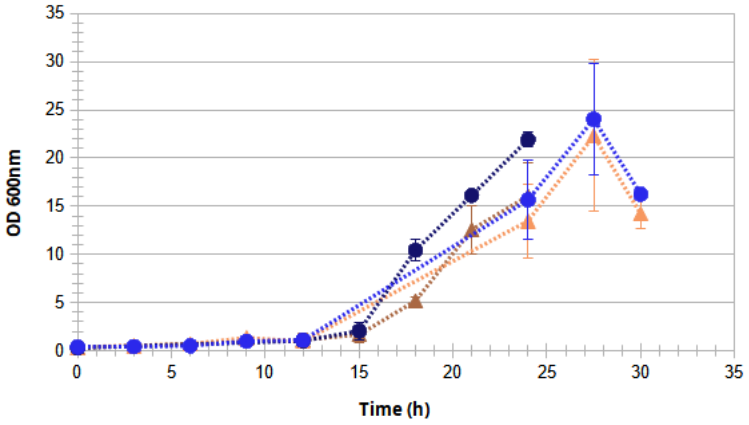
App.Figure 5: Linear regression models of the growth of ST6512 (the Cas9 expressing strain used for construction of the ST9202 strain) and ST9202 (the monoterpene production chassis strain) strains in experiment “A”, in YPD and common shake flasks, at 30° C and 200 rpm. Graphs show ln(OD) against time (in h). (A) Representation of exponential growth of replicate number 1 of ST6512 strain; (B) Representation of exponential growth of replicate number 2 of ST6512; (C) Representation of exponential growth of replicate number 3 of ST6512; (D) Representation of exponential growth of replicate number 1 of ST9202 strain; (E) Representation of exponential growth of replicate number 2 of ST9202; (F) Representation of exponential growth of replicate number 3 of ST9202. Experiment run in triplicate for each strain.



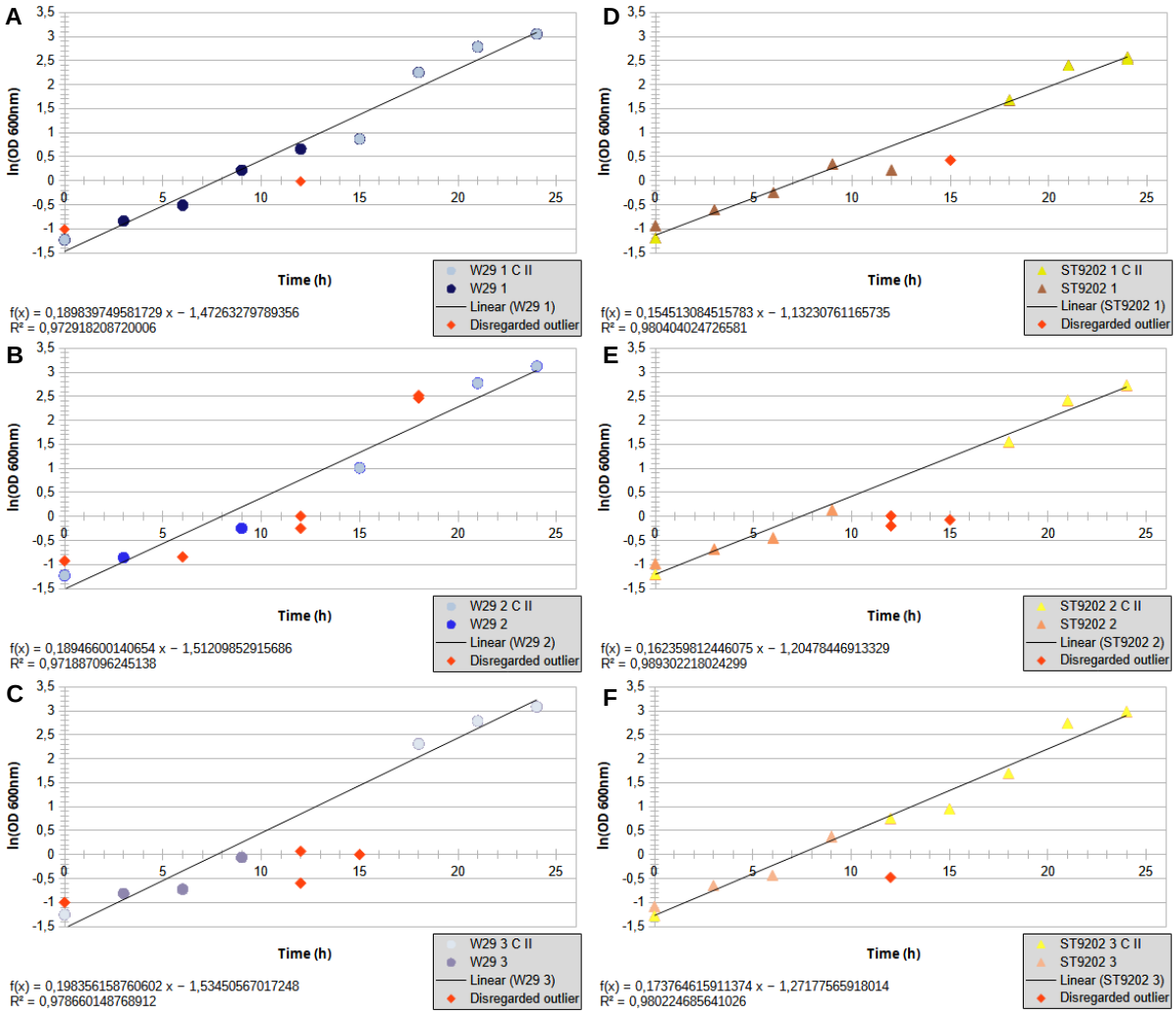
App.Figure 6: Growth curves of ST6512 (the Cas9 expressing strain used for construction of the ST9202 strain) and ST9202 (the monoterpene production chassis strain) strains obtained with data from growth kinetics experiment "B" in YPD medium and baffled flasks, at 28º C and 250 rpm. (—■—) ST6512 strain. (—▲—) ST9202 strain. Graph in OD_{600} measurements versus time (in h). Error bars represent standard deviation of triplicates.



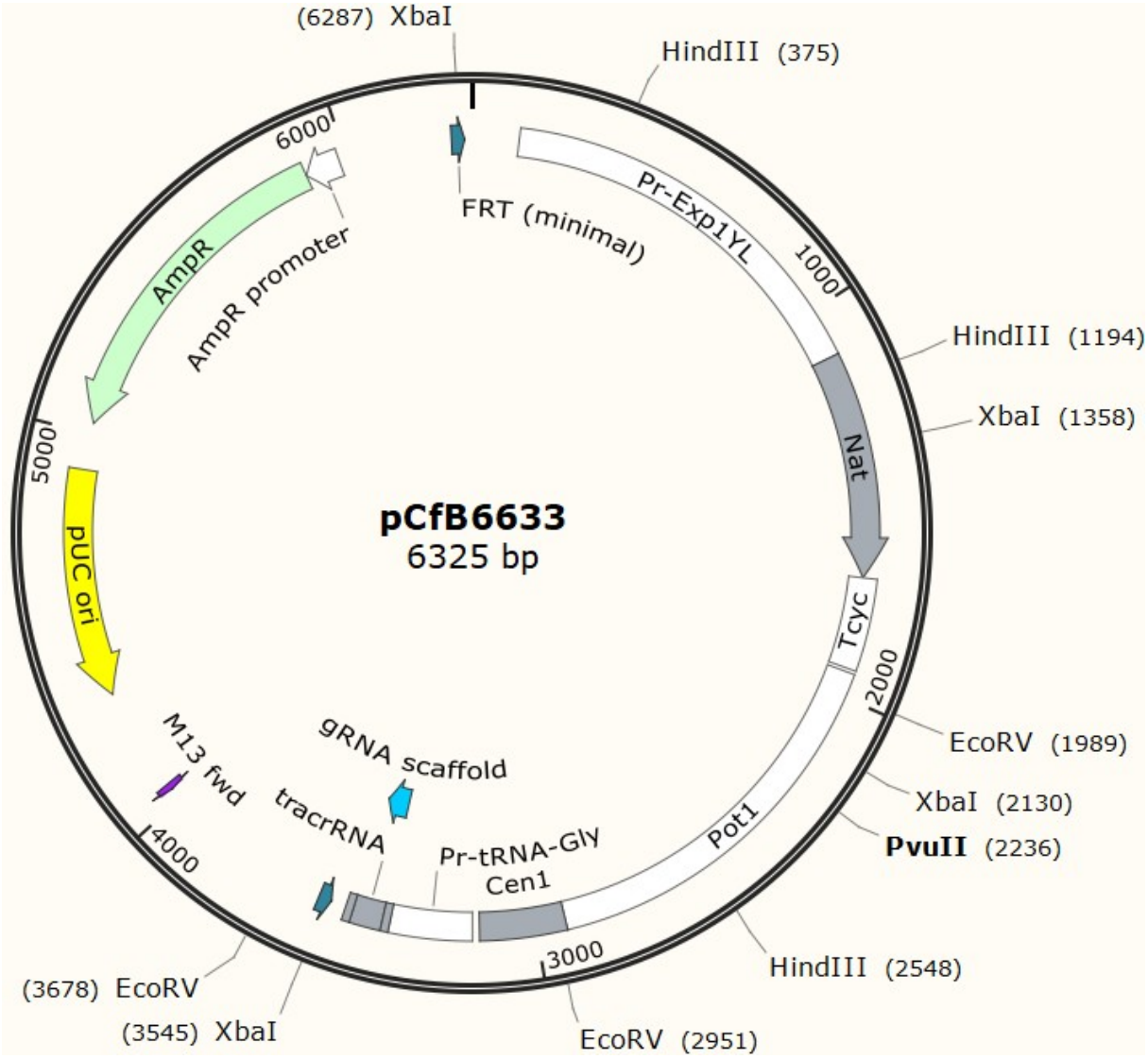
App.Figure 7: Linear regression models of the growth of ST6512 (the Cas9 expressing strain used for construction of the ST9202 strain) and ST9202 (the monoterpene production chassis strain) strains in experiment “B”, in YPD medium and baffled flasks, at 28º C and 250 rpm. Graphs show ln(OD) against time (in h). (A) Representation of exponential growth of replicate number 1 of ST6512 strain; (B) Representation of exponential growth of replicate number 2 of ST6512; (C) Representation of exponential growth of replicate number 3 of ST6512; (D) Representation of exponential growth of replicate number 1 of ST9202 strain; (E) Representation of exponential growth of replicate number 2 of ST9202; (F) Representation of exponential growth of replicate number 3 of ST9202. Experiment run in triplicate for each strain. Red dots were determined as outliers and were not used for the fit of the linear regression model to achieve $R^2 > 0.965$.



App.Figure 8: Growth curves of W29 (the wild type strain) and ST9202 (the monoterpene production chassis strain) strains obtained with data from growth kinetics experiment "C" in BDM-D medium and baffled flasks, at 28º C and 250 rpm. (—●—) W29 strain in "CI". (—●—) W29 strain in "CII". (—▲—) ST9202 strain in "CI". (—▲—) ST9202 strain in "CII". Graph in OD_{600} measurements versus time (in h). Error bars represent standard deviation of triplicates.



App.Figure 9: Linear regression models of the growth of W29 (the wild type strain) and ST9202 (the monoterpene production chassis strain) strains in experiments “C”, in BDM-D medium and baffled flasks, at 28º C and 250 rpm. Graphs show ln(OD) against time (in h). (A) Representation of exponential growth of replicates number 1 of W29 strain; (B) Representation of exponential growth of replicates number 2 of W29; (C) Representation of exponential growth of replicates number 3 of W29; (D) Representation of exponential growth of replicates number 1 of ST9202 strain; (E) Representation of exponential growth of replicates number 2 of ST9202; (F) Representation of exponential growth of replicates number 3 of ST9202. Experiments run in triplicate for each strain. Linear model points in different colors in same graph represent the two different experiments “CI” and “CII”, with same conditions but different sampling times. Red dots were determined as outliers and were not used for the fit of the linear regression model to achieve $R^2 > 0.965$.

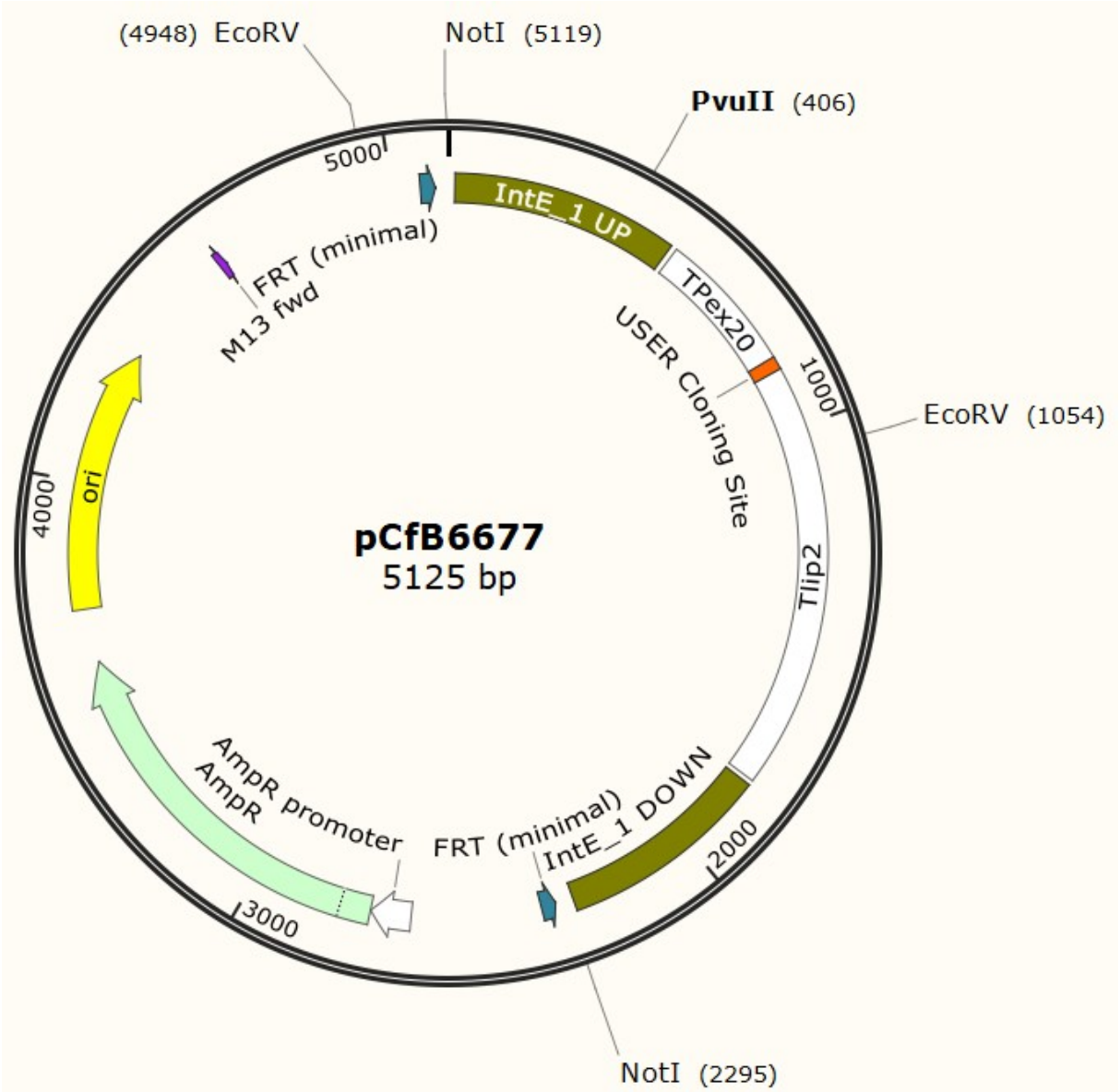


App.Figure 10: Representation of the pCfB6633 episomal plasmid used for IntE_1 sgRNA expression. Some of its features, including the Nouseothricin resistance gene (*NAT*) and the sequence for IntE_1 guide RNA, and relevant restriction sites, are shown.

App. Sequence 1: DNA sequence of the pCfB6633 episomal used for IntE_1 sgRNA expression. Source: Addgene (Holkenbrink et al. 2018). *NAT* (nourseothricin resistance) gene sequence is shown in upper case letters. The Cen1 sequence for episomal plasmid replication is shown in italic and underlined. The specific gRNA sequence for the IntE_1 region is shown in light purple highlight.

gcggccgctatgtctgataaaaggatgtaacataggcaagctgctcgtgagtgttgagtacgaaccttagatccaaatcacccgcacccacggatatacttgcttgaa
tatacagtagtataaggagtttggcgcccgttttttcgagccccacacgtttcggtgagtatgagcggcggcagattcgagcgtttccggtttccgcggctggacgag
agcccatgatgggggctcccaccaccagcaatcagggccctgattacacacccacctgtaatgtcatgctgttcatcgtggttaatgctgctgtgtgctgtgtgtgtgt
gttgtttggcgctcattgttgcgttatgcagcgtacaccacaatattggaagcttattagcctttctatttttcgtttgcaaggcttaacaacattgctgtggagagggatg
gggatatggaggccgctggagggagtcggagaggcgttttggagcggcttggcctggcgcccagctcgcgaaacgcacctaggacccttttggcacgccgaaat
gtgccacttttcagtctagtaacgccttacctacgtcattccatgcatgcatgtttgcgcctttttttcccttgcccttgatcgccacacagtacagtgcactgtacagtgga
ggttttgggggggtcttagatgggagctaaaagcggcctagcggtacactagtgggattgtatggagtggcatggagcctaggtggagcctgacaggacgcacg
accggctagcccgtgacagacgatgggtggctcctgttgtccaccgcgtacaaatgtttgggccaaagtcttgtcagccttgcttgcgaacctaattcccaattttgtc
acttcgcacccccattgatcgagccctaacccctgcccatcaggcaatccaattaagctcgcattgtctgccttgtttagtttggctcctgcccgtttcggcgtccacttg

cacaaacacaaacaagcattatatataaggctcgtctctccctcccaaccacactcacttttttgcccgtcttcccttgctaacacaaaagtcaagaacacaaacaacc
accccaaccccccttacacacaagacatatctacagcaATGGGTACTACTTTGGATGATACTGCTTACAGATACAGAACTTCT
GTTCCAGGTGATGCTGAAGCTATTGAAGCTTTGGATGGTTCTTTCACTACTGACACTGTTTTCAGAGT
TACTGCTACTGGTGATGGTTTCACTTTGAGAGAAGTTCCAGTTGATCCACCATTGACTAAGGTTTTTC
CAGATGATGAATCCGATGACGAATCTGATGATGGTGAAGATGGTGATCCAGATTCTAGAACTTTTGTT
GCTTATGGTGATGACGGTGACTTGGCTGGTTTTGTTGTTGTTTCTTATTCTGGTTGGAATAGAAGATT
GACCGTCGAAGATATTGAAGTTGCTCCAGAACATAGAGGTCATGGTGTTGGTAGAGCTTTGATGGGT
TTGGCTACTGAATTTGCTAGAGAAAGAGGTGCTGGTCATTTGTGGTTGGAAGTTACTAATGTTAACG
CTCCAGCTATTCATGCCTATAGAAGAATGGGTTTTACCTTGTGTGGTTTGGATACTGCATTATACGATG
GTACTGCATCAGATGGTGAACAAGCCTTGTATATGTCTATGCCATGTCCATGAtcatgtaattagttatgtcacgcttac
attcacgccctccccccacatccgctctaaccgaaaaggaaggagttagacaacctgaagtctaggtccctatttatttttttatagttatgttagtattaagaacgttattt
atatttcaaatttttctttttttttctgtacagacgcgtgtacgcatgtaacattatactgaaaaccttgcttgagaaggttttgggacgctcgagcaggcttggaggcgacgt
ggcagtggctcgcgggaggatcggcgggaaactccgggatatccgtcgagagtttacacgtgaatgggcagcgcaatccgttgacgacgatacgactggcaaa
gtagcgacgatacctgccagacaggtgacatgtgcaggccgcactaacaaggaaacgggcgctgggggggcgggcttctagactttgcccttgaacaggaat
ctagtgggggcttgtctttccgccaatgggggagcgcctgttgagcgaccgtgcatgctggaacgccaagtgtatgtacagctggtggtctcgcagcggtatgtga
cgggacttacatctctcgttttttcatgaccacgttttcacaggctcggaggtacgttaaagttttgaaggctgcatctgaaccgaggtatgggggagtttgaagagcaa
cagtgttggggctgaggggccaagatcggggcaagcagaggtcttagatcaattgtggggatcccaaagggctcgttatcacctttttccacccaattcgggtccc
aattgatccactactggcttgcccaagttaccccagaaatgccgccccggatttctccaaaaacctaataagcttcatggaacttggtggaagtgactttctacagagt
ggagagaaccgtggacacgtggcaatggcgctgaccgtgtccccgagccgaatcgacgtgaggggagaacggagtatctgcggtcatgtgaccttccagagcg
gcgtcgccagtgtgcacgcggtgacccccagtttgggtctctgtcacacgcatactacctcggctctccacatgctgaactttatctttcgtggggatcataccgaaag
ttgcaactaccaggtgtatataaagcctggtagactccccccactttggacctcatccaaccaagacacacaaaaatggaccgacttaacaacctcgccacccagct
cgagcagaaccccgccaa*ttcgtagataatggaatacaaatggatatccagagtatacacatggatagtatacactgacacgacaattctgtatctctttatgttaa*
actactgtgaggcattaaatagagcttgatatataaaatgttacatttcacagtctgaacttttgcagattacctaatttggtaagatattaattatgaactgaaagtt
*gatggcatccctaaatttgatg*caacggaatgcgtgcgatagtgaatcattgctaacagatcaagtaatcgtatcaatcgttccacgtcttatccagactcgatttccgc
cctaacctaaggaaactgctgagtaatttcgaggttggttcctaacaaaagccaagaaacgcgccgatggtttagtggtaaaatccatcgttgccatcgatgggcccc
cggttcgattccgggtcggcgcaggttggttaacaagcatacagccctcggggttttagagctagaaatagcaagttaaaataaggctagtccgttatcaacttgaaa
aagtggcaccgagtcggtggtgcttttttgtgtgggtacggtatcgcgtgcattcgcggccgcatttaaatccccacttcagaagttcctatactttctagagaataggа
acttctatagtgagtcgaataagggcgacacaaaatttattctaaatgcataataaatactgataacatcttatagtttgtattatattttgtattatcgttgacatgtataatttt
gatatcaaaaactgattttccctttattattttcgagatttattttcttaattctctttaacaaactagaaatattgtatatacaaaaaatcataaataatagatgaatagtttaatta
taggtgttcatcaatcgaaaaagcaacgtatcttatttaaagtgcgttgctttttctcatttataaggttaaataattctcatatatcaagcaaagtgacaggcgcccttaaa
tattctgacaaatgctctttccctaaactcccccatataaaaaacccgccgaagcgggtttttacgttatttgcggattaacgattactcgttatcagaaccgcccaggg
ggcccgagcttaagactggccgtcgttttacaacacagaaagagtttgtagaaacgcaaaaaggccatccgtcaggggccttctgcttagtttgatgcctggcagttc
cctactctcgccttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaatacggttatccacaga
atcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgttttccataggctccgcccc
cctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagctccctcgtgcgctct
cctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcggtgtaggtcgtt
cgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaagacacgacttatcgcc
actggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtgggctaactacggctacactagaagaacagt
atttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggttttttgtttgcaagc
agcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgacgcgcgcgtaactcacgttaagggatttt
ggtcatgagcttgcgccgtcccgtcaagtcagcgtaatgctctgcttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgc
ctgactccccgtcgtgtagataactacgatacgggagggcttaccatctggccccagcgctgcgatgataccgcgagaaccacgctcaccggctccggatttatca
gcaataaaccagccagccggaagggccgagcgcagaagtggtcctgcaactttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttc
gccagttaatagtttgcgcaacgttgttgccatcgctacaggcatcgtggtgtcacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcga
gttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccgatcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcact
gcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgccc
ggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgag
atccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttactttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaa
gggaataagggcgacacggaaatgttgaatactcatattcttccttttttcaatattattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtattta
gaaaaataaacaaatagggtcagtgttacaaccaattaaccaattctgaacattatcgcgagcccatttatacctgaatatggctcataacacccttgtttgcctggc
ggcagtagcgcggtggtcccacctgaccccatgccgaactcagaagtgaaacgccgtagcgccgatggtagtgtggggactccccatgcgagagtagggaact
gccaggcatcaaataaaacgaaaggctcagtcgaaagactgggcctttcgcccgggctaattatggggtgtcgcccttattcgactctatagtgaagttcctattctct
agaaagtataggaacttctgaagtggggatttaaat

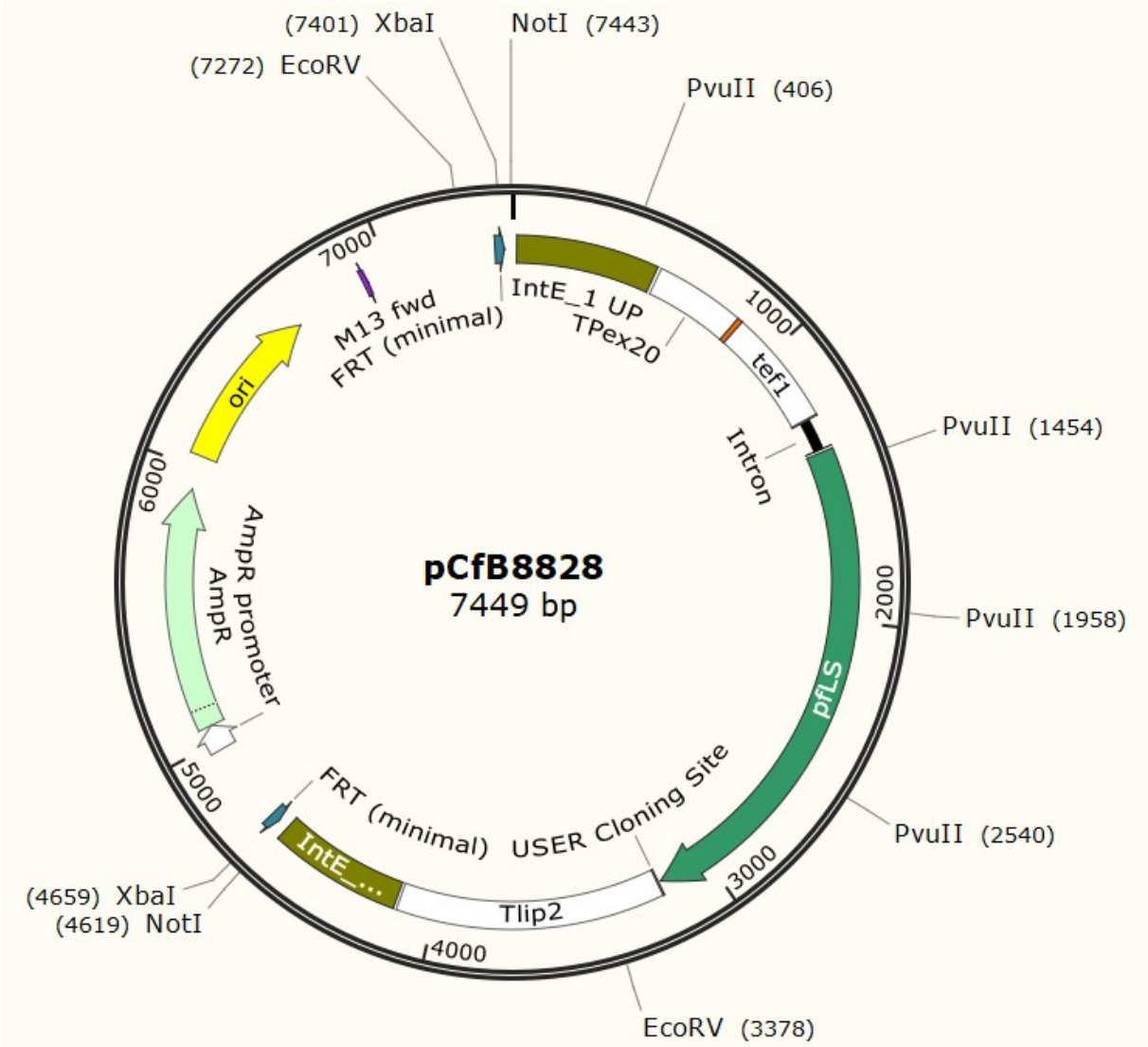


App.Figure 11: Representation of the pCfB6677 plasmid, the backbone vector for expression cassette construction. Some of its features, including the IntE_1 homology arms and the USER cloning site for cloning of promoters and genes, and relevant restriction sites, are shown

App. Sequence 2: DNA sequence of the pCfB6677 plasmid, the backbone vector for expression cassette construction. Source: Addgene (HOLKENBRINK et al. 2018). The TPex20 terminator sequence is shown in black highlight. The Tlip2 terminator sequence is shown in light gray highlight. The IntE_1 Upstream homology arm is shown in italic and underlined. The IntE_1 Downstream arm is shown in italic and in bold.

gaatgcgtgcgatttgatatggtgtaacaatgataaaccaaggcccattgattgagacggagatattttcggcgtgtctagggtgtctacttgtgcaccggaacatggaagatcgtgcgctgttcacttgcaacagaaggctaacgagagaaactgggccaacccgtctatgctcttcatgctactcaatctattcaaaatacgccagtgctccccttgccttatacacaggctcttcactcacatcttgaccatcttggatgtgggtactgtagctcctgggaaccagtcgtgcgggagcatgagaaccacgtcgaagaagatggtgacccaggaaagcaaaaacttgggcacattcttgtagataccctgacctgtcgtggtcgaagataccgtcgaaaaccagctgtcccattgacagaaaccctcctcctgtatctagcagaacgctggcaatagaaaagccattagtggactttcgcgtgtagttcgacgcagcagaggatagtgcttacgtgcaacgcttacgcaactaacatgaatgaatacgatatacatcaaagactatgatacgcagtattgcacactgtacgagtaagggcactagccactgcactcaagtgaaaccgt

tgcccgggtacgagtatgagtatgtacagtatgtttagtattgtacttggacaatgcttgtatcgtacattctcaagtgtcaaacataaatatccgttgctatatcctcgcac
caccacgtagctcgctatatccctgtgttgaatccatccatcttggattgccaattgtgcacacagaaccgggcactcacttccccatccacacttcaacggaatgcgt
gcgatcgcgtgcattccttctgttcggaatcaacctcaaggttaacggccacgatcccctcgttgttactcttggtcagcccattgtcggtaacgctggctttgctaactg
ggtcgataaactcttctttggccaggagaaccccgatgtctccaaggtgtccaaagaccgaaagctctaccgaatcacccaccgaggagatatcgtccctcaagtg
cccttctgggacggttaccagcactgctctggtgaggtctttattgactggcccctgatccaccctcctctctccaacgttgtcatgtgccagggccagagcaataaac
agtgctctgccggtaacactctgctccagcaggtcaatgtgattggaaaccatctgcagtacttcgtcaccgagggtgtctgtggtatctaagctatttatcactctttac
aacttctacctcaactatctactttaataaatgaatatcgtttattctctatgattactgtatatgcgttcctctaagacaaatcgaaaccagcatgcgatcgaatggcataca
aaagtttcttccgaagttgatcaatgtcctgatagtcaggcagcttgagaagattgacacaggtggaggccgtagggaaccgatcaacctgtctaccagcgttacga
atggcaaatgacgggttcaaagccttgaatccttgcaatggtgccttggatactgatgtcacaaacttaagaagcagccgcttgtcctcttcctcgaaactctcaaaca
cagtccagaagtcctttatagtttgatctgtatccagatagcctccgtaattggtgtgtgtcttcaaatcccagacgtccacattggcatgtcctccactgataagcatttg
aagttcatctgcgttgaacattgagacccacgaagggtcaatgagctggtatagaccgcccaagaatgcatctgagtggcctttg***tgtcgaaatacaacagccagt***
cctccacgatccagctcagatagccctgggtgttgagatagagacaagtcgagcgagatgttctcagtggacttgagtttgatgttgtgaaagatggttggattg
agagcacaactccatagaacatagtagtctcagtccatgagctaggagagaatctccagtgctgatgtcatttgaagaagaagaggaagaagaagacagat
agatcaactcagccccttccagttgccgtacaatgactcaaccaaaatacgttcaagcctgctttcaagcctgcgtgctctttctgtctcattagttgaacgtagatca
tctgagagaacttggcagaccgtgcaaacacaagggcgagactgataggacacttgatctgatccaaaatttgacagaggaaaactagttttgaagttactg
aacgtttcttctccaattgcttctaacatcgcgtgcattcgcggccgcatttaaatccccacttcagaagttcctatactttctagagaataggaacttcactatagagtcg
aataagggcgacaccccataattagcccgggcgaaaggcccagtctttcgactgagcctttcgttttatttgatgcctggcagttccctactctcgcatggggagtccc
cacactaccatcggcgctacggcgtttcacttctgagttcggcatggggtcaggtgggaccaccgcgctactgccgccaggcaaacaaggggtgttatgagccat
attcaggtataaatgggctcgcgataatgttcagaattggttaattggttgtaacactgaccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagaca
ataaccctgataaatgcttcaataatattgaaaaaggaagaatatgagtattcaacatttccgtgtcgcccttattcccttttttgcggcattttgccttcctgtttttgctcacc
cagaaacgctggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgcccc
gaagaacgttttccaatgatgagcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctca
gaatgacttggttgagtactcaccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcgg
ccaacttacttctgacaacgatcggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaat
gaagccataccaaacgacgagcgtgacaccacgatgcctgtagcgatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaac
aattaatagactggatggaggcggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatccggagccggtgagcgtggtt
ctcgcggtatcatcgcagcgctggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagat
cgctgagataggtgcctcactgattaagcattggtaagcagagcattacgctgacttgacgggacggcgcaagctcatgaccaaaatcccttaacgtgagttacgcg
cgcgtcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgcta
ccagcggtggtttgtttgccggatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagtt
agcccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggac
tcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacc
tacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgaggga
gcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaa
aacgccagcaacgcggcctttttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgag
tgagctgataccgctcgccgcagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaggcgagagtagggaactgccaggcatcaaactaag
cagaaggcccctgacggatggcctttttgcgtttctacaaactctttctgtgttgtaaaacgacggccagtcttaagctcgggcccccctgggcggttctgataacgagt
aatcgttaatccgcaaataacgtaaaaacccgcttcggcgggtttttttatgggggggagtttagggaaagagcatttgtcagaatatttaagggcgcctgtcactttgctt
gatatatgagaattatttaaccttataaatgagaaaaaagcaacgcactttaaataagatacgttgctttttcgattgatgaacacctataattaaactattcatctattatttat
gatttttttgtatatacaatatttctagtttgttaaagagaattaagaaaataaatctcgaaaatataaagggaaaatcagtttttgatatcaaaattatacatgtcaacgataa
tacaaaatataatacaaactataagatgttatcagtatttattatgcatttagaataaatttttgtgtcgcccttattcgactcactatagaagttcctattctctagaaagtatag
gaacttcacttcatttttatttaaatttgcggccgc



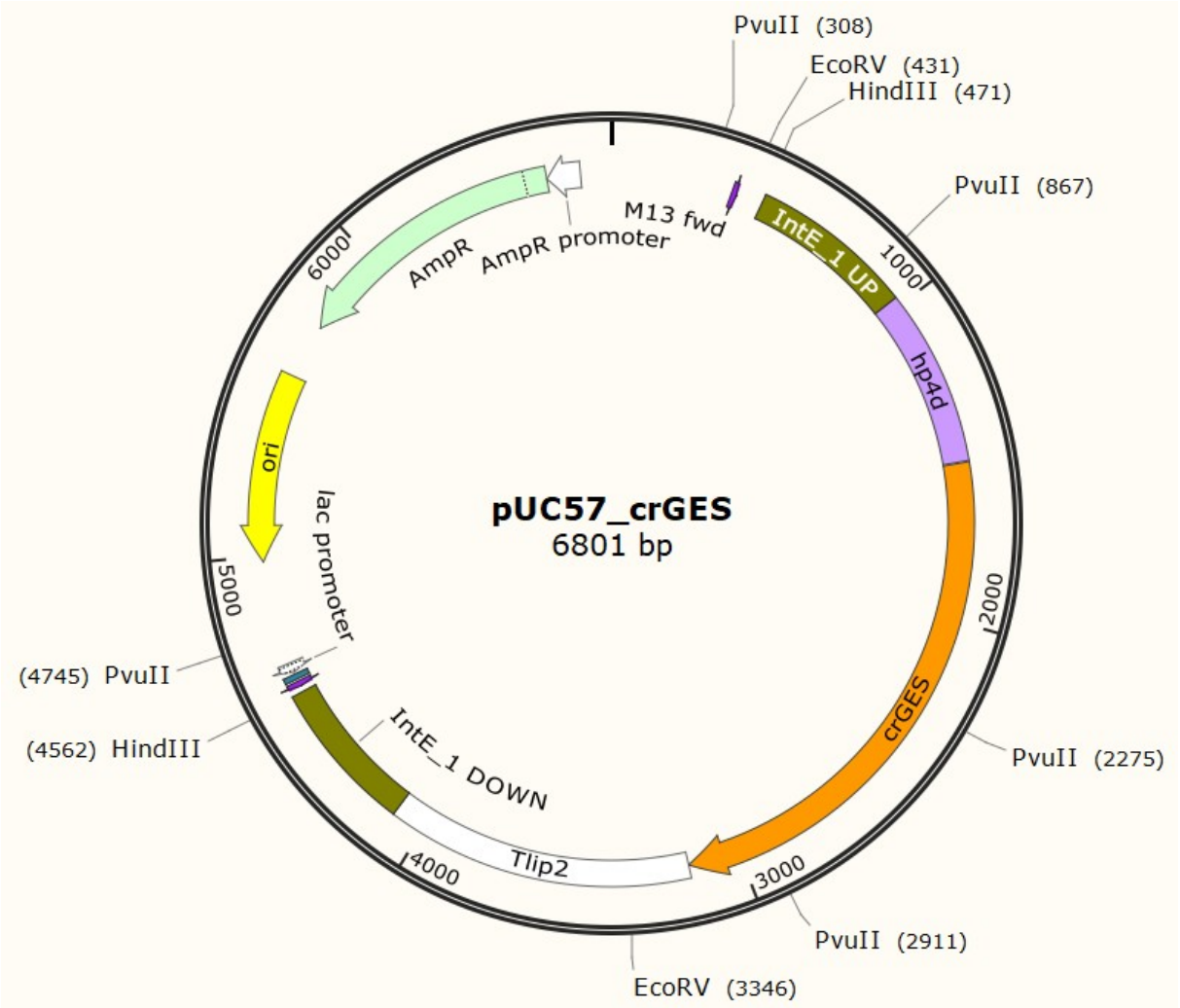
App.Figure 12: Representation of the pCfB8828 plasmid with the *pfLS* expression cassette and used as the transformation positive control. Some of its features, including the *pfLS* expression cassette flanked by the IntE_1 homology arms, and relevant restriction sites, are shown.

App. Sequence 3: DNA sequence of the pCfB8828 plasmid with the *pfLS* expression cassette and used as the transformation positive control. Source: Addgene (HOLKENBRINK et al. 2018) for pCfB6677 vector skeleton, ARNESEN et al. 2020 for *pfLS* and TEFin sequences. The TPex20 terminator sequence is shown in black highlight. The TEFin promoter is shown in light purple highlight. The *Y. lipolytica* codon-optimized *LS* gene from *P. frutescens* sequence is shown in upper case letters. The Tlip2 terminator sequence is shown in light gray highlight. The IntE_1 Upstream homology arm is shown in italic and underlined. The IntE_1 Downstream arm is shown in italic and in bold.

Gaatgcgtgcgatttgatatggtgtaacaatgataaaccaaggcccattgattgagacggagatattttcggcgtgtctagggtgtctacttgtgcaccggaacatggaagatcgtgcgctgttcacttgcaacagaaggctaacgagagaaactgggccaacccgtctatgctcttcatgctactcaatctattcaaaatacgccagtgctccccttgccttatacacaggctcttcactcacatcttgaccatcttggatgtgggtactgtagctcctgggaaccagtcgtgcgggagcatgagaaccacgtcgaagaagatggtgacccaggaaagcaaaaacttgggcacattcttgtagataccctgacctgtcgtggtcgaagataccgtcgaaaaccagctgtcccattgacagaaaccctcctcctgtatctagcagaacgctggcaatagaaaagccattagtggactttcgcgtgtagttcgacgcagcagaggatagtgcttacgtgcaacgcttacgcaactaacatgaatgaatacgatatacatcaaagactatgatacgcagtattgcacactgtacgagtaagggcactagccactgcactcaagtgaaacc

gttgcccgggtacgagtatgagtatgtacagtatgtttagtattgtacttggacaatgcttgtatcgtacattctcaagtgtcaaacataaatatccgttgctatatcctcgc
accaccacgtagctcgctatatccctgtgttgaatccatccatcttggattgccaattgtgcacacagaaccgggcactcacttccccatccacacttcaacggaatgc
gtgcgatagagaccgggttggcggcgcatttgtgtcccaaaaaacagccccaattgccccaattgaccccaaattgacccagtagcgggcccaaccccggcgag
agcccccttctccccacatatcaaacctcccccggttcccacacttgccgttaagggcgtagggtactgcagtctggaatctacgcttgttcagactttgtactagtttct
ttgtctggccatccgggtaacccatgccggacgcaaaatagactactgaaaatttttttgctttgtggttgggactttagccaagggtataaaagaccaccgtccccga
attacctttcctcttcttttctctctctccttgtcaactcacacccgaaatcgttaagcatttccttctgagtataagaatcattcaaaatggtgagtttcagaggcagcagca
attgccacgggctttgagcacacggccgggtgtggtcccattcccatcgacacaagacgccacgtcatccgaccagcacttttgcagtactaaccgcagcacAT
GGCTATCCCCATCAAGCCCGCTCACTACCTGCACAACTCTGGCCGATCTTACGCCTCTCAGCTGTGCG
GCTTCTCTTCTACCTCTACTCGAGCCGCCATTGCTCGACTGCCCCTGTGCCTGCGATTCCGATGCTCT
CTGCAGGCCTCTGACCAGCGACGATCTGGCAACTACTCTCCCTCTTTCTGGAACGCCGACTACATCC
TGTCTCTGAACTCTCACTACAAGGACAAGTCTCACATGAAGCGAGCCGGCGAGCTGATCGTGCAGG
TCAAGATGGTGATGGGCAAGGAAACTGACCCCGTCGTCCAGCTTGAGCTGATTGACGACCTGCAGA
AGCTGGCCCTGTCTCACCACGTCGAGAAGGAAATCAAGGAAATTCTGTTCAAGATCTCTACCTACGA
CCACAAGATCATGGTCGAGCGAGATCTGTACTCTACCGCTCTGGCCTTCCGACTGCTGCGACAGTAC
GGCTTCAAGGTGCCCCAAGAGGTGTTCGACTGCTTCAAGAACGACAACGGCGAGTTCAAGCGATCT
CTGTCCTCTGACACCAAGGGCCTGCTGCAGCTGTACGAGGCTTCGTTCCTGCTGACCGAGGGCGAG
ATGACCCTCGAGCTGGCCCGAGAGTTCGCCACCAAGTCTCTGCAAGAGAAGCTGAACGAAAAGAC
CATTGACGACGACGATGACGCCGACACCAACCTGATCTCTTGCGTGCGACACTCTCTGGACATCCCC
ATCCACTGGCGAATTCAGCGACCCAACGCCTCTTGGTGGATCGACGCCTACAAGCGAAGATCCCACA
TGAACCCTCTGGTGCTGGAACTGGCTAAGCTGGACCTGAACATCTTCCAGGCTCAGTTTCAGCAAG
AGCTGAAGCAGGACCTCGGCTGGTGGAAGAACACCTGTCTGGCCGAGAAGCTGCCCTTCGTGCGA
GATCGACTGGTCGAGTGCTACTTCTGGTGCACCGGCATCATTCAGCCCCTGCAGCACGAGAACGCCC
GAGTGACCCTGGCCAAGGTGAACGCCCTGATCACCACTCTGGACGACATCTACGACGTGTACGGCA
CCCTCGAGGAACTCGAGCTGTTCACCGAGGCCATCCGACGATGGGACGTGTCCTCTATCGACCATCT
GCCTAACTATATGCAGCTGTGCTTTCTGGCTCTGAACAACTTCGTGGACGACACCGCCTACGACGTC
ATGAAGGAAAAGGACATCAACATCATTCCCTACCTGCGAAAGTCTTGGCTGGACCTCGCCGAGACTT
ACCTGGTCGAGGCCAAGTGGTTCTACTCTGGACACAAGCCCAACCTGGAAGAGTACCTGAACAACG
CCTGGATCTCTATCTCTGGCCCCGTGATGCTGTGCCACGTGTTCTTTCGAGTGACCGACTCTATCACC
CGAGAGACTGTCGAGTCCCTGTTCAAGTACCACGACCTGATCCGATACTCCTCTACCATCCTGCGAC
TGGCCGACGACCTGGGCACCTCTCTGGAAGAGGTGTCCCGAGGCGACGTGCCCAAGTCTATCCAGT
GCTACATGAACGATAACAACGCTTCTGAGGAAGAGGCCCGACGACACATCCGATGGCTGATCGCTG
AGACTTGGAAGAAGATCAACGAAGAGGTCTGGTCTGTGGACTCGCCCTTCTGCAAGGACTTCATTG
CCTGTGCCGCCGACATGGGCCGAATGGCCCAGTTCATGTACCACAACGGTGACGGCCACGGCATTC
AGAACCCTCAGATCCACCAGCAGATGACCGACATCCTGTTCGAGCAGTGGCTGTAAatcgcgtgcattccttct
gttcggaatcaacctcaaggttaacggccacgatcccctcgttgttactcttggtcagcccattgtcggtaacgctggctttgctaactgggtcgataaactcttctttgg
ccaggagaaccccgatgtctccaaggtgtccaaagaccgaaagctctaccgaatcacccaccgaggagatatcgtccctcaagtgcccttctgggacggttacca
gcactgctctggtgaggtctttattgactggcccctgatccaccctcctctctccaacgttgtcatgtgccagggccagagcaataaacagtgctctgccggtaacact
ctgctccagcaggtcaatgtgattggaaaccatctgcagtacttcgtcaccgagggtgtctgtggtatctaagctatttatcactctttacaacttctacctcaactatctac
tttaataaatgaatatcgtttattctctatgattactgtatatgcgttcctctaagacaaatcgaaaccagcatgcgatcgaatggcatacaaaagtttcttccgaagttgatc
aatgtcctgatagtcaggcagcttgagaagattgacacaggtggaggccgtagggaaccgatcaacctgtctaccagcgttacgaatggcaaatgacgggttcaaa
gccttgaatccttgcaatggtgccttggatactgatgtcacaaacttaagaagcagccgcttgtcctcttcctcgaaactctcaaacacagtccagaagtcctttatagtt
tgatctgtatccagatagcctccgtaattggtgtgtgtcttcaaatcccagacgtccacattggcatgtcctccactgataagcatttgaagttcatctgcgttgaacattg
agacccacgaagggtcaatgagctggtatagaccgcccaagaatgcatctgagtggcctttg***tgtcgaaatacaacagccagtcctccacgatccagctcagat***
agccctgggtgttgagatagagacaagtcgagcgagatgttctcagtggacttgagtttgatgttgtgaaagatggttggattgagagcacaactccatagaac
atagtagtctcagtccatgagctaggagagaatctccagtgctgatgtcatttgaagaagaagaggaagaagaagacagatagatcaactcagccccttcca
gttgccgtacaatgactcaaccaaatacgttcaagcctgctttcaagcctgcgtgctctttctgtctcattagttgaacgtagatcatctgagagaacttggcagac
cgtgcaaacacaagggcgagactgataggacacttgatctgatccaaaatttgacagaggaaaactagttttgaagttactgaacgtttcttctccaattgcttct
aacatcgcgtgcattcgcggccgcatttaaatccccacttcagaagttcctatactttctagagaataggaacttcactatagagtcgaataagggcgacaccccataa
ttagcccgggcgaaaggcccagtctttcgactgagcctttcgttttatttgatgcctggcagttccctactctcgcatggggagtccccacactaccatcggcgctacg
gcgtttcacttctgagttcggcatggggtcaggtgggaccaccgcgctactgccgccaggcaaacaaggggtgttatgagccatattcaggtataaatgggctcgc
gataatgttcagaattggttaattggttgtaacactgacccctatttgtttatttttctaaatacattcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaat
aatattgaaaaaggaagaatatgagtattcaacatttccgtgtcgcccttattccctttttgcggcattttgccttcctgtttttgctcacccagaaacgctggtgaaagtaa
aagatgctgaagatcagttgggtgcacgagtgggttacatcgaactggatctcaacagcggtaagatccttgagagttttcgccccgaagaacgttttccaatgatga

gcacttttaaagttctgctatgtggcgcggtattatcccgtattgacgccgggcaagagcaactcggtcgccgcatacactattctcagaatgacttggttgagtactca
ccagtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgctgccataaccatgagtgataacactgcggccaacttacttctgacaacgat
cggaggaccgaaggagctaaccgcttttttgcacaacatgggggatcatgtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacga
gcgtgacaccacgatgcctgtagcgatggcaacaacgttgcgcaaactattaactggcgaactacttactctagcttcccggcaacaattaatagactggatggagg
cggataaagttgcaggaccacttctgcgctcggcccttccggctggctggtttattgctgataaatccggagccggtgagcgtggttctcgcggtatcatcgcagcg
ctggggccagatggtaagccctcccgtatcgtagttatctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgctgagataggtgcctcac
tgattaagcattggtaagcagagcattacgctgacttgacgggacggcgcaagctcatgaccaaaatcccttaacgtgagttacgcgcgcgtcgttccactgagcgt
cagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccg
gatcaagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttagcccaccacttcaagaact
ctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccgga
taaggcgcagcggtcgggctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaa
agcgccacgcttcccgaagggagaaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcct
ggtatctttatagtcctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggcctt
tttacggttcctggccttttgctggccttttgctcacatgttctttcctgcgttatcccctgattctgtggataaccgtattaccgcctttgagtgagctgataccgctcgccg
cagccgaacgaccgagcgcagcgagtcagtgagcgaggaagcggaaggcgagagtagggaactgccaggcatcaaactaagcagaaggcccctgacggat
ggccttttgcgtttctacaaactctttctgtgttgtaaaacgacggccagtcttaagctcgggcccccctgggcggttctgataacgagtaatcgttaatccgcaaataac
gtaaaaacccgcttcggcgggtttttttatgggggggagtttagggaaagagcatttgtcagaatatttaagggcgcctgtcactttgcttgatatatgagaattatttaacc
ttataaatgagaaaaaagcaacgcactttaaataagatacgttgcttttcgattgatgaacacctataattaaactattcatctattatttatgatttttgtatatacaatatttc
tagtttgttaaagagaattaagaaaataaatctcgaaaataataaagggaaaatcagtttttgatatcaaaattatacatgtcaacgataatacaaaatataatacaaacta
taagatgttatcagtatttattatgcatttagaataaattttgtgtcgcccttattcgactcactatagaagttcctattctctagaaagtataggaacttcacttcattttatttaa
atttgcggccgc



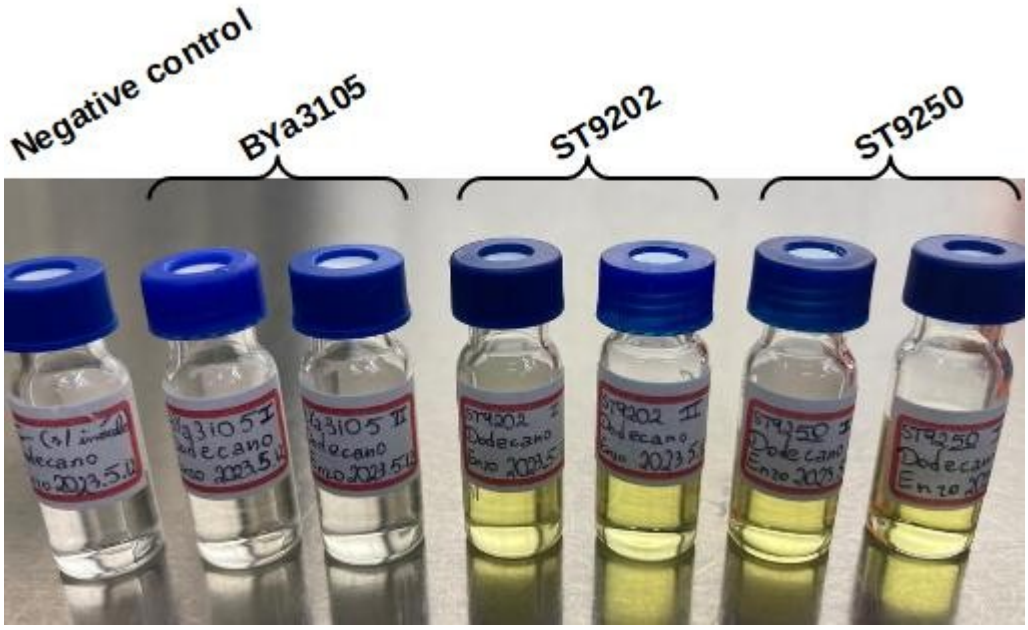
App.Figure 13: Representation of the pUC57_ crGES plasmid with the *crGES* expression cassette. Some of its features, including the *crGES* expression cassette flanked by the IntE_1 homology arms, and relevant restriction sites, are shown.

App. Sequence 4: DNA sequence of the pUC_crGES plasmid with the *crGES* expression cassette. Source: this work. The hp4d promoter sequence is shown in light purple highlight. The *Y. lipolytica* codon-optimized and truncated *GES* gene from *C. roseus* sequence is shown in upper case letters. The Tlip2 terminator sequence is shown in light gray highlight. The IntE_1 Upstream homology arm is shown in italic and underlined. The IntE_1 Downstream homology arm is shown in italic and in bold.

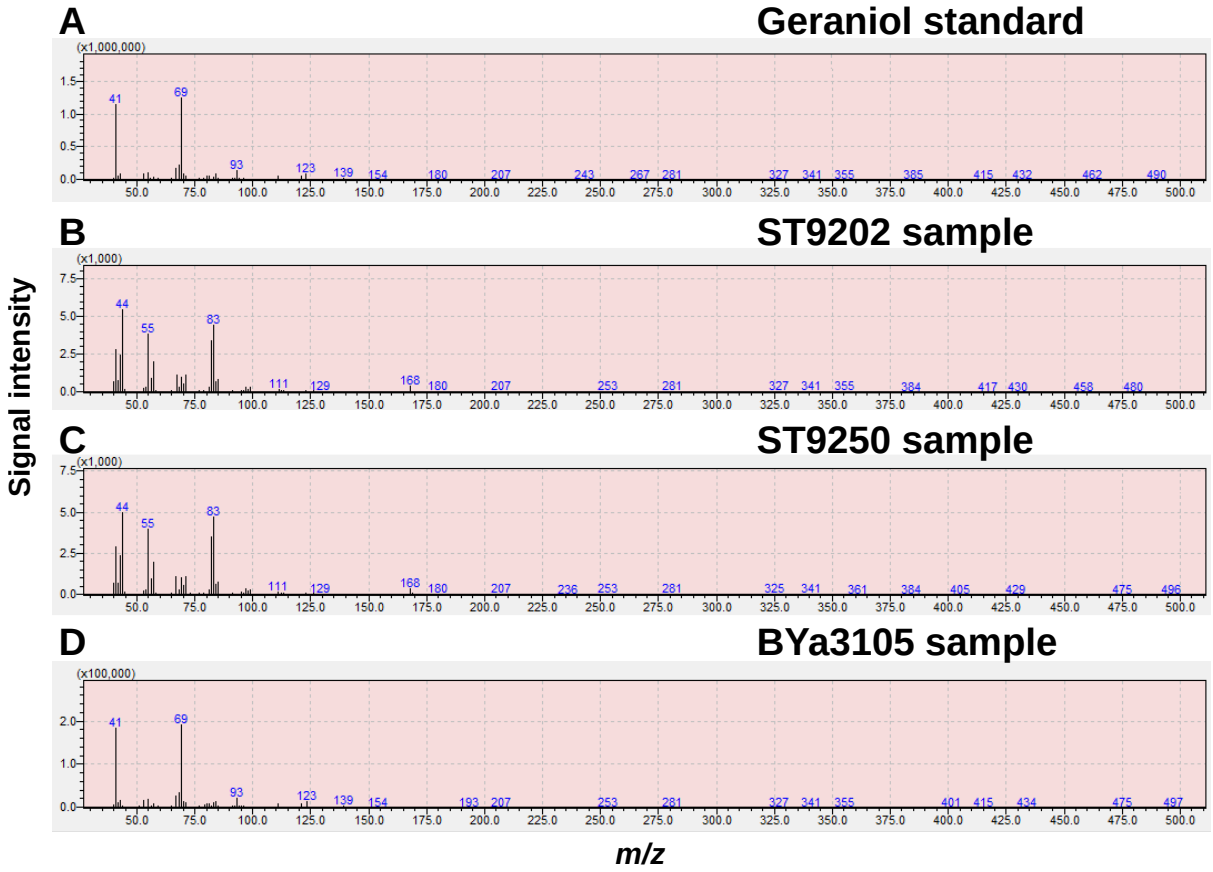
Tcgcgcgtttcggtgatgacggtgaaaacctctgacacatgcagctcccggagacggtcacagcttgtctgtaagcggatgccgggagcagacaagcccgtcag
ggcgcgtcagcgggtgttggcgggtgtcggggctggcttaactatgcggcatcagagcagattgtactgagagtgcaccatatgcggtgtgaaataccgcacaga
tgcgtaaggagaaaataccgcatcaggcgccattcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctattacgccagctggcg
aaaggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttgtaaaacgacggccagtgaattcgagctcggtacctcgcgaa
tgcatctagatatcggatcccgggcccgtcgactgcagaggcctgcatgcaagctt*gatatggtgtaacaatgataaaccaaggcccattgattgagacggaga*
tattttcggcgtgtctagggtgtctacttgtgcaccggaacatggaagatcgtgcgctgttcacttgcaacagaaggctaacgagagaaactgggccaacccgt
ctatgctcttcatgctactcaatctattcaaaatacgccagtgctccccttgccttatacacaggctcttcactcacatcttgaccatcttggatgtgggtactgtagct
cctgggaaccagtcgtgcgggagcatgagaaccacgtcgaagaagatggtgacccaggaaagcaaaaacttgggcacattcttgtagataccctgacctgt
cgtggtcgaagataccgtcgaaaaccagctgtcccattgacagaaaccctcctcctgtatctagcagaacgctggcaatagaaaagccattagtggacttttcg
*cgtgtagttcgacgcagcagaggatagtgctt*gcatgctgaggtgtctcacaagtgccgtgcagtcccgcccccacttgcttctctttgtgtgtagtgtacgtacatta
tcgagaccgttgttcccgcccacctcgatccggcatgctgaggtgtctcacaagtgccgtgcagtcccgcccccacttgcttctctttgtgtgtagtgtacgtacattat

cgagaccgttgttcccgcccacctcgatccggcatgctgaggtgtctcacaagtgccgtgcagtcccgcccccacttgcttctctttgtgtgtagtgtacgtacattatc
gagaccgttgttcccgcccacctcgatccggcatgctgaggtgtctcacaagtgccgtgcagtcccgcccccacttgcttctctttgtgtgtagtgtacgtacattatcg
agaccgttgttcccgcccacctcgatccggcatgcactgatcacgggcaaaagtgcgtatatatacaagagcgtttgccagccacagattttcactccacacaccac
atcacacatacaaccacacacatccacaATGTCCTCTTCTTCTTCGTCCTCTTCCTCTATGTCTCTGCCCCTGGCTACC
CCTCTGATCAAGGACAACGAGTCCCTGATCAAGTTCCTGCGACAGCCCCTGGTGCTGCCCCACGAG
GTGGACGACTCTACCAAGCGACGAGAGCTGCTCGAGCGAACCCGAAAGGAACTCGAGCTGAACGC
CGAGAAGCCCCTCGAGGCCCTGAAGATGATCGACATCATCCAGCGACTGGGCCTGTCTTACCACTTC
GAGGACGACATCAACTCTATCCTGACCGGCTTCTCTAACATCTCTTCGCAGACCCACGAGGACCTGC
TGACCGCCTCTCTGTGCTTCCGACTGCTGCGACACAACGGCCACAAGATCAACCCCGACATCTTCCA
AAAGTTCATGGACAACAACGGCAAGTTCAAGGACTCTCTGAAGGACGACACCCTGGGCATGCTGTC
TCTGTACGAGGCCTCTTACCTGGGCGCCAACGGTGAAGAGATCCTGATGGAAGCCCAAGAGTTCAC
CAAGACTCACCTGAAGAACTCGCTGCCCGCCATGGCTCCCTCGCTGTCTAAGAAGGTGTCTCAGGC
CCTGGAACAGCCCCGACACCGACGAATGCTGCGACTCGAGGCTCGACGATTCATCGAGGAATACGG
CGCCGAGAACGATCACAACCCTGACCTGCTGGAACTGGCCAAGCTGGACTACAACAAGGTGCAGTC
TCTGCACCAGATGGAACTGTCTGAGATCACCCGATGGTGGAAGCAGCTGGGCCTCGTGGACAAGCT
GACTTTCGCCCGAGATCGACCCCTCGAGTGCTTCCTGTGGACCGTGGGACTGCTGCCCGAGCCTAA
GTACTCTGGCTGCCGAATCGAGCTGGCCAAGACCATTGCCATCCTGCTGGTGATCGACGACATTTTC
GACACCCACGGCACCCTGGACGAGCTGCTGCTGTTCACCAACGCCATCAAGAGATGGGACCTCGAG
GCTATGGAAGATCTGCCCGAGTACATGCGAATCTGCTACATGGCCCTGTACAACACCACCAACGAGA
TCTGTTACAAGGTGCTGAAGGAAAACGGCTGGTCTGTTCTGCCCTACCTGAAGGCCACCTGGATCG
ACATGATCGAGGGCTTCATGGTCGAGGCCGAGTGGTTCAACTCTGACTACGTGCCCAACATGGAAG
AATACGTCGAGAACGGCGTGCGAACCGCCGGCTCTTACATGGCTCTGGTGCACCTGTTCTTTCTGAT
CGGCCAGGGCGTGACCGAGGACAACGTGAAGCTGCTGATCAAGCCCTATCCTAAGCTGTTCTCTTCT
TCCGGCCGAATCCTCCGACTGTGGGACGACCTGGGCACCGCCAAGGAAGAACAAGAGCGAGGCGA
CCTGGCCTCTTCTATCCAGCTGTTCATGCGAGAGAAGGAAATCAAGTCTGAGGAAGAAGGCCGAAA
GGGCATCCTCGAGATCATCGAGAACCTGTGGAAGGAACTGAACGGCGAGCTGGTGTACCGAGAGG
AAATGCCTCTGGCCATCATCAAGACCGCCTTCAACATGGCCCGAGCTTCTCAGGTGGTGTACCAGCA
CGAGGAAGATACCTACTTCTCTTCTGTGGACAACTACGTGAAGGCCCTGTTCTTCACCCCTTGCTTCT
AActtctgttcggaatcaacctcaaggttaacggccacgatcccctcgttgttactcttggtcagcccattgtcggtaacgctggctttgctaactgggtcgataaaact
cttctttggccaggagaaccccgatgtctccaaggtgtccaaagaccgaaagctctaccgaatcacccaccgaggagatatcgtccctcaagtgcccttctgggac
ggttaccagcactgctctggtgaggtctttattgactggcccctgatccaccctcctctctccaacgttgtcatgtgccagggccagagcaataaacagtgctctgccg
gtaacactctgctccagcaggtcaatgtgattggaaaccatctgcagtacttcgtcaccgagggtgtctgtggtatctaagctatttatcactctttacaacttctacctca
actatctactttaataaatgaatatcgtttattctctatgattactgtatatgcgttcctctaagacaaatcgaaaccagcatgcgatcgaatggcatacaaaagtttcttccg
aagttgatcaatgtcctgatagtcaggcagcttgagaagattgacacaggtggaggccgtagggaaccgatcaacctgtctaccagcgttacgaatggcaaatgac
gggttcaaagccttgaatccttgcaatggtgccttggatactgatgtcacaaacttaagaagcagccgcttgtcctcttcctcgaaactctcaaacacagtccagaagt
cctttatagtttgatctgtatccagatagcctccgtaattggtgtgtgtcttcaaatcccagacgtccacattggcatgtcctccactgataagcatttgaagttcatctgcg
ttgaacattgagacccacgaagggtcaatgagctggtatagaccgcccaagaatgcatctg***tgtcgaaatacaacagccagtcctccacgatccagctcagata***
gccctgggtgttgagatagagacaagtcgagcgagatgttctcagtggacttgagtttgatgttgtgaaagatggttggattgagagcacaactccatagaaca
tagtagtctcagtccatgagctaggagagaatctccagtgctgatgtcatttgaagaagaagaggaagaagaagacagatagatcaactcagccccttccag
ttgccgtacaatgactcaaccaaatacgttcaagcctgctttcaagcctgcgtgctctttctgtctcattagttgaacgtagatcatctgagagaacttggcagacc
gtgcaaacacaagggcgagactgataggacacttgatctgatccaaaatttgacagaggaaaactagttttgaagttactgaacgtttcttctccaattgcttcta
acaagcttggcgtaatcatggtcatagctgtttcctgtgtgaaattgttatccgctcacaattccacacaacatacgagccggaagcataaagtgtaaagcctggggtg
cctaatgagtgagctaactcacattaattgcgttgcgctcactgcccgctttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggccaacgcgcgggg
agaggcggtttgcgtattgggcgctcttccgcttcctcgctcactgactcgctgcgctcggtcgttcggctgcggcgagcggtatcagctcactcaaaggcggtaat
acggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcgtttttcc
ataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggactataaagataccaggcgtttcccccctggaagc
tccctcgtgcgctctcctgttccgaccctgccgcttaccggatacctgtccgcctttctcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagt
tcggtgtaggtcgttcgctccaagctgggctgtgtgcacgaaccccccgttcagcccgaccgctgcgccttatccggtaactatcgtcttgagtccaacccggtaag
acacgacttatcgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttcttgaagtggtggcctaactacggctac
actagaagaacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccggcaaacaaaccaccgctggtagcggtggt
ttttttgtttgcaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatcttttctacggggtctgacgctcagtggaacgaaaactcacgttaagg
gattttggtcatgagattatcaaaaaggatcttcacctagatccttttaaattaaaaatgaagttttaaatcaatctaaagtatatatgagtaaacttggtctgacagttacca
atgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacgggagggcttaccatctgg

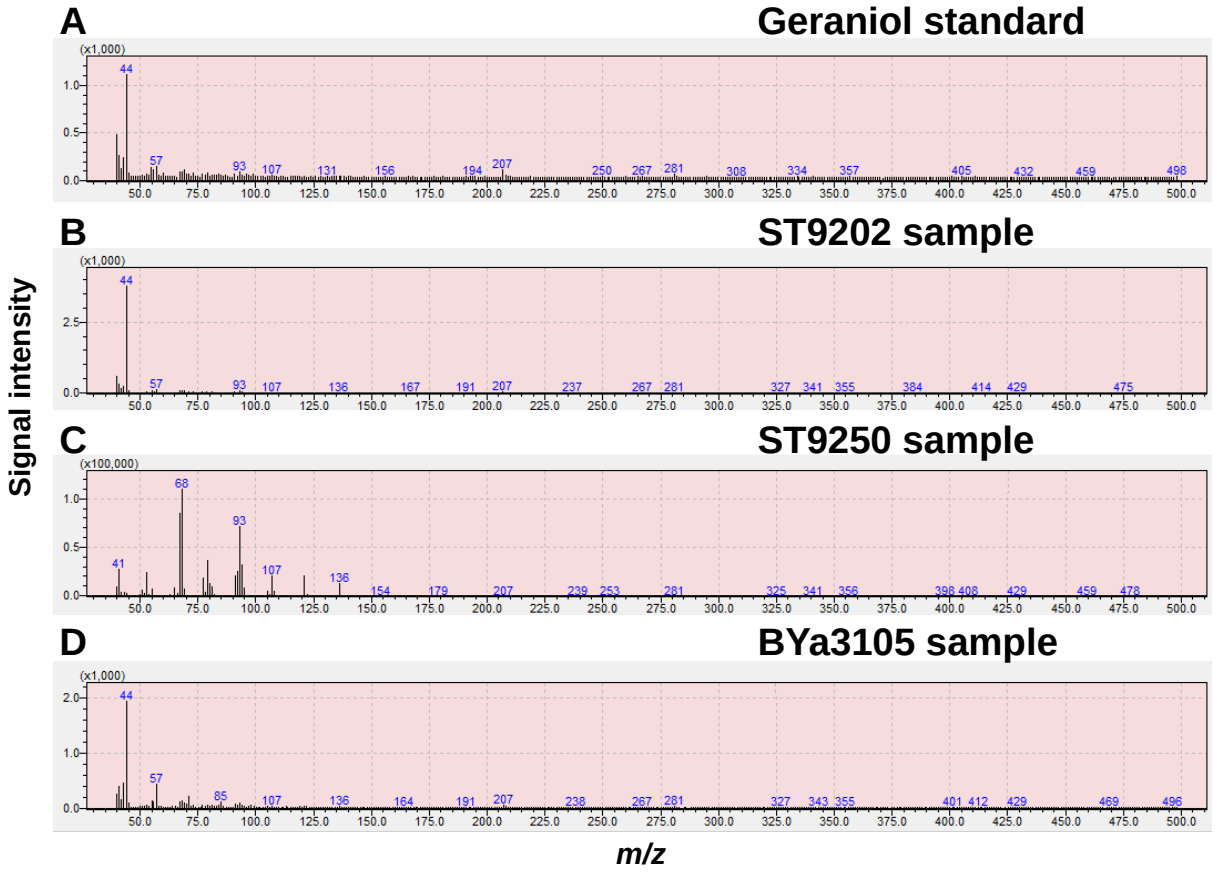
ccccagtgctgcaatgataccgcgagacccacgctcaccggctccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggtcctgca
actttatccgcctccatccagtctattaattgttgccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttgttgccattgctacaggcatcgtggtgt
cacgctcgtcgtttggtatggcttcattcagctccggttcccaacgatcaaggcgagttacatgatcccccatgttgtgcaaaaaagcggttagctccttcggtcctccg
atcgttgtcagaagtaagttggccgcagtgttatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgcttttctgtgactggtgagt
actcaaccaagtcattctgagaatagtgtatgcggcgaccgagttgctcttgcccggcgtcaatacgggataataccgcgccacatagcagaactttaaaagtgctc
atcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcgatgtaacccactcgtgcacccaactgatcttcagcatcttttact
ttcaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaataagggcgacacggaaatgttgaatactcatactcttccttttcaatat
tattgaagcatttatcagggttattgtctcatgagcggatacatatttgaatgtatttagaaaaataaacaaatagggttccgcgcacatttccccgaaaagtgccacct
gacgtctaagaaaccattattatcatgacattaacctataaaaataggcgtatcacgaggccctttcgtc



App.Figure 14: Photo of dodecane overlay samples harvested from cultures of geraniol production assay. Cultures were grown in YP4D medium at 30° C and 250 rpm for approximately 72h to asses monoterpene production through dodecane extraction and GC-MS identification. Negative control was harvested from culture without cell inoculum. Cultures in duplicates were conducted for BYa3105 (the genetically-modified ST9202 with the *crGES* cassette integrated into its genome), ST9202 (the monoterpene production chassis strain) and ST9250 (the genetically-modified ST9202 with the *pfLS* cassette integrated into its genome).



App.Figure 15: Mass spectra of samples from geraniol production assay (cultures were grown in YP4D medium at 30° C and 250 rpm for approximately 72h to asses monoterpene production through dodecane extraction and GC-MS identification) at the geraniol retention time (16.5). (A) Mass spectra of geraniol standard solution in dodecane; (B) Mass spectra of dodecane sample from ST9202 (the monoterpene production chassis strain) culture; (C) Mass spectra of dodecane sample from ST9250 (the genetically-modified ST9202 with the *pfLS* cassette integrated into its genome) culture; (D) Mass spectra of dodecane sample from BYa3105 (the genetically-modified ST9202 with the *crGES* cassette integrated into its genome) culture.



App.Figure 16: Mass spectra of samples from geraniol production assay (cultures were grown in YP4D medium at 30° C and 250 rpm for approximately 72h to asses monoterpene production through dodecane extraction and GC-MS identification) at the limonene retention time (9.5). (A) Mass spectra of geraniol standard solution in dodecane; (B) Mass spectra of dodecane sample from ST9202 (the monoterpene production chassis strain) culture; (C) Mass spectra of dodecane sample from ST9250 (the genetically-modified ST9202 with the *pfLS* cassette integrated into its genome) culture; (D) Mass spectra of dodecane sample from BYa3105 (the genetically-modified ST9202 with the *crGES* cassette integrated into its genome) culture.