Optimization of *de novo* 10-hydroxygeraniol production in *Saccharomyces cerevisiae*

Meghan Davies

A Thesis
In the Department
of
Biology

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science (Biology) at
Concordia University
Montreal, Quebec, Canada
September 2019

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Meghan Davies

Entitled: Optimization of *de novo* 10-hydroxygeraniol production in *Saccharomyces cerevisiae*

and submitted in partial fulfillment of the requirements for the degree of

Master of Science (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair
Dr. Malcolm Whiteway

_____ Examiner
Dr. Peter Pawelek

_____ Examiner
Dr. Malcolm Whiteway

_____ Examiner
Dr. Brandon Findlay

_____ Supervisor
Dr. Vincent Martin

Approved by _____
Dr. Patrick Gulick, Chair of Department

Date September 30, 2019

_____
Andre Roy, Dean of Arts and Science

ABSTRACT

Optimization of *de novo* 10-hydroxygeraniol production in *Saccharomyces cerevisiae*

Meghan Davies

Monoterpenes and their indole alkaloid derivates are plant secondary metabolites valuable to the pharmaceutical and industrial sectors for their bioactivities in humans, cats and insects. The biosynthetic pathway to strictosidine, the central monoterpene indole alkaloid scaffold, proceeds via an industrially relevant intermediate – nepetalactol. The oxidized counterpart of this monoterpene, nepetalactone, is known for its euphoric influence on cats but is also a potent insect repellent. Unfortunately, low *in planta* levels of these metabolites have resulted in high production costs and environmentally unfeasible cultivation strategies for an industrial product with mass market appeal. Engineering a high-producing nepetalactol or strictosidine yeast would benefit several commercial sectors, however recent studies attempting *de novo* strictosidine in *S. cerevisiae* only achieved trace titers (0.5 mg/L). Activity of geraniol-10-hydroxylase, a cytochrome P450, was an identified bottleneck in nepetalactol/strictosidine production. This thesis presents the application of a combination of strategies aimed at overcoming the geraniol-10-hydroxylase bottleneck and improving titres of 10-hydroxygeraniol. Geraniol-10-hydroxylase and cytochrome P450 reductase variants were explored to increase enzyme efficiency, wherein several reductases were shown to beneficially influence hydroxylation activity. We further demonstrated that a cytochrome P450-reductase fusion increased carbon flux towards 10-hydroxygeraniol by 1.7-fold. Accumulation of a heterologous metabolite, putatively identified as isopulegol, was observed, implying inefficient channeling of carbon into the heterologous pathway. Deletion of two native reductases, *oye2* and *oye3*, resulted in a 4.6-fold increase in carbon channeled to 10-hydroxygeraniol, resulting in a final accumulation of 144 mg/L of 10-hydroxygeraniol. These pathway optimization steps will aid in the development of high yielding monoterpene *S. cerevisiae* strains.

ACKNOWLEDGEMENTS

I would like to extend gratitude to my supervisor Dr. Vincent Martin for giving me the opportunity to work and learn in his lab. This invaluable learning experience would not have been possible without his mentorship, guidance and, especially, his patience. I am grateful for being given the chance to be a part of his lab, thank you Vince!

I would also like the thank my committee members Dr. Malcom Whiteway and Dr. Brandon Findlay for reading this thesis and for providing their thoughtful feedback throughout my project.

Next, I would like to give a big thanks to Bjorn Bean and Shoham Mookerjee for editing of my thesis, especially when exceptionally long paragraphs were involved; Bjorn Bean, again, for his wonderful microscopy work; Lauren Narcross for experimental set-up and overall help with everything; and Audrey Morin for keeping this lab running smoothly!

Many thanks to current and past Martin lab members for all their help and encouragement along the way, and putting up my incessant habit of talking to myself: Mohamed Nasr, Michael Pyne, Nicholas Gold, Elena Fossati, Mindy Melgar, Leanne Bourgeois, Daniel Tsyplenkov, James Bagley, Kaspar Kevvai, Mathieu Husser, Will Cheney and Smita Amarnath. Thank you everyone.

And last, but not least, I would like to thank my mom, dad and brother for their support and encouragement!

AUTHOR CONTRIBUTIONS

This project was conceived by Dr. Vincent J.J. Martin. Dr. Bjorn Bean built the microscopy strains and imaged and processed the final figures used in this study. All other experiments were designed, implemented and analyzed by Meghan Davies.

CONTENTS

LIST OF FIGURES viii
LIST OF TABLES ix
LIST OF ABBREVIATIONS: x
1. INTRODUCTION 1
1.1 Monoterpene Indole Alkaloids (MIA) 1
1.2 Economics and production strategies of MIAs and nepetalactone 3
1.2.1 Economic feasibility of MIA and nepetalactone production 3
1.2.2 Optimization of *C. roseus* cell cultures for the production of MIAs 6
1.2.3 Genetic modification of *C. roseus* 6
1.2.4 Direct chemical synthesis of MIAs 7
1.3 Biosynthesis of nepetalactol and the global MIA precursor, strictosidine 8
1.3.1 Precursor pathways 8
1.3.2 Strictosidine and nepetalactone formation *in planta* 11
1.4 Optimization of nepetalactol production in *S. cerevisiae* 12
1.4.1 Optimization of native yeast pathways 12
1.4.2 Introduction and optimization of plant pathways 13
1.5 Strategy for optimizing 10HG production in *S. cerevisiae* 18
2. MATERIALS AND METHODS 19
2.1 Strains and media 19
2.2.1 Plasmids and genes 19
2.2.2 Library construction and phylogenetic analysis 20
2.2.3 Gene modifications 20
2.2.4 Linearization of pCAS and gRNA design 20
2.3 Yeast transformation and CRISPR/Cas9-mediated genome integration 21
2.4 Colony PCR 23
2.5 Microscopy 23
2.6 Analysis of Monoterpenes (GC-MS) 23
3. RESULTS 24
3.1 Increasing *in vivo* geraniol titers 24
3.2 Testing G10H variant library for geraniol hydroxylation 26
3.3 N-terminal exchange of G10H variants 29
3.4 Optimizing *Cr*G10H activity in *S. cerevisiae* 31
3.5 Combinatorial CPR analysis 33

3.6 Assessing the 3x-*CrG10H* strain with non-cognate CPRs 37
3.7 Increasing carbon flux towards 10HG 39
3.7.1 Decreasing off-target reactions 39
3.7.2 Overall increase in carbon flux towards 10-hydroxygeraniol 43
4. DISCUSSION 43
4.1 Redirecting carbon for GPP production 45
4.2 Assessing the activity of the G10H variant library 45
4.3 Assessing differential activity from chimeric G10H variants 47
4.4 CPR influence on CYP activity 48
4.5 Identification and reduction of off-target metabolism 49
4.6 Potential effect of *OYE* deletions on cell growth 50
5. CONCLUSION 53
REFERENCES 54
APPENDIX 67

LIST OF FIGURES

Figure 1. Overview of MIA biosynthesis beginning with the synthesis of strictosidine from nepetalactol. 2
Figure 2. Structure and stereochemistry of nepetalactol and nepetalactone. 4
Figure 3. Overview of nepetalactone and strictosidine biosynthesis in engineered *S. cerevisiae*. 10
Figure 4. Design of the N-terminal exchange for the G10H variants. 22
Figure 5. Geraniol titers of engineered *S. cerevisiae* and its influence on strain growth. 25
Figure 6. Monoterpene production in engineered *S. cerevisiae* strains harboring G10H variants 28
Figure 7. The effect of the N-terminal sequence on geraniol hydroxylation *in vivo*. 30
Figure 8. GFP localization of *Cr*G10H and *St*G10H(Nt)-*Cr*G10H. 30
Figure 9. Monoterpene production from strains expressing multi-copy *CrG10H* and *Cr*G10H-*Cr*CPR fusion. 32

LIST OF TABLES

Table 1. Percent molar conversion of geraniol to 10HG for the multi-copy *CrG10H* and *Cr*G10H-*Cr*CPR fusion strain..... 32
Table 2. Percent molar conversion of geraniol to 10HG for different CPRs paired with *Cr*G10H..... 36

LIST OF ABBREVIATIONS:

(in order of appearance)

MIA: monoterpene indole alkaloid

DW: dry weight

DEET: N,N-Diethyl-m-toluamide

G10H: geraniol-10-hydroxylase

TDC: tryptophan decarboxylase

STR: strictosidine synthase

IPP: isopentenyl pyrophosphate

DMAPP: dimethylallyl pyrophosphate

MVA: mevalonate

ACAT: acetoacetyl-CoA thiolase

HMG-CoA: 3-hydroxy-3-methylglutaryl CoA

HMGS; Erg13: 3-hydroxy-3-methylglutaryl CoA synthase

HMGR: 3-hydroxy-3-methylglutaryl CoA reductase

MVK; Erg12: mevalonate kinase

GPP: geranyl pyrophosphate

FPP: farnesyl pyrophosphate

PMVK; Erg8: phosphomevalonate kinase

MVD: mevalonate pyrophosphate decarboxylase

IDI: isopentenyl diphosphate isomerase

FPPS; Erg20: farnesyl pyrophosphate synthase

GPPS: geranyl pyrophosphate synthase

GES: geraniol synthase

10HG: 10-hydroxygeraniol

CYP: cytochrome P450

CPR: cytochrome P450 reductase

10HGO: 10-hydroxygeraniol oxidoreductase

IS: iridoid synthase

NEPS: nepetalactol-related short-chain reductases

IO: iridoid oxidase

7DLH: 7-deoxyloganic acid hydroxylase

SLS: secologanin synthase

tHMGR: truncated 3-hydroxy-3-methylglutaryl CoA reductase

ER: endoplasmic reticulum

OYE: old yellow enzyme

KO: kaurene oxidase

KAH: kaurenoic acid hydroxylase

PCR: polymerase chain reaction

DNA: deoxyribonucleic acid

SOE: sequence overlap extension

gRNA: guide ribonucleic acid

CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats

Cas9: CRISPR-associated protein 9 from *Streptococcus pyogenes*

OD_{600}: optical density measured with a wavelength of 600

1. INTRODUCTION

1.1 Monoterpene Indole Alkaloids (MIA)

Monoterpene indole alkaloids (MIAs) are a diverse and structurally complex family of bioactive secondary metabolites ubiquitous among the Apocynaceae, Loganiaceae and Rubiaceae plant families - with certain species producing upwards of 100 different MIAs[1]. These metabolites have become industrially relevant over the past few decades for their potent pharmaceutical activities; particularly their analgesic, antimicrobial, anticancer and antispasmodic properties[2]. In 1958, Canadian chemists Robert Noble and Charles Beer isolated two such MIAs, vinblastine and vincristine, from the leaves of *Catharanthus roseus*[3]. These metabolites became pharmaceutically important for their anti-tumor properties involving the potent inhibition of microtubule assembly[4,5]. However, as with most secondary metabolites, *in planta* quantities are negligible and are often present as complex mixtures making isolation economically unviable. For instance, vinblastine and vincristine constitute approximately 0.01% and 0.003% dry plant weight (DW) respectively, with quantities greatly dependent on plant growth conditions[4]. These low titers can be attributed to the biological complexity of MIA biosynthesis. In the case of vinblastine, the biosynthetic pathway is composed of at least 35 intermediates, 30 enzymes and 2 regulatory genes spread over several intra and inter cellular compartments[6]. This elaborate pathway is representative of most native MIA biochemical pathways. In addition, the majority of these biosynthetic routes are poorly characterized, as many steps remain to be elucidated[7]. The low natural abundance of the pharmaceutically important vinblastine and vincristine explains why *C. roseus* is one of the most extensively studied medicinal plants. This research on *C. roseus* identified a common metabolite, strictosidine, involved in all MIA synthesis[6,8,9]. All identified MIAs can be derived from various oxidative, substitution and rearrangement reactions of strictosidine[10,11]. Thus, any optimization strategies for MIA production involves optimizing the production of strictosidine and its precursors, one of which is nepetalactol (Figure 1).

Nepetalactol is a monoterpene intermediate in the strictosidine biosynthetic pathway that is industrially relevant for its oxidized counterpart – nepetalactone (Figure 1 and 2A). Nepetalactone is generally found among the *Nepeta* plant species, commonly referred to as

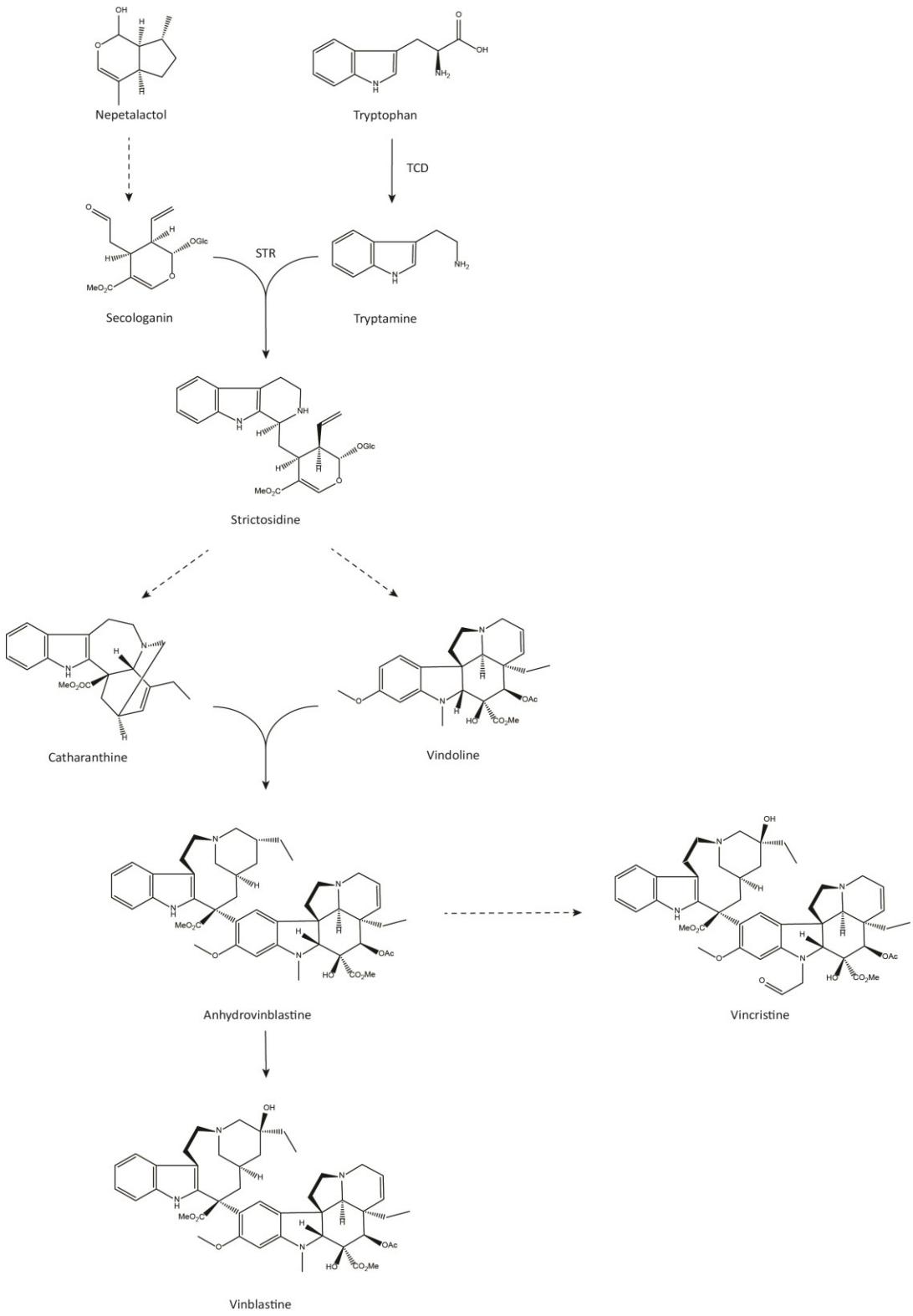

Figure 1. Overview of MIA biosynthesis beginning with the synthesis of strictosidine from nepetalactol.
Hashed lines denote mulitple enzymatic steps. Abbreviations: TDC, tryptophan-decarboxylase; STR, strictosidne synthase.

catnip, where it is the biologically active ingredient in commercial catnip. The intoxication effects of this molecule on cats is well known, however it has also occupied a less well-known niche since the 1960's as insect repellent. Nepetalactone is effective at repelling various species of insects, namely mosquitos, ticks, mites and cockroaches[12–15]. It can be at least as effective a repellent as N,N-Diethyl-m-toluamide (DEET), which makes nepetalactone valuable given the growing population of DEET resistant mosquitoes[16].

1.2 Economics and production strategies of MIAs and nepetalactone

1.2.1 Economic feasibility of MIA and nepetalactone production

Commercial production strategies for both nepetalactone and most MIAs rely heavily on the mass cultivation and extraction from the desired plant species. Generally, plants must reach maturity before the leaves are harvested and prepared for extraction[17]. With respect to MIAs, several direct extraction methods have been explored in attempts to increase extraction efficiency. These include ionic liquid ultrasound assisted extraction[18], supercritical CO_2 fluid extraction with methanol as a co-solvent (achieving 0.0466 mg vinblastine/g DW)[19] and negative-pressure cavitation extraction (0.126 mg vinblastine/g DW)[20]. Though these optimized methods succeeded in increasing the extraction efficiency, *in planta* levels of MIAs are too low to be economically feasible for pharmaceutical interest.

Given the biological importance of vinblastine and vincristine, MIAs may be an untapped source of pharmaceutically relevant molecules. However, many drug screening efforts exclude plant natural products in favor of screening larger numbers of simpler molecules, which can be produced inexpensively using chemical synthesis[21,22]. The current capital needed to develop and bring a drug to market is around $0.8–1.7 billion, allocated to discovery (10%), preclinical (15%), manufacturing and processing (15%), clinical trials (55%) and post-marketing (5%)[23]. Taking the high failure rates of clinical development into consideration, currently estimated at 94-96%, a high market potential is required for a big pharmaceutical company to be interested in any new product[23]. Although the chemical syntheses of plant natural products, particularly alkaloids, are improving[24], MIAs pose too great of a financial risk where many syntheses are still too complicated or low yielding for drug discovery and commercial production, or require industrially impractical separation steps[21]. With the development of strategies to produces MIAs

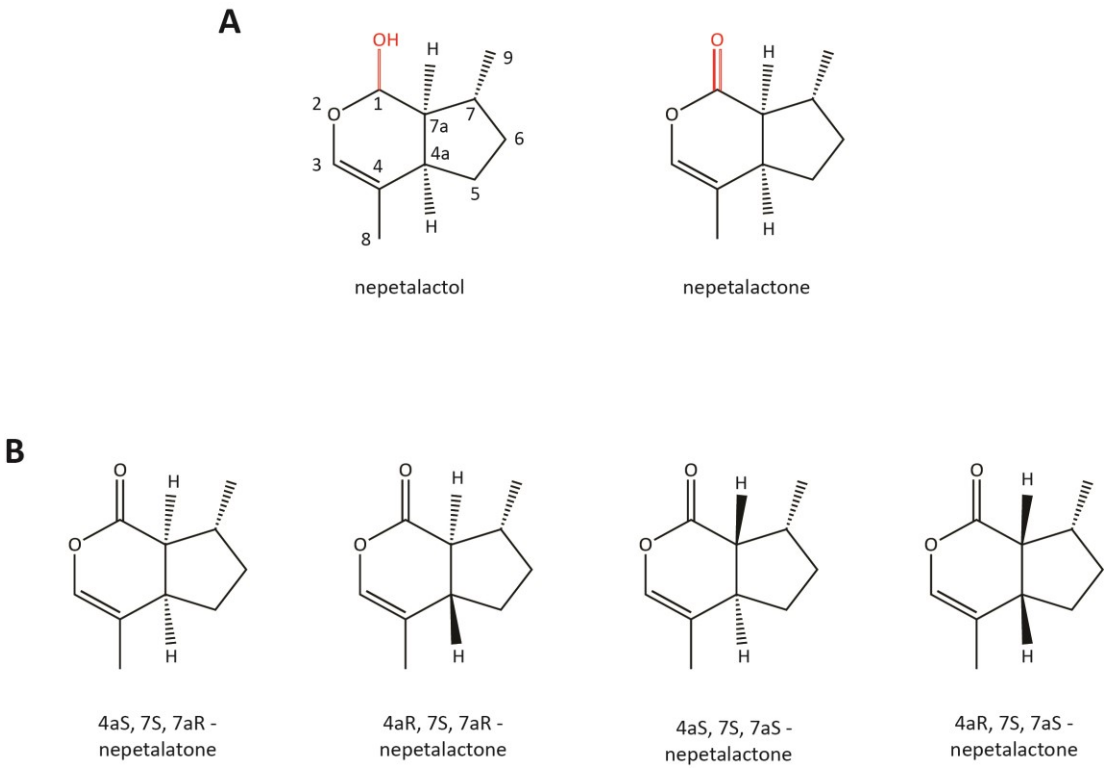

Figure 2. Structure and stereochemistry of nepetalactol and nepetalactone.
A) The point of oxidation (red) for nepetalactol to nepetalactone. The numbers designate the position of the 4a and 7a chiral centers. **B)** Structures of the various nepetalactone stereoisomers.

in an efficient and economically viable manner, the pharmaceutical potential of MIAs could be realized.

Compared to MIAs, nepetalactone has a shorter biosynthetic pathway and higher *in planta* yields, though not high enough to meet growing demand. Currently, catnip production is centered in Western USA and Canada where the limited commercial crop area is dedicated to the production of essential oils or dry leaves for catnip toys and herbal uses[25]. The extraction of nepetalactone oil from high yielding plants hovers around 15 mg nepetalactones/g DW[26]. Depending on the plant species, this 15 mg can be composed of several stereoisomers produced by the 4a and 7a chiral centers (Figure 2B). *Nepeta cataria* is the most commonly cultivated species with three nepetalactone enantiomers present in its extracted oil, the predominant being (4aS,7S,7aR) (up to 90%) and (4aS,7S,7aS) (up to 78%)[27]. It has been observed that optimal insect repellency requires a specific 3:2 ratio of the (4aS,7S,7aR) and (4aS,7S,7aS) enantiomers[12,28], while cats prefer the (4aS,7S,7aS) enantiomer[27]. Keeping this ratio consistent, it was found that 23 ug/cm^2 of body surface area is adequate for effective repellency of mosquitoes for up to 6 hours[28]. Assuming an average adult has a body surface area around 1.85 m^2 (160lbs, 170cm) and applies the repellant to half their body, it would require approximately 213 mg of nepetalactones per 6 hours.

A study conducted on a commercial herb farm in Richland, New Jersey found that the highest yielding crop of nepetalactone involved 80 724 plants/ha, resulting in 770 kg/ha DW from which 12.5 kg/ha of total essential oil was extracted[25]. Seventy-three percent of the extracted oil was attributed to (4aS,7S,7aR)-nepetalactone, totalling 9.13 kg/ha. Extrapolating to a per plant basis, this corresponds to approximately 113 mg of this enantiomer per plant. Returning to the assumption that one person requires 213 mg (of a 3:2 ratio) per 6 hours, this suggests that ~ 1 plant/person/day is required if an insect repellent is desired. In 2016, the mosquito repellent market size in North America alone was estimated at $969.6 million USD and is expected to reach $4.8 billion in 2022 coinciding with the rising threat of mosquito-borne diseases[29–31]. To this extent, the amount of land allocation required to sustain any reasonable demand of a nepetalactone based insect repellent would be unfeasible, with one hectare of land providing enough nepetalactone for around 2700 people for one month (assuming use is once

every day). Factoring in the amount of water to sustain the required crops[32] further emphasizes the wastefulness and inefficiency of current strategies.

The remainder of the section will focus on optimization strategies for MIA biosynthesis. To date, there have been no apparent attempts to optimize nepetalactone yields either through direct optimization of *in vitro*/*in vivo* methods or total chemical synthesis, further emphasizing the need to develop efficient production strategies.

1.2.2 Optimization of *C. roseus* cell cultures for the production of MIAs

The use of plant cell bioreactors is an attractive alternative to improve secondary metabolite formation, including MIA production. They provide favorable conditions in which several factors can be tested to boost MIA yields without environmental, ecological or climatic limitations and allow cells to proliferate at higher growth rates than whole plants in cultivation[33]. Such optimization methods include screening and selecting for high-producing cell lines, exogenously supplying specific pathway regulators or precursors, changing medium conditions or culture conditions such as light, temperature, and aeration[6]. A bench top bioreactor was designed for *C. roseus* plant cell cultures allowing continuous extraction of secondary metabolites. The addition of *trans*-cinnamic acid, an inhibitor of phenylalanine ammonia lyase activity[34], paired with mannitol-induced osmotic stress increased the total alkaloid content to around 110 mg/L, an approximate 11-fold increase compared to the control[35]. Although this is a substantial increase, it is for total alkaloid content and not specific to the valuable MIAs. Production of complex MIAs in *C. roseus* cultures is limited by the complexity of metabolic regulation involving intracellular compartmentation, several different cell types and tissue differentiation[36]. In terms of complex MIAs, UV-B irradiation of *C. roseus* cell cultures increased yields of the vinblastine precursors, vindoline and catharanthine (Figure 1) to 0.06 ± 0.0023 and 0.12 ± 0.0054 mg/g DW corresponding to a 117- and 3-fold improvement over the control, respectively[37]. Even with minor improvements, MIA yields are still economically non-viable thus driving the need to find better methods.

1.2.3 Genetic modification of *C. roseus*

Another potential method to increase MIA yields is through genetic modification of *C. roseus*. Transgenic *C. roseus* strains have been created in which MIA synthesizing or regulating

genes have been modified, including the overexpression of transcription factors and transporters[38,39]. The overexpression of the *Orca3* transcription factor and the *C. roseus* geraniol-10-hydroxylase (G10H) in leaf tissue increased yields to 2.13 ± 0.88 mg vindoline/g DW and 4.57 ± 0.79 mg catharanthine/g DW, 3- and 2.3- fold improvements, respectively[40]. Recently, the overexpression of the *C. roseus* tryptophan decarboxylase (*Cr*TDC) and strictosidine synthase (*Cr*STR) (Figure 1), two rate-limiting enzymes involved in strictosidine biosynthesis, in transformed leaf tissue increased vindoline and catharanthine content to 3.0 ± 0.3 and 1.0 ± 0.1 mg/g DW respectively, with an approximate 6- and 2.5-fold increase compared to the control[41]. However, transgenic *C. roseus* lines are unstable, with the wildtype phenotype re-emerging after successive generations[6]. While the direct over-expression of encoding genes has resulted in the improved accumulation of desired products, the overall low yields and the instability of genetic modifications highlight the need for better methods.

1.2.4 Direct chemical synthesis of MIAs

As *in planta* genetic modification strategies for MIA production proved to be ineffective and non-robust, several semi-synthetic strategies were developed for MIA production. The more abundant and naturally occurring plant MIAs, vindoline and catharanthine, were extracted and coupled via chemical means to produce vinblastine following the oxidation of the anhydrovinblastine intermediate[42–45] (Figure 1). Initial protocols resulted in an overall yield of 40-43% vinblastine, with a 20-23% loss to leurosidine, a vinblastine regioisomer[45]. While vindoline and catharanthine are more abundant than vinblastine, their yields are limited by plant production. Since this alternative semi-synthetic method relies heavily on limited natural sources of these secondary metabolites, full chemical synthesis of vindoline and catharanthine was investigated. Yokoshima *et al.*[46] accomplished a stereo-controlled total synthesis of vinblastine via chemical derivation of both vindoline and a pseudo-catharanthine partner, involving 42 intermediates (19 for vindoline, 21 for pseudo-catharanthine and 2 for their coupling) with an overall total conversion of less than one percent[46,47]. Though this is an impressive feat and a valuable starting point for optimized MIA production, overall yield was reported as “nearly quantitative” and synthesis was laborious requiring an extensive list of chemicals and conditions which would not be feasible to replicate at industrial scales[48].

Overall, the current mass cultivation strategies for commercial production of both nepetalactol and high value MIAs are suboptimal. High costs and poor yield of plant and chemical synthesis have stimulated interests in developing alternative large-scale production processes. An attractive alternative is to reconstitute these pathways in a microbial host. As the substrates for microbial fermentation are renewable and most processes occur at 30-37˚C, biological production processes are economical and not energy intensive[49–51]. This provides a more scalable, robust and cost-efficient solution for MIA production. Furthermore, extensive metabolic engineering studies with *Saccharomyces cerevisiae* have demonstrated its potential to accumulate high titers of alkaloids and terpenes, with significant potential for improvement in yields[10,52,61,53–60]. To this effect, the introduction and optimization of a nepetalactol production platform in *S. cerevisiae* would benefit both MIA and nepetalactone production.

1.3 Biosynthesis of nepetalactol and the global MIA precursor, strictosidine

The identification of the MIA precursor biosynthetic pathway has enabled the heterologous production of strictosidine in *S. cerevisiae*[57], establishing a potential production platform for MIA synthesis. This platform would also enable scalable and economically feasible production of nepetalactone.

1.3.1 Precursor pathways

The condensation of two isoprene units is responsible for initiating synthesis of the secologanin that, when condensed with tryptamine, forms the structural backbone strictosidine, the precursor to all MIAs[11] (Figure 3). Whereas tryptamine is derived from the decarboxylation of tryptophan by TDC, secologanin is produced in a complex series of transformations which originate from isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). Two distinct and independent biosynthetic routes exist for the formation of these pyrophosphates, the mevalonate-dependent (MVA) and the methylerythritol phosphate (or 1-deoxyxylulose-5-phosphate) pathways. In eukaryotes, other than plants, these isoprenes originate from acetyl-CoA via the MVA pathway. Plants can use either the methylerythritol phosphate pathway, in which pyruvate is condensed with glyceraldehyde-3- phosphate[62], or the MVA pathway. While monoterpenes have been suggested to result from the methylerythritol phosphate

pathway *in planta*[63], yeast employs the MVA pathway, hence this route will be described in further detail.

The MVA pathway is initiated with condensation of three molecules of acetyl CoA by acetoacetyl CoA thiolase (ACAT) and 3-hydroxy-3-methylglutaryl CoA (HMG-CoA) synthase (HMGS; Erg13) to yield the intermediate HMG-CoA, which is subsequently reduced to mevalonate by the rate-controlling enzyme HMG-CoA reductase (HMGR) in the first committed step of the pathway[64–67] (Figure 3). HMGR is regulated through a multivalent feedback mechanism mediated by downstream sterol and non-sterol metabolites[68,69], ensuring maintenance of lipid homeostasis by regulating the accumulation of HMG-CoA, which is known to inhibit fatty acid biosynthesis[70,71]. Two variants of HMGR exist in *S. cerevisiae*, Hmgr1 and Hmgr2, with 83% of enzyme activity attributed to Hmgr1[72,73]. Mevalonate is then converted to melvalonate-5-phosphate following two phosphorylation events catalyzed by mevalonate kinase (MVK; Erg12), which is feedback regulated by the downstream isoprenoids geranyl pyrophosphate (GPP) and farnesyl pyrophosphate (FPP)[74,75]. Melvalonate-5-phosphate is then phosphorylated and decarboxylated by phosphomevalonate kinase (PMVK; Erg8) and mevalonate pyrophosphate decarboxylase (MVD), respectively, to produce the basic isoprene building block for all terpenoids, IPP[67,69]. The isomerization of IPP to its highly electrophilic isomer, DMAPP, by isopentenyl diphosphate isomerase (IDI) represents the final step in the MVA pathway[76] and is suggested to be rate-limited, thereby favouring IPP accumulation (Figure 3)[77,78].

A head-to-tail condensation of IPP and DMAPP by farnesyl diphosphate synthase (FPPS; Erg20) produces GPP, the isoprenoid backbone involved in the synthesis of either monoterpenes, or sterols and quinones by further converting GPP to FPP[79]. While Erg20 in *S. cerevisiae* has both GPP and FPP synthase activity, GPP is not released from the catalytic site and is quickly converted to FPP, hence the lack of monoterpene accumulation in most yeast strains[80]. Conversely, plants have GPP synthases (GPPS) that exclusively produce GPP enabling nepetalactone and strictosidine biosynthesis[81–84].

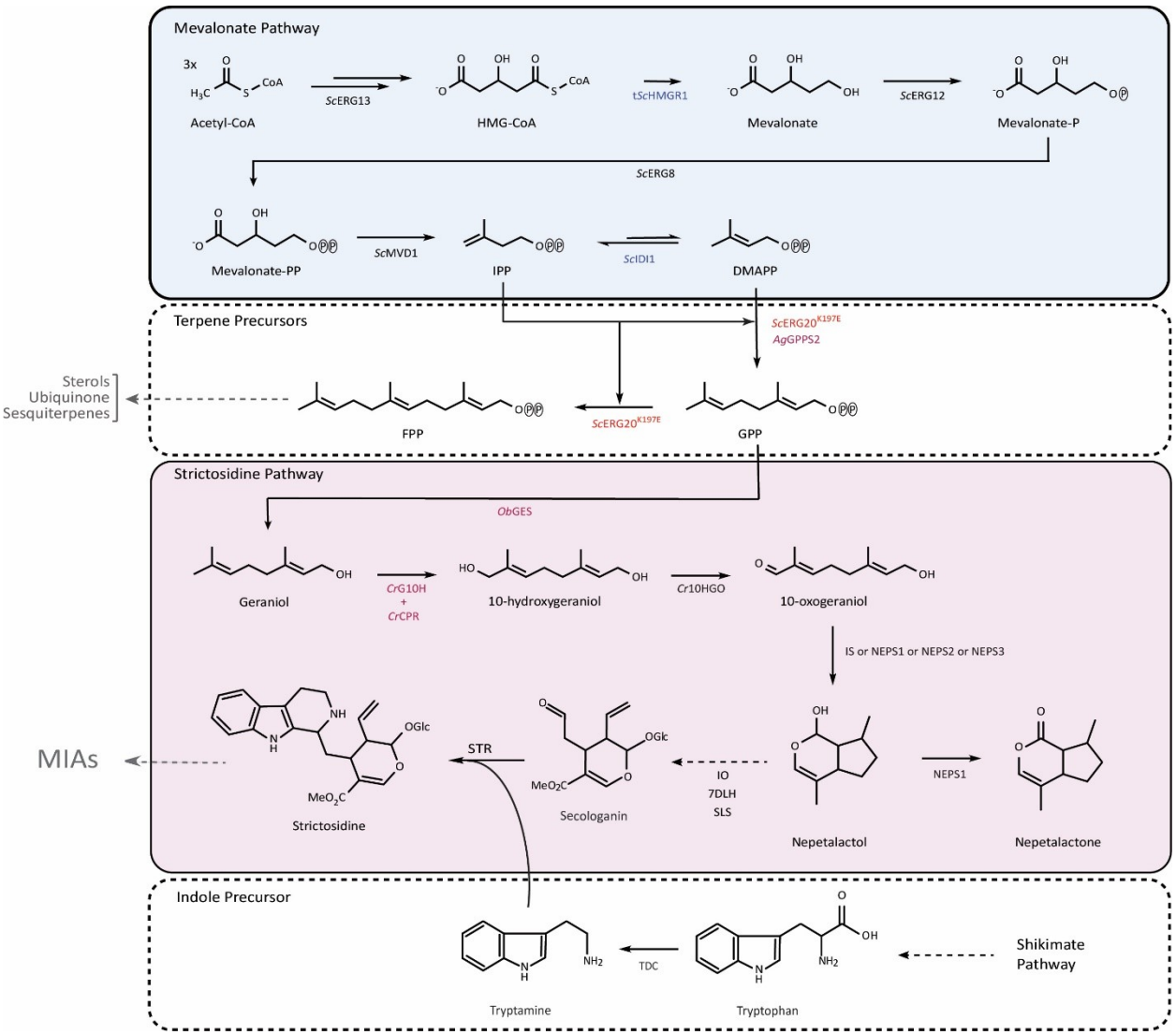

Figure 3. Overview of nepetalactone and strictosidine biosynthesis in engineered *S. cerevisiae*.
The general biosynthetic pathway can be divided into three sections: 1) The MVA pathway (blue), native to *S. cerevisae,* which is responsible for producing the C5 isomer building blocks used to synthesize the monterpnene precursor, GPP. 2) Precursor synthesis (white) which involves GPP formation and production of the shikimate derived tryptamine. 3) Strictosidine biosynthesis (pink), heterlogous to *S. cerevisiae*, wherein a derivitized GPP condenses with the tryptmaine to produce strictosidine. The enzymes involved in optimizing 10HG production in an engineered *S. cerevisiae* strain are colored. Enzymes in blue denote the overexpression of the native yeast genes, those highlighted in red indicate a mutated native yeast enzyme and those in pink denote the expression of heterlogous plant enzymes.

1.3.2 Strictosidine and nepetalactone formation *in planta*

Synthesis of strictosidine in *C. roseus* involves nine enzymatic steps initiated by the dephosphorylation of GPP to geraniol by geraniol synthase (GES)[85] (Figure 3). Hydroxylation of geraniol to 10-hydroxygeraniol (10HG) by G10H, a cytochrome P450 (CYP), and its CYP reductase (CPR) redox partner, define the first committed step towards iridoid monoterpenoids[11,86]. 10HG is oxidized by 10HG oxidoreductase (10HGO)[87] to produce 10-oxogeraniol which is subsequently cyclized by iridoid synthase (IS), an NADPH-dependant reductive cyclase[88], to produce nepetalactol. Production of the nepetalactol intermediate marks a branch point from which nepetalactone biosynthesis can occur. In the *Nepeta* plant species, IS works with three nepetalactol-related short-chain reductases (NEPS) for the cyclisation of 10-oxogeraniol to nepetalactol, determining the stereochemistry of the bridgehead 4a-7a-carbons[89] (Figure 2). IS catalyses the initial reduction of 10-oxogeraniol to the activated oxocitronellyl enol intermediate enabling spontaneous cyclisation to (4aS,7S,7aR)-nepetalactol. During spontaneous cyclisation, Neps1 and Neps2 have been demonstrated to stabilize the enol intermediate thereby reducing off-target iridodial products. Conversely, Neps3 can act on the enol intermediate catalysing its cyclisation to the (4aS,7S,7aS) enantiomer[89]. Subsequent oxidation by Neps1 is responsible for the conversion of the nepetalactol isomers to their cognate nepetalactone isomers[89]. In the continuation of strictosidine synthesis, a complex series of reactions involving three CYPs [iridoid oxidase (IO), 7-deoxyloganic acid hydroxylase (7DLH), and secologanin synthase (SLS)] converts nepetalactol to secologanin, via oxidative ring cleavage[11]. In the final steps, secologanin is condensed with tryptamine by STR to form strictosidine[57,90] (Figure 3). The complete elucidation of the strictosidine biosynthesis pathway has enabled the heterologous production of strictosidine in *S. cerevisiae*. However, multiple points in the native and heterologous pathways have been identified that must be optimized for efficient *in vivo* production of strictosidine in yeast.

1.4 Optimization of nepetalactol production in *S. cerevisiae*

1.4.1 Optimization of native yeast pathways

In vivo GPP accumulation

In order to achieve high titers of strictosidine, the monoterpene biosynthetic pathway in *S. cerevisiae* must be upregulated to ensure high carbon flux to the precursor GPP. The available GPP precursor pool in *S. cerevisiae* is severely limited as it is produced as a tightly enzyme-bound intermediate in the two-step synthesis of FPP (Figure 3). The sequestration of GPP to the enzymatic catalytic site explains why it is not released for the biosynthesis of C10 isoprenoids[79]. Unlike plants, microorganisms do not usually express a specific GPPS, with the exception of a few wine making yeast strains[91]. To overcome this issue, multiple mutants of the Erg20 protein have been characterized in *S. cerevisiae*, of which $Erg20^{K197E}$ specifically, partially disrupts the binding of GPP to the protein, allowing for GPP accumulation[92–94]. Introduction of the K197E mutation into wild type *S. cerevisiae* increased geraniol accumulation from 0.047 ± 0.017 mg/L to 4.2 ± 0.62 mg/L when paired with a plant GES[59]. Several others mutants have been characterized, displaying broad effects on monoterpene and sterol accumulation, and cell growth[59,95]. The additional incorporation of a heterologous plant GPPS in *S. cerevisiae* has also been demonstrated to further enhance GPP accumulation[57,58].

Overexpression of MVA pathway genes for increased precursor supply

Increased production of monoterpenes requires an increased supply of precursors, namely GPP. This can be achieved by upregulating carbon flux through the MVA pathway in yeast. HMGR is the major late-limiting enzyme in the MVA pathway[96]. Several studies aimed at increasing the carbon flux through the MVA pathway for heterologous synthesis of plant secondary metabolites alleviated this bottleneck by the overexpression of all (or most) MVA genes in addition to one to three copies of a truncated *Hmgr1* (t*Hmgr1*)[52,61,97,98]. Since HMGR is an ER bound protein tightly controlled by transcriptional and degradational repression[69], truncation of its N-terminal domain, which is responsible for the ER localization and regulation, circumvents the multivalent feedback mechanism in yeast, resulting in a soluble, degradation-stabilized variant thereby increasing its activity[61,98,99].

In addition to HMGR, IDI was also identified as a rate-limiting enzyme favoring the synthesis of IPP over DMAPP. The overexpression of *IDI* favors the first step of the Erg20 reaction, ie. GPP formation, since under these conditions, DMAPP is more readily available for Erg20 catalysis, leading to a higher GPP/FPP ratio[95]. This study demonstrated that the overexpression of t*Hmgr1* and *IDI* was sufficient to overcome rate-limiting issues by obtaining a 9-fold increase in geraniol titers (6.77 mg/L) when paired with wild type Erg20 and *Ocimum basilicum* GES (*Ob*GES). Further introduction of a mutant Erg20 ($Erg20^{K197G}$) resulted in a 50-fold increase in geraniol, with titers reaching 36.04 mg/L[77].

Reducing off-target activity

Accumulation of heterologous product is often hindered by the activity of native enzymes acting on the pathway intermediates. Geraniol and other strictosidine pathway intermediates are affected in such a manner by the Old Yellow Enzymes (OYE)[100–104]. Two OYE homologs are found in *S. cerevisiae* (Oye2 and Oye3), catalyzing a variety of flavin-dependent, stereospecific reductions of α,β-unsaturated double bonds[105]. There have been several studies that demonstrate Oye2 mediated reduction of geraniol to citronellol , as evidenced by a 60% increase in geraniol accumulation following *Oye2* deletion, which was further corroborated by a 40% increase in citronellol following *Oye2* overexpression[100,101]. In addition to Oye2, Oye3 is also implicated in direct α,β-unsaturated monoterpenol reduction, wherein a 44% reduction of side products was achieved from a strain lacking both *OYE* homologs, an improvement over the 33% reduction from *Oye2* deletion alone[102].

1.4.2 Introduction and optimization of plant pathways

Initial work towards optimizing the heterologous pathway in S. cerevisiae

Several studies have focused on the heterologous optimization of specific enzymatic steps in the strictosidine pathway[106,107], however, there have only been two studies to date that have succeeded in *de novo* production of either nepetalactol[58] or strictosidine[57] in *S. cerevisiae*. Based on these initial studies, several limiting steps were identified. Brown *et al.* were successful in producing trace amounts (0.5 mg/L) of strictosidine following the optimization of both native and heterologous pathways[57]. Their initial strain was built using the above strategies aimed at increasing precursor supply (ie. overexpression of t*Hgmr1*, *IDI* and introduction of an *Erg20*

mutant) coupled with expressing all the required heterologous genes for strictosidine biosynthesis (Figure 3). When no strictosidine was detected in this strain, they supplied strictosidine pathway intermediates individually to the culture media, from which they identified either geraniol supply or G10H activity to be limiting. In terms of G10H activity, the expression of a second G10H enabled trace amounts (0.011 ± 0.002 mg/L) of strictosidine to be detected. In addition, minimizing geraniol metabolism to off-target products, the deletion of *Oye2* and *Atf1*, an alcohol acetyltransferase (previously shown to act on geraniol[100]), similarly resulted in trace titers of strictosidine. To further increase geraniol supply, *Adh2*, an alcohol dehydrogenase responsible for the conversion of ethanol to acetaldehyde[108] (the acetyl-CoA precursor), was overexpressed in the *OYE* deletion strain, increasing strictosidine to 0.017 ± 0.007 mg/L. The amalgamation of these strategies along with an additional three *G10H* copies (total of 5) resulted in the production of 0.5 mg/L strictosidine[57]. Poor G10H activity was suggested to be the main culprit for limiting strictosidine yields in engineered *S. cerevisiae*.

Campbell *et al.* further demonstrated that the low titers due to poor G10H activity is further aggravated by the promiscuous activity of IS using the pathway intermediates geraniol and 10HG as substrates, in addition to 10-oxogeraniol[58]. IS reduces geraniol and 10HG to non-productive pathway intermediates citronellol and 10-hydroxygcitronellol, respectively[58]. The promiscuous activity of IS was further corroborated by Billingsley *et al.* who demonstrated that exogenously supplied 10HG was converted, almost exclusively (90%), to 10-hydroxytetrahydrogeraniol in an *S. cerevisiae* strain harboring only 10HGO and IS[102]. Two distinct, non-productive pathways were identified where 10HG and 10-oxogeraniol were substrates for 1) 'ene'-reduction and 2) reduction by iridoid synthase in a non-productive redox cascade involving yeast alcohol and aldehyde dehydrogenases[102]. With respect to the first non-productive pathway, both Oye2 and Oye3 individually paired with 10HGO reduced 10-oxogeraniol to 10-hydroxytetrahydrogeraniol. The individual knockout of *Oye2* and *Oye3* decreased 10-hydroxytetrahydrogeraniol accumulation by 33% and 44% respectively, with a 99% reduction observed for their combined knockout. For the second non-productive pathway, endogenous yeast dehydrogenases were demonstrated to work with IS wherein Ari1, Adh6, and Adh7 were found to influence 10-hydroxytetrahydrogeraniol production, with their combined deletion causing a 169% increase in nepetalactol in the *OYE* deletion background strain[102].

The specific activity of IS was further investigated following the elucidation of its crystal structure in complex with $NADP^+$/8-oxogeranial[107]. Ninety-five single-site point mutations were individually introduced into the substrate binding pocket, as well as the extra cavity of IS, to assess the catalytic mechanism and substrate specificity. These mutations resulted in a broad spectrum of IS catalytic activity on 10-oxogeraniol with activities ranging from <5% to 100% of the wild type enzyme activity, with only 25 of the 95 mutants retaining at least 65% activity[107]. These results implicated a non-conserved hydrophilic Ser349 residue and the hydrophobic scaffold at the substrate binding pocket to be determinant of substrate specificity. To this extent, certain point mutations may be capable of mitigating IS promiscuity, however since only 10-oxogeraniol was tested as a substrate, further research will need to assess the activities of the mutant library on other substrates.

With these previous studies in mind, several optimization strategies can be employed in *S. cerevisiae* that can build off this previous work to further optimize nepetalactol and strictosidine production. As poor G10H activity was implicated in several of these studies and was identified as a major bottleneck, several strategies can be explored to further increase geraniol supply and optimize geraniol hydroxylation activity.

Modifying GES for optimal expression in S. cerevisiae

Though endogenous yeast enzyme activity can reduce GPP to geraniol, specific plant monoterpene synthases are more suitable for the efficient and specific conversion of GPP to geraniol. *In planta,* GES is localized to the plastid via a targeting peptide at the N-terminus which is cleaved upon insertion into the organelle, forming the mature protein. The expression of this signal peptide in *S. cerevisiae* has been suggested to interfere with enzyme activity, as removal of this targeting sequence increases enzyme functionality[106,109]. The introduction of a truncated *Valeriana officinalis* GES (t*Vo*GES) into wild type *S. cerevisiae* yielded 2.7 mg/L of geraniol, an approximate 3.4-fold increase compared to the full length *Vo*GES. Pairing the t*Vo*GES with the overexpression of t*Hmgr1* and *IDI* increased geraniol titers to 7.9 mg/L, with the introduction of an Erg20 double mutant ($Erg20^{F96W\text{-}N127W}$) further increasing yields to 27.5 mg/L[106].

Potential ways to optimize G10H activity

As demonstrated by Brown *et al.*[57], G10H remains a major limiting step in nepetalactol and strictosidine biosynthesis, wherein no current optimization strategies have been attempted. G10H is a monooxygenase belonging to the extensive CYP superfamily of hemoproteins. This cytosol-facing, N-terminally bound ER protein catalyzes the C10-hydroxylation of geraniol, requiring the transfer of two electrons catalyzed by the ER bound, NADPH-dependent CPR[110–114]. Several studies have identified factors affecting heterologous CYP expression in yeast such as CYP:CPR ratio, CPR choice and protein localization. As alluded by the multiple CYP copies required in the work by Brown *et al.*[57], optimal CYP:CPR ratios, which exist in nature, can be upwards of 15:1[115,116]. For example, Biggs *et al.* optimized heterologous production of the Taxol precursors, oxygenated taxanes, by increasing the CYP:CPR ratio to 12:1, achieving the highest reported titers of oxygenated taxanes (570 ± 45 mg/L)[116]. This high ratio may ensure a type of logistic control required for efficient electron transfer from the CPR to the CYPs[117]. In addition, CPRs can both generate and deplete reactive oxygen species on their own merit, this low CPR ratio may impede this energetically wasteful process[118]. Increasing CYP copy number likely mimics the optimal conditions found in nature.

This intricate relationship is further demonstrated by the high evolutionary conservation of the CYP and CPR interaction domains, where pairs from different species or kingdoms have the ability to alter, influence or compliment enzyme functionality[117,119–122]. Furthermore, CPR sharing is not uncommon, wherein multiple CYPs receive electrons from a single CPR donor. This is exemplified in plants, in which one to three CPR paralogs are the sole providers of electrons for all CYP reactions[117]. Although *S. cerevisiae* expresses a native CPR, heterologous expression of plant CYPs can offer higher activity upon co-expression of a plant CPR. For instance, co-expression of the *Coleus blumei* CYP (*Cb*CYP) with its cognate CPR in yeast increased the enzymes' hydroxylation activity up to 7-fold over the endogenous yeast CPR[123] . The observed activity of *Cb*CYP in the presence of the native yeast CPR alone demonstrates CYP/CPR functionality across kingdoms, however, the increased activity in the presence of a plant CPR may suggest a more productive relationship between enzyme partners within the same kingdom. More recently, a study combinatorially tested various plant CYP-CPR pairs for their effect on steviol biosynthesis in *S. cerevisiae*[119]. Their expression system assessed the sequential

oxidation and hydroxylation of kaurene by two individual CYPs, kaurene oxidase (KO) and kaurenoic acid hydroxylase (KAH), from two plant species (*Stevia rubaudiana* (*Sr*) and an unnamed species) paired with various CPRs on steviol biosynthesis. Despite only certain KO variants from both species displaying marginally higher activity in the presences of a plant CPR compared to the endogenous yeast CPR, all KAHs were generally more productive with the co-expression of plant CPRs. In addition, both the KO and KAH from the unnamed species outperformed the *Sr*KO and *Sr*KAH when paired with the *Sr*CPR by 10% and 30% respectively. These results corroborate the observed increase in activity for *Cb*CYP in the presence of a plant CPR and further expand on the potential of enzyme pairs from different species to influence catalytic functionality. Additionally, their final steviol strain harbored a 2:1:1 ratio of KO:KAH:CPR, reinforcing the necessity to optimize CYP:CPR ratios[119].

Another critical element for of CYP-CPR function is proper ER localization, mediated by ER-targeting N-termini, which is required to maintain CYP-CPR proximity for efficient electron transfer[113,124,125]. Functional heterologous expression of plant CYPs in yeast cells may be influenced by their N-terminal anchor sequence. Previous studies have shown that, when heterologously expressed, a non-functional CYP can gain activity if its N-terminus is exchanged with one from a CYP shown to be functionally expressed in *S. cerevisiae*[124,126]. Larbat *et al.* were able to characterize monooxygenase activity from their plant CYP, psoralen synthase, in *S. cerevisiae* following the exchange of its N-terminus for that of a functionally active CYP[124]. Prior to the swap, no monooxygenase activity was detected, presumed to be caused by insufficient membrane association, or to low translational efficiency caused by an unfavorable codon bias on yeast expression[124]. However, exchange of the N-terminus for that of a cinnamate 4-monooxygenase, shown to be expressed in yeast[127], allowed for initial characterization of substrate specificity and enzyme kinetics.

Furthermore, in eukaryotic systems, CYPs and CPRs are their own entity both independently localizing and inserting into the membrane via their N-termini. However, an increasing number of bacterial CYP–CPR fusion enzymes have been identified in nature such as the fusion CYP-CPR (BM3) isolated from *Bacillus megaterium* displaying the highest reported monooxygenase activity for a CYP enzyme[128,129]. Since their discovery, a large number of CYP–CPR fusion enzymes have been created by genetic engineering[128]. Artificial fusions of

eukaryotic CYP-CPR enzymes can produce efficient, catalytically self-sufficient entities. The proximity of the reductase to the CYP could enhance the frequency of productive interactions, hence increase the electron transfer and overall turnover rates of the engineered CYP system[128]. Using BM3 as an architectural example, Leonard *et al.* improved the low turnover rate of their heterologously expressed plant CYP, amassing a 20-fold increase in isoflavone synthesis, demonstrating the importance of domain interaction between the cytochrome P450 and its redox partner protein[130]. Although the CYP-CPR fusion is juxtaposed by the high CYP to CPR ratios generally found in nature, a fusion may alleviate the issue with 1:1 ratio by decreasing off-target CPR reactions.

1.5 Strategy for optimizing 10HG production in *S. cerevisiae*

Nepetalactol production in *S. cerevisiae* would be a large step towards the efficient production of both nepetalactone and strictosidine, ultimately enabling the production of a wide array of high value MIA compounds. Previously, both 10HGO[102] and IS[107] activities have been investigated and scrutinized for optimizing *in vivo* nepetalactol production in *S. cerevisiae*, the final step requiring improvement is the C10 hydroxylation of geraniol by G10H.

I hypothesized that the optimization of G10H activity via enzyme variant assessment and catalytic efficiency adjustments, coupled with mitigating off-target reactions would increase substrate turnover, thereby improving 10HG production. The optimization strategies first involved the production of the monoterpene geraniol using a combination of MVA enzyme overexpressions in a strain harboring an $Erg20^{K197E}$ mutation. Next, a G10H enzyme variant library, followed by modification of their N-terminus was assessed for their geraniol hydroxylation activity. Furthermore, the expression of various non-cognate CPRs was characterized for their ability to modulate 10HG activity. I also assessed the negative effect of promiscuous native *S. cerevisiae* enzymes on heterologous product accumulation. This information will be vital to reconstitute and a further optimize an efficient nepetalactol pathway in yeast.

2. MATERIALS AND METHODS

2.1 Strains and media

All *S. cerevisiae* strains built in this study are listed in Appendix Table A1. Yeast cultures were grown at 30°C with shaking at 200 rpm in YPD (10 g/L yeast extract, 20 g/L tryptone, 20 g/L dextrose (Thermo Fisher Scientific)) or synthetic complete (SC) medium made with 6.8 g/L Yeast Nitrogen Base (YNB) with amino acids and supplemented with 2% glucose. To maintain selection for the Erg20^{K197E} background, strains were under constant selection with 200 µg/ml geneticin (G418) (Sigma Aldrich). When appropriate 200 µg/ml of hygromycin was added for the selection of the pCas plasmid. All plasmids used in this study are listened in Appendix Table A2 and were maintained and propagated in *Escherichia coli* DH5α. *E. coli* cultures were grown at 37°C with shaking at 200 rpm in Lysogeny Broth (LB) (Sigma Aldrich) supplemented with either 100 µg/mL of ampicillin or 100 µg/mL of hygromycin.

2.2 DNA manipulation

2.2.1 Plasmids and genes

All genes and primers used in this study are described in Appendix Table A3 and Appendix Tables A4-A10 respectively. Promoters, terminators and the yeast MVA genes, *IDI* and truncated *Hmgr1* (*tHmg1*), were amplified from *S. cerevisiae* CEN.PK genomic DNA. Genes originating from plant species, apart from the G10H variant library, were amplified from pYES2-GerOH[58]. The pCas plasmid used in this study was previously modified from the Addgene (plasmid #60847) in which the G418 resistance KanMX cassette was exchange for an hphNT1 cassette to confer hygromycin resistance (pCAS-Hyg). G10H variants were codon optimized and synthesized by GeneArt (Thermo Fisher Scientific) whereas the codon optimized *Pk*CPR and *Vv*CPR were synthesized by Twist Bioscience. These genes were further sub-cloned individually into pJet2.1 using the CloneJET PCR cloning kit (Thermo Fisher Scientific) and transformed into *E. coli*. Plasmids were purified from *E. coli* stocks using the GeneJET plasmid mini prep kit (Thermo Fisher Scientific) where the DNA parts were amplified by PCR using Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific), resolved using 0.8% (w/v) agarose gel electrophoresis and individually purified using Qiagen Gel Purification kit (Valencia, CA, USA).

2.2.2 Gene library phylogenetic analysis

The G10H gene library to be used in this study was selected using the *Cr*G10H amino acid sequence as the search query for three separate databases: Plant Metabolic Network[131], 1k Plant Collection[132] and NCBI using DeltaBLAST[133]. From each database, sequences with over 65% homology were selected. This list was consolidated using CD-HIT[134] with a sequence identity cut-off of 90%. Bootstrapped (500 replicates) maximum likelihood phylogenetic trees for the G10H and CPR libraries were built (default settings) following MUSCLE alignment (default settings) using MEGA7[135]. The phylogenetic trees were rooted using the various outgroups, amino acid sequences can be found in Appendix Table A3.

2.2.3 Gene modifications

The N-terminal signal peptide of GES was identified using ChloroP 1.1 software[136] which predicts the chloroplast transit peptide and its possible cleavage site. Following identification, a truncated version was generated by PCR using primers listed in Appendix Table A5. The N-termini of the G10H and CPR variants were identified using the TMHMM server[137]. The desired N-terminus and truncated G10Hs were individually amplified using primers describe in Appendix Table A7 that added overlapping sequences to the 3' end of the N-terminus and the 5' end of the tG10H (Figure 4). The two parts were then assembled by sequence overlap extension (SOE) PCR to generate the final chimera product. The CYP-CPR fusion was similarly assembled using primers from Appendix Table A9 that insert overlapping sequences containing the GST linker sequence and then assembled by SOE PCR.

2.2.4 Linearization of pCAS and gRNA design

For CRISPR-Cas9 directed gene modification or insertion, a 20 nucleotide guide RNA (gRNA) sequence targets the CRISPR-associated protein, Cas9, to a genomic specific loci where its endonuclease activity introduces a double stranded break[138]. By supplying an exogenous DNA fragment with regions of homology flanking the double-stranded break, yeast homologous recombination can initiate homology-directed repair, resulting in the integration of the desired DNA repair template[139,140].

The pCas-Hyg plasmid was isolated from *E. coli* and linearized following a double digest with *Bg*lI and *Bg*llI restriction enzymes (NEB). Synthetic gRNA targeting sequences (N_{20}) towards previously characterized genomic loci[141,142] were generated with a minimum GC content of 40% and queried against the *S. cerevisiae* genome to ensure the target sites were unique[143]. The gRNA expression cassette was amplified in two individual parts using the pCas vector backbone as a template and primers that insert overlapping sequences containing the novel N_{20} gRNA target sequence (Appendix Table A4). These two parts were then assembled by SOE PCR to generate the final gRNA expression cassette.

2.3 Yeast transformation and CRISPR/Cas9-mediated genome integration

S. cerevisiae strains were transformed using a method modified from the Gietz PEG/LiAc protocol[144] and were genetically modified using the CRISPR-Cas9 system[138,145]. Strains were initially grown overnight in YPD + G418 medium. The strains were then diluted to an optical density (OD_{600}) of 0.05 and grown to a final OD_{600} of 0.6 to 0.8. Two mL of culture per transformation was harvested and washed initially in sterile water and then in 100 µM lithium acetate. The cell pellet was then suspended in 3M lithium acetate (5.6 µL per transformation) and DNA + water (total volume of 40 µL) and incubated at room temperature for 10 minutes before adding the transformation mix, which included per reaction: 100 µL 50% (w/v) PEG3350 and 5 µL boiled salmon sperm DNA (10 mg/mL). Transformation conditions included a 15-minute incubation period at 30°C, followed by a 30 minute heat shock at 42°C. The cells were recovered in 500 µL of YPD before plating on selective medium contain hygromycin and G418 to maintain the pCas plasmid and select for the $Erg20^{K197E}$ genotype.

DNA present in the transformation reactions included: linearized pCas vector backbone harboring *cas9* (300ng), the gRNA expression cassette (600ng) and the donor DNA (1-4 µg) with homology to the desired integration site. In certain instances, UP and DOWN regions targeting adjacent genomic regions to the desired integration site were included, providing extra ~ 400 bp homology regions to the donor DNA. The UP and DOWN regions were amplified from *S. cerevisiae* CEN.PK genomic DNA using primers described in Supplemental Table S10. These primers add an LV3 linker[146] to the 3' UP region and an LV5 linker[146] to the 5' DOWN region.

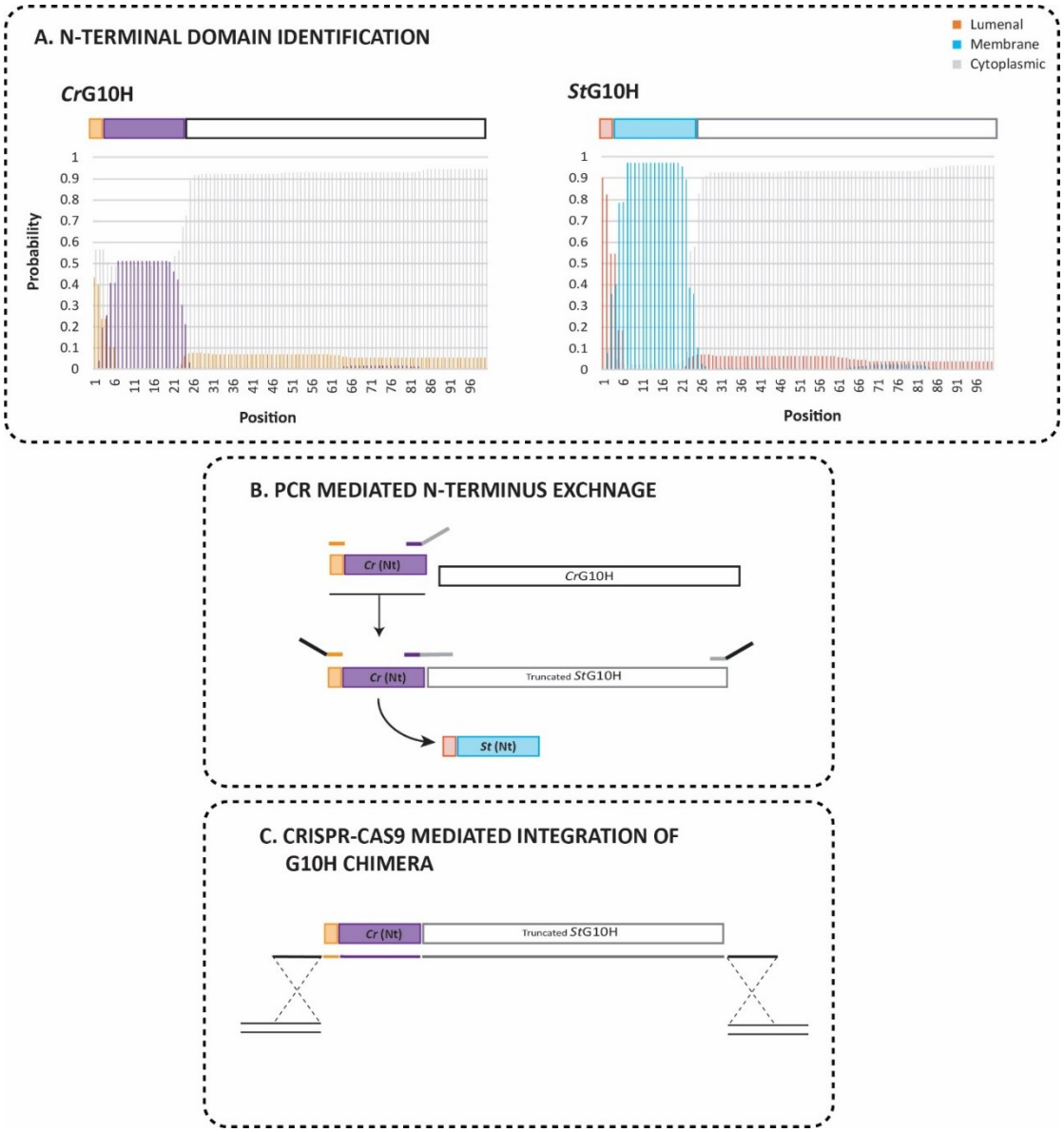

Figure 4. Design of the N-terminal exchange for the G10H variants.
The N-termini were exhanged to assess their infuence on G10H activity. The strategy employed in this experiment was as follows: A) The transmembrane regions were identified using TMHMM server where the probability of luminal (orange), transmembrane (TM, blue) and cytoplasmic (grey) domains were assessed. B) Following TM identfication, the N-termini and truncated G10Hs were individually amplified with specific primers that added homology for subseqent PCR reactions and integration. C) Integration of the G10H chimeras.

These linkers are complimentary to the PCR added LV3 and LV5 linkers flanking the donor DNA.

2.4 Colony PCR

Individual colonies were suspended in 40 µL of sterile water. Twenty µL of the suspended colony was heated in a microwave oven for two minutes, from which 1.5 µL was used as template for a 5 µL colony PCR using Phire™ Plant Direct PCR Master Mix (Thermo Fisher Scientific). Primers used for the colony PCR were site specific and annealed to genomic DNA adjacent to the integration site. PCR products showing positive integrations were identified by agarose gel electrophoresis.

2.5 Microscopy

Cells were grown overnight at 30°C in SC media, reinoculated at a 1/10 dilution and incubated at 30°C for 4 hours. Log phase cells were then imaged on a Zeiss Axioplan microscope equipped with an Infinity 3 camera and a Ph3 Plan-NEOFLUAR 100x/1.30 Oil objective using Infinity Capture Software. All images were acquired using either a 750ms or 900ms exposure and evenly processed in Adobe Photoshop CC.

2.6 Analysis of Monoterpenes (GC-MS)

To analyze monoterpene production, metabolites from cultures of the various recombinant yeast strains were analyzed by GC-MSD. Strains were inoculated in triplicate to an OD_{600} of 0.05 in a 96-well deep well plate containing 500 µL of YPD + G418 medium and incubated at 30°C and 200 rpm shaking for either 24, 48 or 72 hours. For sample preparation, 400 µL culture aliquots were transferred to a new 96-well deep well plate and were extracted into 0.5 volumes of ethyl acetate containing 10 mg/L of eugenol (Sigma Aldrich) as an internal standard. Extracts were analyzed by GC-EI-MS using an Agilent 6890N GC system coupled to an Agilent 5875C mass selective detector. Separation, following splitless injection, of 1 µL of extract was carried using an HP-5ms column (30 m ×0.25 mm × 0.25 µm film thickness) and hydrogen gas as the carrier with a constant flow of 1.3 mL/min and an inlet temperature of 240°C. Separation conditions were as follows: initial oven temperature set to 60°C for 2 minutes, then an increase to 150 at 30°C/min, then increase to 220 at 10 °C/min and then a final increase

to 325 at 30°C/min (hold for 5 min). For monoterpene quantification, the Total Ion Chromatogram (TIC) for the positive ion electron-impact spectra at 70 eV was recorded in the scan range *m/z* 50–220 and Selective Ion Monitoring (SIM) monitored ions 67, 71 and 91 *m/z*. Total amount of citronellol, geraniol, eugenol and 10HG were determined using calibration curves with a minimum R^2 value 0.98 over the concentration range of 3.90 to 1000 µM. Isopulegol was identified using the NIST 08 standard reference database[147].

3. RESULTS

3.1 Increasing *in vivo* geraniol titers

All titers reported in this thesis are normalized to the yeast culture OD_{600} of 1, unless indicated otherwise. The synthesis of 10HG in *S. cerevisiae* requires the modification of endogenous pathways to enable *in vivo* production of geraniol. *S. cerevisiae* does not naturally accumulate sufficient monoterpene precursor for geraniol production (ie. GPP), hence the modification of the native mevalonate pathway in conjunction with the expression of exogenous plant derived genes is essential for increasing GPP, and subsequently geraniol concentrations. Previously Campbell *et al.* introduced the $Erg20^{K197E}$ mutation[59] into wild type CEN.PK (denoted strain MD1) to allow for *in vivo* GPP production[58]. This strain was used as the foundation for all further strain modifications reported in this present study. GPP is synthesized from two isoprene units that are derived from the mevalonate pathway where carbon flux is limited by *Sc*HMGR and *Sc*IDI. To this effect, a second copy of *tScHmgr1* and *ScIDI* were integrated into the genome to overexpress the MEV pathway, as previously described[77], resulting in strain MD2 which accrued trace titers of geraniol (0.07 ± 0.04 mg/L/OD_{600}) (Figure 5A). Once the isoprenoid precursor supply was increased, *ObGES* was integrated into strain MD2 leading to the accumulation of 1.26 ± 0.60 mg/L/OD_{600} in strain MD3. *In planta*, GES is localized to the plastids via a N-terminal targeting peptide that can interfere with enzyme activity when expressed in *S. cerevisiae*[106]. Exclusion of the predicted signal peptide from *Ob*GES increased geraniol yields in strain MD4 by approximately 3.8-fold, accumulating 4.81 ± 0.52 mg/L/OD_{600} of geraniol. The integration of the *Abies grandis GPPS* (*AgGPPS*) into MD4 further increased geraniol titers by over 2-fold, resulting in strain MD5 and titers of 10.79 ± 0.70 mg/L/OD_{600} geraniol, corresponding to a total accumulation (non-OD normalized) of 43.45 ± 1.84 mg/L after

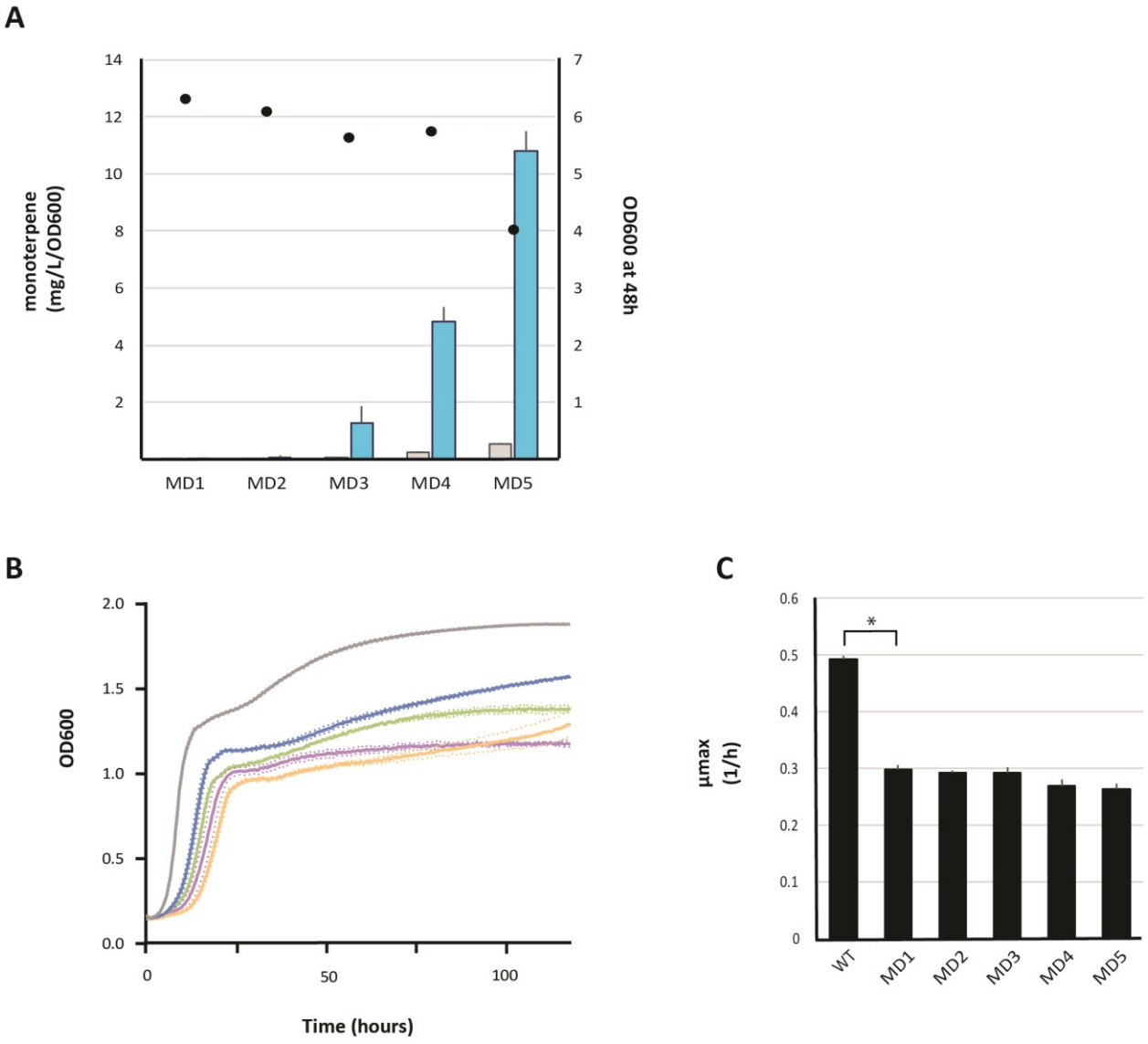

Figure 5. Geraniol titers of engineered *S. cerevisiae* and its influence on strain growth.
A) Citronellol (grey), geraniol (blue) and OD_{600} values (•) of engineered strains. Error bars represent the standard deviation from triplicate cultures. **B)** Comparison of growth profiles for the geraniol production strains. Dotted lines correspond to the standard deviation for quadruplicate cultures. **C)** Specific growth rate (µmax) for engineered strains from quadruplicate cultures. Asterisk indicates statistically significant ($p<0.05$) growth rate compared to the wild type strain as determined by Student's two-sample *t*-test assuming unequal variances with two tailed distribution. For A, B and C strain designation is as follows: WT (—); **MD1,** WT + *Sc*Erg20^{K197E} (—); **MD2,** MD1 + t*Sc*HMGR + *Sc*IDI (—); **MD3,** MD2 + *Ob*GES; **MD4,** MD2 + t63*Ob*GES (—); **MD5,** MD4 + GPPS (—).

48 hours (Appendix Figure 5A). For each round of integrations, citronellol titers were quantified to observe the extent of carbon loss in the form of non-productive reductions. In MD5, a 4.7 percent molar loss of carbon in the form of citronellol was observed, corresponding with 0.54 ± 0.03 mg/L/OD_{600} citronellol titers.

An inverse relationship between cell growth and increasing geraniol titer was observed in these strains with each subsequent round of integrations aimed at increasing geraniol titers (Figure 5A and B). The most notable effect on cell growth can be observed between the wild type CEN.PK strain and MD1, following the introduction of the $Erg20^{K197E}$ mutation. Due to the involvement of the FPP synthase activity of Erg20 in sterol and ubiquinone biosynthesis, it is presumed that decreasing the carbon flux towards these essential pathways will negatively affect cell growth. This effect is shown by a prolonged lag phase followed by a decrease in specific growth rate (μmax) from 0.49 ± 0.00 h^{-1} to 0.29 ± 0.00 h^{-1} between the wild type and MD1, respectively ($p<0.05$) (Figure 5B and C). This impaired growth is further shown by the lower growth yields for MD1 compared to wild type. A yeast culture with an OD_{600} of 1 has an average of 3.0 x 10^7 cells/mL, the introduction of the $Erg20^{K197E}$ mutation decreased growth yields from 5.1 x 10^7 cells/mL to 3.8 x 10^7 cells/mL at 72 hours (Figure 5B). Moreover, a further decrease in cell growth occurs with each consecutive integration aimed at increasing geraniol titers. Though not significant ($p>0.05$), a decrease in growth rate is observed following GES expression (MD4) (Figure 5C). Longer lag phases and lower growth yields similarly coincide with increased geraniol titers, where strain MD4 accumulates 2.9 x 10^7 cells/mL at 72 hours, a 57 percent decrease compared to wild type (Figure 5B).

3.2 Testing G10H variant library for geraniol hydroxylation

Efficient *de novo* synthesis of 10HG in *S. cerevisiae* for nepetalactol and strictosidine biosynthesis requires optimizing the rate-limiting hydroxylation of geraniol by G10H. With the *Catharanthus roseus* G10H (*Cr*G10H) being the only functionally characterized G10H to date, initial strategies to optimize this hydroxylation step involved testing other potential G10Hs for improved activity. Using the *Cr*G10H amino acid sequence as the search query, a preliminary list of predicted G10Hs was established using three separate databases: Plant Metabolic Network[131], 1k Plant Collection[132] and NCBI using DeltaBLAST[133]. Since G10H is a cytochrome P450 (CYP, CYP76b6), the initial search results were extensive owing to the vast sequence diversity

of CYP families. Fifty-five percent sequence identity is sufficient for CYPs to belong to the same subfamily[148]. For the purpose of reducing the list to a more manageable number, sequences with over 65 percent homology were kept. This list was further consolidated to 269 sequences using CD-HIT[134] to minimize redundancy among the sequences with a maximum percent identity set to 90 percent (Appendix Figure A1). This dictates that when there is a cluster of sequences with over 90 percent identity, the sequence that best represents the clade is kept and the rest are discarded. From the condensed preliminary phylogenetic tree, 15 sequences were chosen and verified for the presence of the CYP conserved domains (Figure 6A, Appendix Figure A2). This final list included three sequences isolated from plants known to produce 10HG (red), two of which have already been identified as G10Hs based on plant transcriptomics (*Pk*G10H[149] and *Gri*G10H[150]), and two isolated from plant species known to produce geraniol (blue). The remaining sequences were isolated from plant species from which there is no prior knowledge of geraniol or 10HG production (Figure 6A).

To test the activity of these predicted G10Hs, a CYP reductase (CPR) is required to donate two electrons to catalyze hydroxylation. Accordingly, the *C. roseus CPR* (*CrCPR*) was previously integrated alongside *GPPS* in the making of strain MD5. Each of the predicted *G10H*s was then individually integrated into MD5 and 10HG production was analysed by GC-MS. Analysis showed that *Cr*G10H remained the highest producer of 10HG with titers reaching 2.22 ± 0.24 mg/L/OD_{600} (Figure 6B) following trace amounts observed from *Gr*G10H (0.18 ± 0.04 mg/L/OD_{600}) and *In*G10H (0.13 ± 0.03 mg/L/OD_{600}). However, *In*G10H was plagued with instability between consecutive rounds of analysis wherein varying amounts of geraniol and 10HG were observed and was therefore discarded from further analysis. The majority of strains displayed a reduction in geraniol titers compared to MD5, suggesting geraniol is being used as a substrate, yet the lack of 10HG accumulation may imply the enzymes are involved in a non-C10 hydroxylation reaction. *St*G10H and *Pp*G10H achieved the most notable loss of geraniol titers, showing a near complete reduction. In all instances, the GC-MS chromatograms did not show the appearance of any new product peak. Similar to *In*G10H, *Eg*G10H showed a sporadic loss or gain of geraniol between subsequent rounds of analysis and was discarded from further analysis (Figure 6B). The lack of geraniol hydroxylation activities from the predicted G10H library sparked further investigation into the difference between *Cr*G10H and the predicted enzymes.

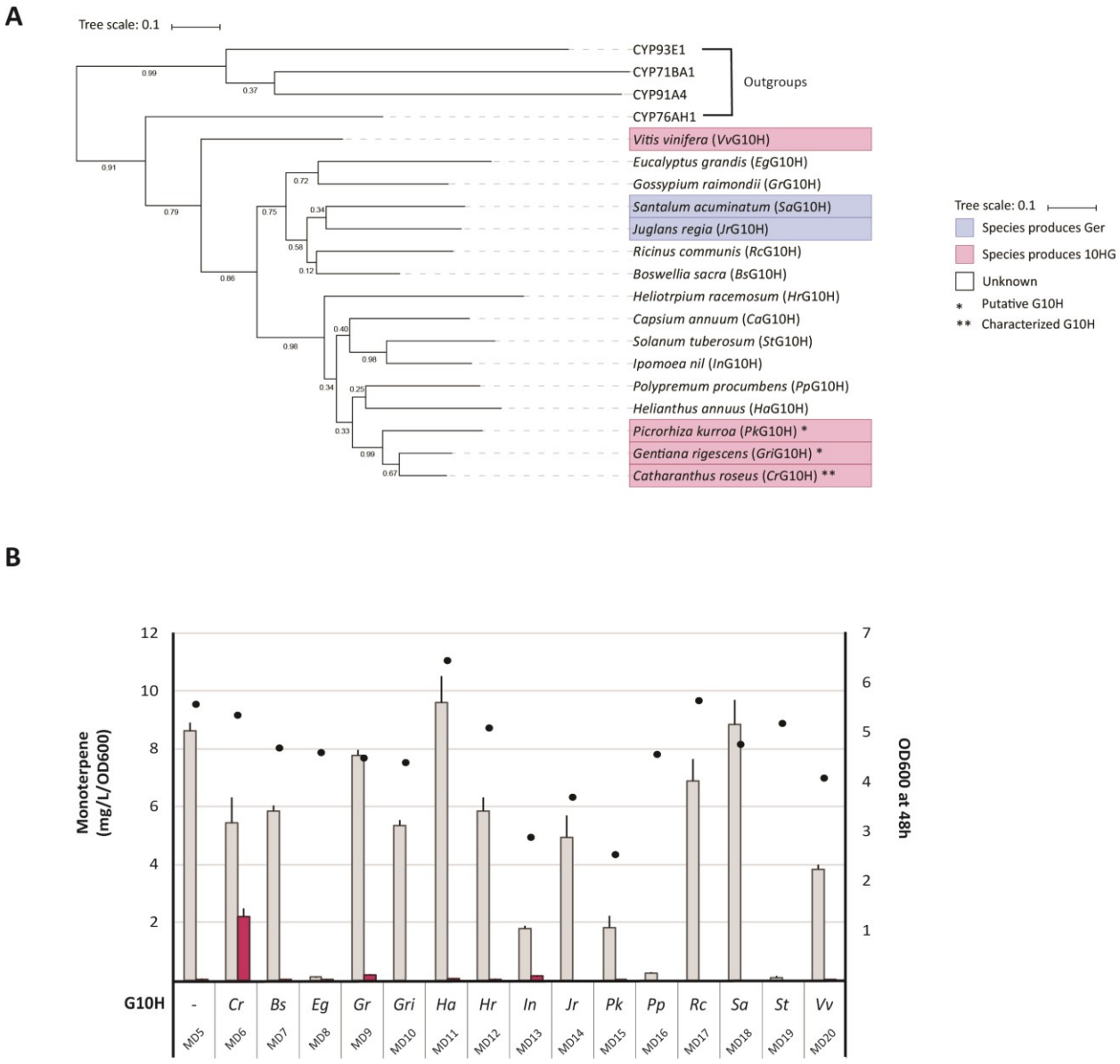

Figure 6. Monoterpene production in engineered *S. cerevisiae* strains harboring G10H variants.
A) A bootstrapped maximum likelihood phylogenetic tree of predicted/putative G10Hs, rooted using the various outgroups. Certain variants were identified from plant species known to produce geraniol (blue) or 10HG (red). The remaining variants belong to plant species wherein the plant oil content has not been characterized. Certain variants have been characterized as a putative G10H (*) whereas others have been functionally characterized (**). **B)** Monoterpene production of the G10H variants in the MD5 background strain. Geraniol (grey) and 10HG (red). Error bars represent the standard deviation from triplicate cultures.

3.3 N-terminal exchange of G10H variants

As the initial screening of the G10H variant library did not identifying any candidate enzymes with improved activity over *Cr*G10H, we pursued the targeted modification of CYP domains. CYPs are N-terminally bound to the ER which is important for functional expression[113,124,125]. Proper ER localization is essential due to the required proximity to the ER bound CPR. Replacing the N-terminus of a functionally impaired CYP for an N-terminus of a CYP shown to have activity in *S. cerevisiae* can stabilize expression[124,126]. Since *Cr*G10H shows geraniol hydroxylation activity in *S. cerevisiae*, its N-terminus was used to replace the N-termini of the predicted G10Hs. Using TMHMM prediction software[137], the transmembrane domains of all the G10Hs were annotated and the new chimeras were designed and assembled (Figure 4). As a control, the *Cr*G10H N-terminus was exchanged for that of *St*G10H which had no observable *in vivo* geraniol C10 hydroxylation activity.

A marginal change was observed in 10HG titers between the wild type *Cr*G10H (4.28 ± 0.63 mg/L/OD_{600}) and the $_{St}$G10H(Nt)*Cr*G10H chimera (3.26 ± 0.44 mg/L/OD_{600}) (Figure 7), suggesting the N-terminus of *St*G10H is not interfering with its expression and is capable of localizing to the ER. This hypothesis can be corroborated by microscopy data wherein GFP tagged versions of both proteins are seen to localize to the ER (Figure 8). It is worth noting the difference in 10HG titers for the *Cr*G10H strain between Figure 6B and Figure 7, which may be attributed to slight differences in culture growth between experiments. Interestingly, the relative signal intensity from $_{St}$G10H(Nt)*Cr*G10H-GFP displays a lower expression profile compared to *Cr*G10H-GFP, contrasting the GC-MS data showing similar 10HG titers. This may indicate an increased substrate turnover for the chimera or increased coupling efficiencies with the CPR since lower protein levels achieve similar titers. Furthermore, the $_{St}$G10H(Nt)*Cr*G10H chimera displays lower levels of geraniol accumulation (0.62 ± 0.05 mg/L/OD_{600}) compared to the *Cr*G10H strain (6.05 ± 0.54 mg/L/OD_{600}). These geraniol levels observed for the chimera seem to remain relatively consistent with the wild type *St*G10H (Figure 6B and 7). However, this result is not consistent among chimeras, where the *Cr*G10H N-terminus seems to alter geraniol titers in most of the G10H variants.

In addition to the activity observed from the $_{St}$G10H(Nt)*Cr*G10H, the $_{Cr}$G10H(Nt)*Gr*G10H and $_{Cr}$G10H(Nt)*Rc*G10H also showed trace levels of 10HG with titers of 0.60 ± 0.21 mg/L/OD_{600} and

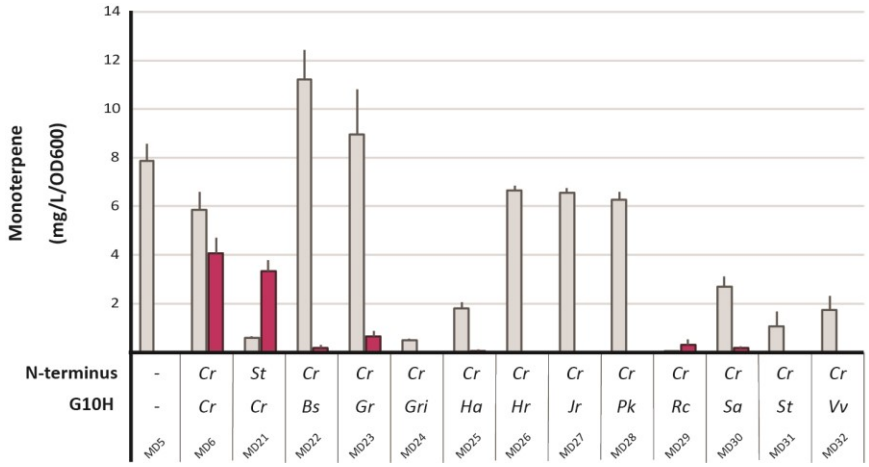

Figure 7. The effect of the N-terminal sequence on geraniol hydroxylation *in vivo.*
The N-termini of the G10H library was exchanged for the N-terminus of the functional *Cr*G10H whereas the *Cr*G10H N-terminus was exhanged with the *St*G10H, a variant lacking hydroxylation. Geraniol (grey) and 10HG (red) titers. Error bars represent the standard deviation from triplicate cultures.

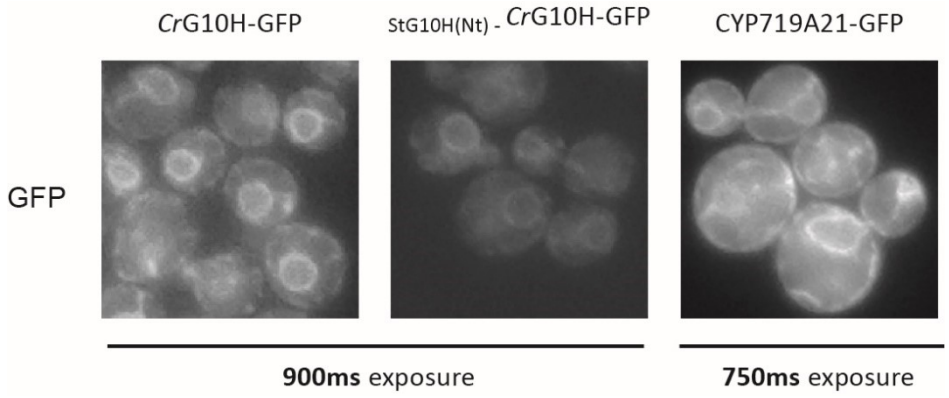

Figure 8. GFP localization of *Cr*G10H and *St*G10H(Nt)-*Cr*G10H.
All constructs were expressed under the TDH3p and were seen to localized to the ER. CYP719A21 was included as a control. **Note: Images were taken by Dr. Bjorn Bean.**

0.24 ± 0.21 mg/L/OD_{600}, respectively. However, the standard deviation observed from *Cr*G10H(Nt)*Rc*G10H between triplicate cultures may be resulting in a misrepresentation of 10HG accumulation. With the lack of observed 10HG from all but two of the chimeric constructs and varying geraniol titers from several strains, the GC-MS spectrums were scrutinized for the appearance of new compounds. A unique peak was identified in the *Cr*G10H(Nt)*Gr*G10H spectrum putatively identified as neric acid (Appendix Figure A3B, peak 5), which was not present prior to the N-terminal exchange (Appendix Figure A3A). Though no increase in 10HG was observed, the appearance of neric acid suggests the N-terminus can influence enzyme activity.

3.4 Optimizing *Cr*G10H activity in *S. cerevisiae*

Since optimal CYP:CPR ratios and CYP-CPR fusions have been demonstrated to influence substrate turnover[115,116,130], both strategies were assessed for their effect on *Cr*G10H activity. To this effect, *CrG10H* copy number was investigated to identify an optimal CYP:CPR ratio, wherein multiple copies of *CrG10H* were iteratively integrated into the single copy *CrG10H* strain. Consistency between each integration was controlled for by ensuring the gene expression profiles at each integration site were comparable, as previously identified[151], and the same promoter-terminator pair was maintained for each integrated *CrG10H* copy. Furthermore, a fusion protein was expressed in the multi-copy *CrG10H* strain that consisted of a full length *Cr*G10H linked to a truncated *Cr*CPR lacking the transmembrane domain.

As was expected, a higher *Cr*G10H:*Cr*CPR ratio improved substrate turnover (Figure 9A, Table 1). A constant increase in 10HG titers can be observed with each additional copy of *CrG10H*, where 3x-*CrG10H* accumulates a total of 6.22 ± 0.27 mg/L/OD_{600} (32.09 ± 6.36 mg/L after 48 hours). Addition of the CYP-CPR fusion into the 3x-*CrG10H* strain further improved titers to 8.49 ± 0.06 mg/L/OD_{600}, reaching 58.24 ± 3.26 mg/L after 48 hours. With each improvement, it can be noted that geraniol reduction coincides with 10HG accumulation where the 3x-*CrG10H* strain increased the molar conversion of geraniol from 37.6 to 87.0 percent (Table 1). A further increase to 92.6 percent molar conversion is seen with the CYP-CPR fusion where geraniol titers reach as low as 0.71± 0.23 mg/L/OD_{600}, with assumption that no geraniol is lost from evaporation or non-productive reactions.

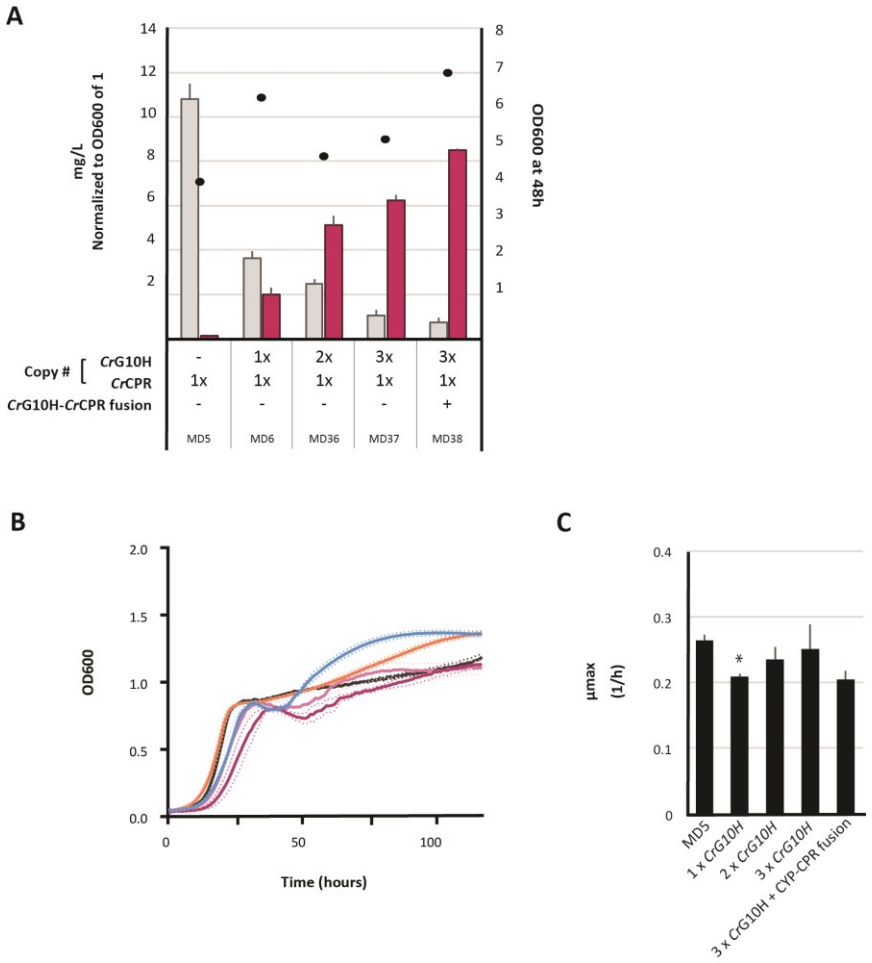

Figure 9. Monoterpene production from strains expressing multi-copy *Cr*G10H and *Cr*G10H-*Cr*CPR fusion. **A)** *Cr*G10H was integrated iteratively into strain MD5. The *Cr*G10H-*Cr*CPR fusion was integrated into the 3x-*Cr*G10H strain. Geraniol (grey) and 10HG (red) titers and OD_{600} (•). Error bars represent the standard deviation from triplicate cultures. **B)** Comparison of growth profiles for the multi-copy *Cr*G10H and *Cr*G10H-*Cr*CPR fusion production strains. Dotted lines correspond to the standard deviation for quadruplicate cultures. Strain designations are as follows: MD5(—), 1x-*Cr*G10H (—), 2x-*Cr*G10H (—), 3x-*Cr*G10H (—) and the *Cr*G10H-*Cr*CPR fusion (—). **C)** Specific growth rate (µmax) for engineered strains from quadruplicate cultures. Asterisk indicates a statistically significant ($p<0.05$) growth rate compared to MD5 strain as determined by Student's two-sample *t*-test assuming unequal variances with two tailed distribution.

Table 1. Percent molar conversion of geraniol to 10HG for the multi-copy *Cr*G10H and *Cr*G10H-*Cr*CPR fusion strain.

G10H	CPR	Geraniol [uM]	10-hydroxygeraniol [uM]	Total [uM]	% molar conversion
*Cr*G10H	*Cr*	21.4	12.9	34.2	37.6
2x *Cr*G10H	*Cr*	14.6	33.1	47.7	69.4
3x *Cr*G10H	*Cr*	6.0	40.3	46.4	87.0
3x *Cr*G10H + *Cr*G10H-*Cr*CPR fusion	*Cr*	4.2	55.0	59.2	92.6

With respect to culture growth (ie. OD_{600}) after 48 hours, it appears that the CYP:CPR ratio impacts culture cell-density and therefore cell growth (Figure 9A). A decrease in cell density can be observed after the addition of a second and third copy of *CrG10H* compared to the 1x-*CrG10H* strain at 48 hours. Although the integration of *CrG10H* lowers the growth rate compared to MD5 ($p<0.05$), little change in growth rate is observed following multi-copy integration compared to 1x-*CrG10H* ($p>0.05$) (Figure 9C). A look into the growth profiles demonstrates that each subsequent increase in copy number results in a prolonged lag phase and lower growth yields, with a decrease from 3.2×10^7 cells/mL to 2.7×10^7 cells/mL for 1x-*CrG10H* and 3x-*CrG10H*, respectively at 72 hours (Figure 9B). Furthermore, the OD_{600} of the multi-copy strains drop directly proceeding diauxic shift, possibly due to toxic product build-up. It appears that the incorporation of the CYP-CPR fusion further alters the growth phenotype, where culture densities at 48 hours are comparable to those expressing 1x-*CrG10H* (Figure 9A). The growth profile of the CYP-CPR fusion strain shows similar lag and exponential phases as those expressing the 2x-*CrG10H* but post diauxic shift (around 48 hours) shows a notable change in growth. This change in growth may indicate the alleviation of toxic product build-up, which is corroborated by the increased molar conversion for this strain resulting in lower geraniol titers (Table 1). Since the results in Figure 9A show a continual increase in 10HG synthesis with no observable limit of production following subsequent *CrG10H* gene additions, the ideal ratio has yet to be determined. These results demonstrate the importance of optimizing the proper CYP:CPR ratio. Furthermore, although the inclusion of CYP-CPR fusion benefits 10HG synthesis in the strain expressing 3x-*CrG10H*, its individual influence should be assessed to determine whether a fourth copy of a *G10H* alone would have the same effect on 10HG synthesis.

3.5 Combinatorial CPR analysis

Owing to the high conservation of CYP-CPR interactions, sets of non-cognate enzymes from different species or kingdoms can pair effectively and display unique enzyme functionality[117,119]. In other cases, two enzymes may not pair properly. Thus, the lack of activity seen for all predicted G10Hs could be caused by a poor interaction with the *Cr*CPR partner. A small library of plant CPRs was previously established in our lab with 21 sequences from various plant species. The amino acid sequences were aligned with the *Cr*CPR and two cognate CPRs

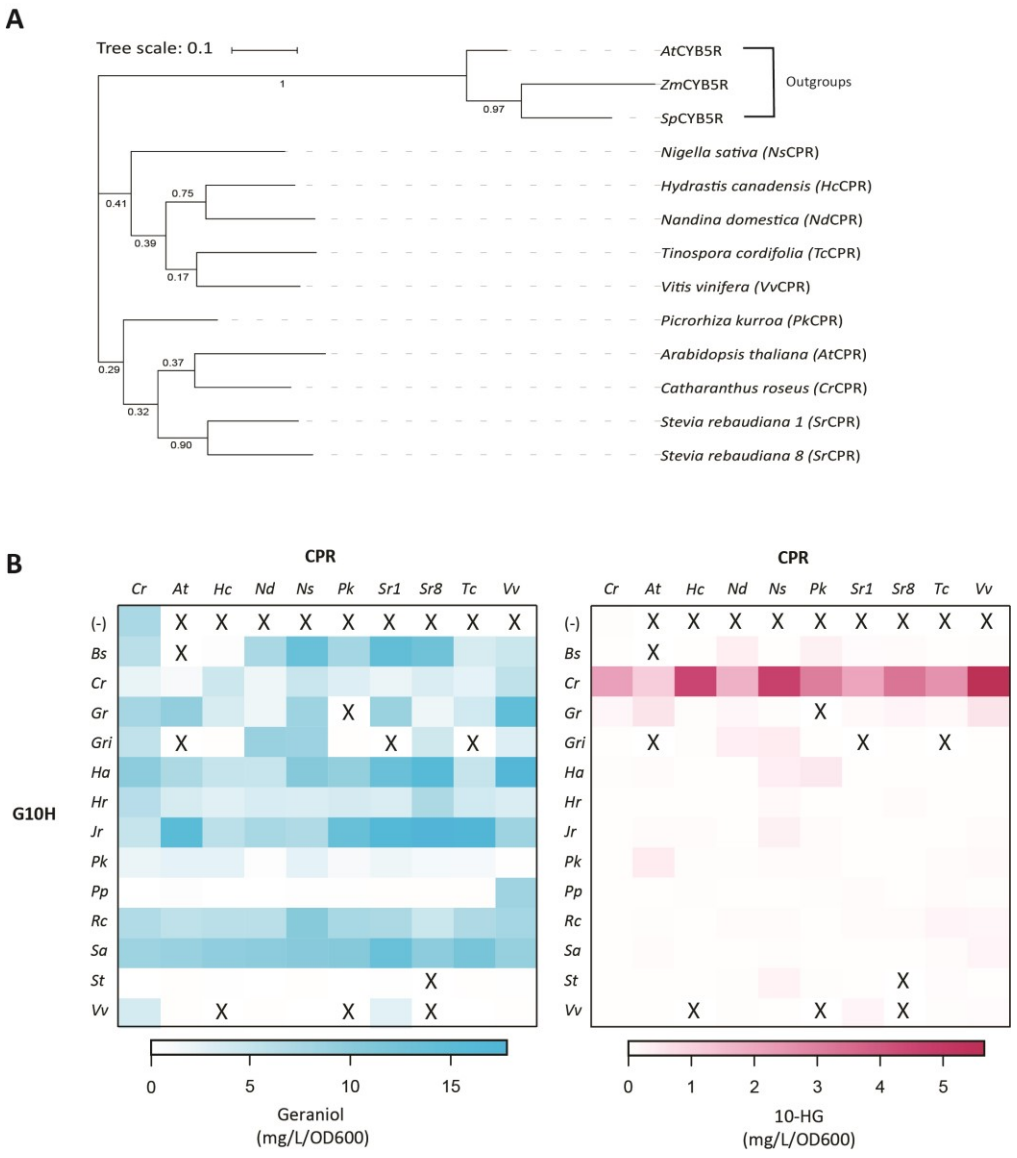

Figure 10. Analysis of CYP-CPR pairs on monoterpene production.
A) A bootstrapped maximum likelihood phylogenetic tree of the CPR library rooted by the cytochrome b5 reductase (CYB5R) outgroups. **B)** Monoterpene production. Geraniol (blue) and 10HG (red) production where the results are displayed in the above heatmaps. Squares with an 'X' represent combinations that were not constructed. The data is based on triplicate cultures, average values and standard deviations are in Appendix Tables A11 and A12.

(*Pk*CPR and *Vv*CPR) for G10Hs in the predicted library. A final list of nine candidate CPRs was selected (Figure 10A). The *CrCPR* was exchanged combinatorially with the candidate *CPR*s in all G10H background strains, and the functionality of each combination was assessed. This resulted in a total of 130 different combinations (including the *Cr*CPR combinations); albeit, nine combinations were not successfully constructed (Figure 10B, squares with an 'X' excluding the (-) G10H controls).

The results from the combinatorial CPR exchange conform with previous reports that different CPR partners can influence CYP activity. Varying degrees of geraniol and 10HG titers can be observed when the G10Hs are paired with different CPRs (Figure 10B). In both heatmaps, the results for the *Cr*CPR strains are displayed in the far-left column with MD5 included as a control in which the *Cr*CPR is present but no G10H. The CPRs seem to have a dramatic effect on metabolite profiles in comparison to the *Cr*CPR. Most notably are the increases in geraniol titers, with several strains improving by over 2-fold, which is especially apparent for *Jr*G10H in which geraniol accumulation increased from 5.15 ± 0.74 mg/L/OD_{600} with the *Cr*CPR to 17.46 ± 1.06 mg/L/OD_{600} and 17.79 ± 1.97 mg/L/OD_{600} with *Tc*CPR and *Sr8*CPR, respectively (Figure 10B, Appendix Table A11A and B). Although an increase in geraniol was observed, this did not correspond to an increased production of 10HG. Furthermore, several of the new combinations resulted in a further decrease in geraniol titers compared to the *Cr*CPR without a corresponding increase in 10HG, the most apparent being *Gri*G10H paired with *Hc*CPR or *Pk*CPR. There are a few observed instances of 10HG production, *Gr*G10H displays an increase in geraniol accumulation that corresponded to the appearance of trace 10HG titers when paired with *At*G10H (0.57± 0.15 mg/L/OD_{600}) and *Vv*G10H (0.57 ± 0.12 mg/L/OD_{600}) (Appendix Figure A11B, Appendix Table A12A and B).

Although several of the new CPRs allowed for trace amounts of 10HG from certain predicted G10Hs, *Cr*G10H remained the most efficient enzyme at C10 hydroxylation of geraniol (Figure 10B, Appendix Table A1 and A2). Additionally, several of the new CPRs increased substrate turnover and 10HG accumulation when paired with *Cr*G10H (Figure 10B and 11). *Hc*CPR, *Ns*CPR, *Pk*CPR and *Sr8*CPR increased the molar conversion of geraniol to 10HG from 26.3% to 83.8, 69.1, 75.9 and 49.1%, respectively (Table 2). Moreover, *Hc*CPR, *Ns*CPR and

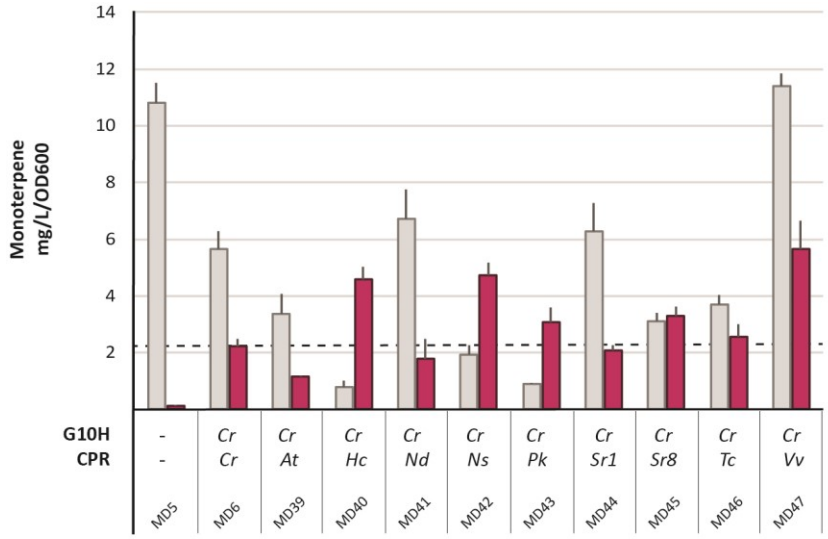

Figure 11. Graphical representation of heatmap for monoterpene production of *Cr*G10H with various CPRs. Geraniol (grey) and 10HG (red) titers for *Cr*G10H paired with non-cognate CPRs. Error bars represent the standard deviation of triplicate cultures and the hashed line is included as a comparison to *Cr*G10H + *Cr*CPR 10HG titers.

Table 2. Percent molar conversion of geraniol to 10HG for different CPRs paired with *Cr*G10H

G10H	CPR	Geraniol [uM]	10-Hydroxygeraniol [uM]	Total [uM]	% molar conversion
Cr	*Cr*	36.6	13.1	49.7	26.3
Cr	*At*	21.7	6.8	28.5	23.7
Cr	*Hc*	5.2	26.8	32.0	83.8
Cr	*Nd*	43.7	10.5	54.2	19.5
Cr	*Ns*	12.4	27.8	40.2	69.1
Cr	*Pk*	5.7	18.1	23.8	75.9
Cr	*Sr1*	40.8	12.2	53.0	23.0
Cr	*Sr8*	20.1	19.4	39.5	49.1
Cr	*Tc*	23.9	15.1	39.0	38.7
Cr	*Vv*	74.0	33.2	107.2	31.0

*Vv*CPR were the most notable in terms of 10HG accumulation with titers reaching 4.57 ± 0.46 mg/L/OD_{600}, 4.73 ± 0.45 mg/L/OD_{600} and 5.65 ± 0.96 mg/L/OD_{600}, respectively. With the identification of several new combinations with improved substrate turnover and product accumulation, the same strain optimization rationale (as shown in Figure 9) was used to further improve 10HG synthesis.

3.6 Assessing the 3x-*CrG10H* strain with non-cognate CPRs

With the identification of several CPRs that increased activity in the 1x-*CrG10H* strain (Figure 11, Table 2), the same optimization rational was implemented as section 3.4. CPRs were selected based on molar conversion, with the assumption that a higher percent molar conversion would be advantageous for 10HG biosynthesis – especially once the appropriate copy number is determined. With a higher molar conversion, less geraniol would be lost to evaporation and further optimization of geraniol synthesis could then improve 10HG titers. It was also hypothesized that 10HG would be less toxic to the cells due to the addition of the hydrophilic hydroxyl group at the end of the hydrocarbon chain, impairing its ability to insert into hydrophobic membranes. To this effect, *Hc*CPR, *Ns*CPR, *Pk*CPR and *Tc*CPR stains were chosen as the pairs with the highest observed molar conversion (Table 2). There were difficulties with integrating the *Vv*CPR in the 3x-*CrG10H* background strain and the results for that combination are not be present in this thesis. Additionally, *Sr1*CPR displayed a lower percent molar conversion compared to *Cr*CPR and was included to observe its effect at higher copy numbers.

Integration of the newly identified CPRs into the 3x-*CrG10H* strain was predicted to further increase 10HG titers in proportion to how they fared compared to the *Cr*CPR in the 1x-*CrG10H* strain. Using the same design as the combinatorial CPR exchange, the new CPRs were integrated into the 3x-*CrG10H* strain. A time course was carried out for all strains with samples taken at 24, 48 and 72 hours with the 1x-*CrG10H* strain as a control. Figure 12 reports the monoterpene profiles after 48 and 72 hours; the 24-hour time point was excluded since the low OD_{600} values resulted in an overestimation of monoterpene titers. The strain harboring *Ns*CPR accumulated the most 10HG at both 48 and 72 hours with yields of 6.98 ± 0.52 mg/L/OD_{600} and 9.48 ± 0.45 mg/L/OD_{600}, respectively. A similar trend was observed for the strain with *Cr*CPR wherein yields increased from 5.81 ± 0.40 mg/L/OD_{600} (48 hours) to 8.17 ± 0.71 mg/L/OD_{600} (72 hours). Unlike the observed trends from *Cr*CPR and *Ns*CPR strains, the *Hc*CPR strain had

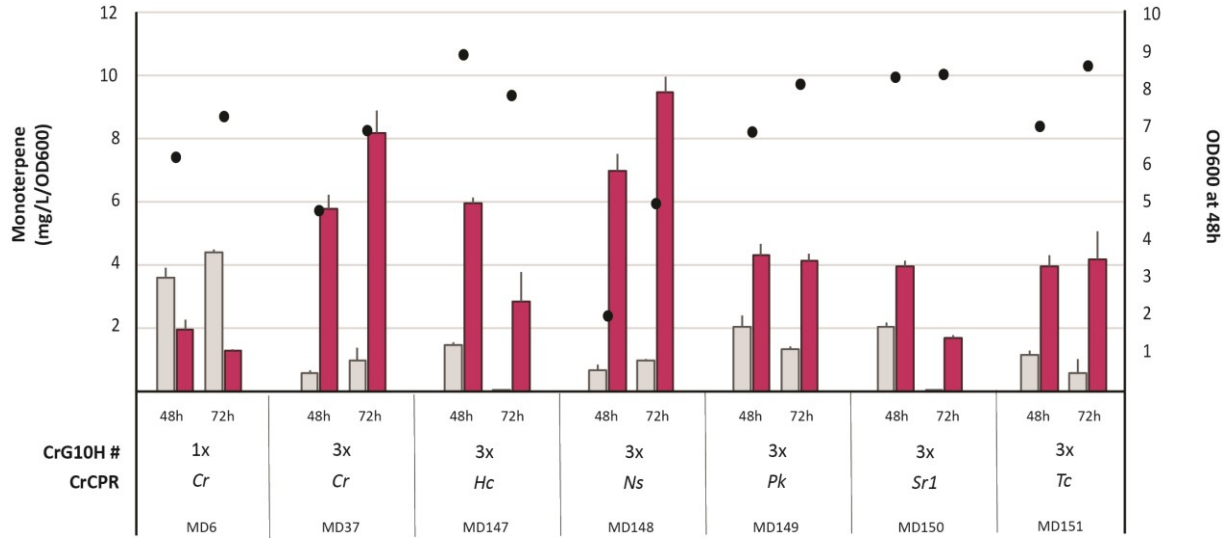

Figure 12. Monoterpene production from different CPRs in 3x-*Cr*G10H strain.
*Sr1*CPR was included as it had a similar percent molar conversion as the *Cr*CPR strain. Geraniol (grey) and 10HG (red) titers and OD_{600} (•) values were measured after 48 and 72 hours, respectively. Error bars represent the standard deviation from triplicate cultures.

relatively high titers at 48 hours (5.95 ± 0.19 mg/L/OD_{600}) but 10HG accumulation diminishes after 48 hours. The remaining CPRs accumulate around 4 mg/L/OD_{600} of 10HG at 48 hours and display similar or decreased titers at 72 hours (Figure 12).

Although *Ns*CPR yielded the highest geraniol hydroxylation when normalized to an OD_{600} of 1, the strain showed low optical densities at both time points. Taking this into account, the *Ns*CPR strain shows a total 10HG accumulation (non-OD normalized) at 48 and 72 hours of 14.13 ± 3.33 mg/L and 47.19 ± 3.62 mg/L, respectively, which is lower than the 27.93 ± 4.43 mg/L and 56.56 ± 4.43 mg/L for the *Cr*CPR strain at both time points (Appendix Figure A4). However, considering the observed error of both the *Ns*CPR and *Cr*CPR strains at 72 hours, their 10HG titers are not statistically different ($p>0.05$). With respect to total accumulation, the *Hc*CPR strain outperforms all strains at 24 and 48 hours accruing 53.09 ± 0.45 mg/L of 10HG. Interestingly, only the *Hc*CPR and *Sr1*CPR strains showed 10HG production at 24 hours with 48 hours showing the highest 10HG levels followed by a decrease at 72 hours.

These results may suggest phenotype stability issues with some of the CPRs, namely *Hc*CPR and *Sr1*CPR, possibly due to degradational control where enzyme activity seems to decrease after 48 hours. Furthermore, these results suggest a difference in copy-number effects between different CPRs, since several of the new CPRs outperformed *Cr*CPR in the 1x-*CrG10H* strain but not in the 3x-*CrG10H* strain. This is in accordance with previous reports of varying optimal ratios between different CYP:CPR pairs[118]. Moreover, these results identified two potential CPRs that may prove to be better partners for the *Cr*G10H.

3.7 Increasing carbon flux towards 10HG

3.7.1 Decreasing off-target reactions

Throughout the optimization process, the presence of an off-target metabolite became more apparent in the GC-MS spectra, coinciding with increased 10HG titers (Figure 13A, peak 3). This peak was predicted to be isopulegol, according to the NIST08 MS database[147], which identifies compounds based on reference mass spectra and compound retention time. However, due to the lack of an appropriate standard, this cannot yet be confirmed. Chemical synthesis of isopulegol usually involves the cyclisation of citronellal using various chemical catalysts[152–154]. In nature, terpene synthases and cyclases have been implicated in the cyclisation of citronellal to

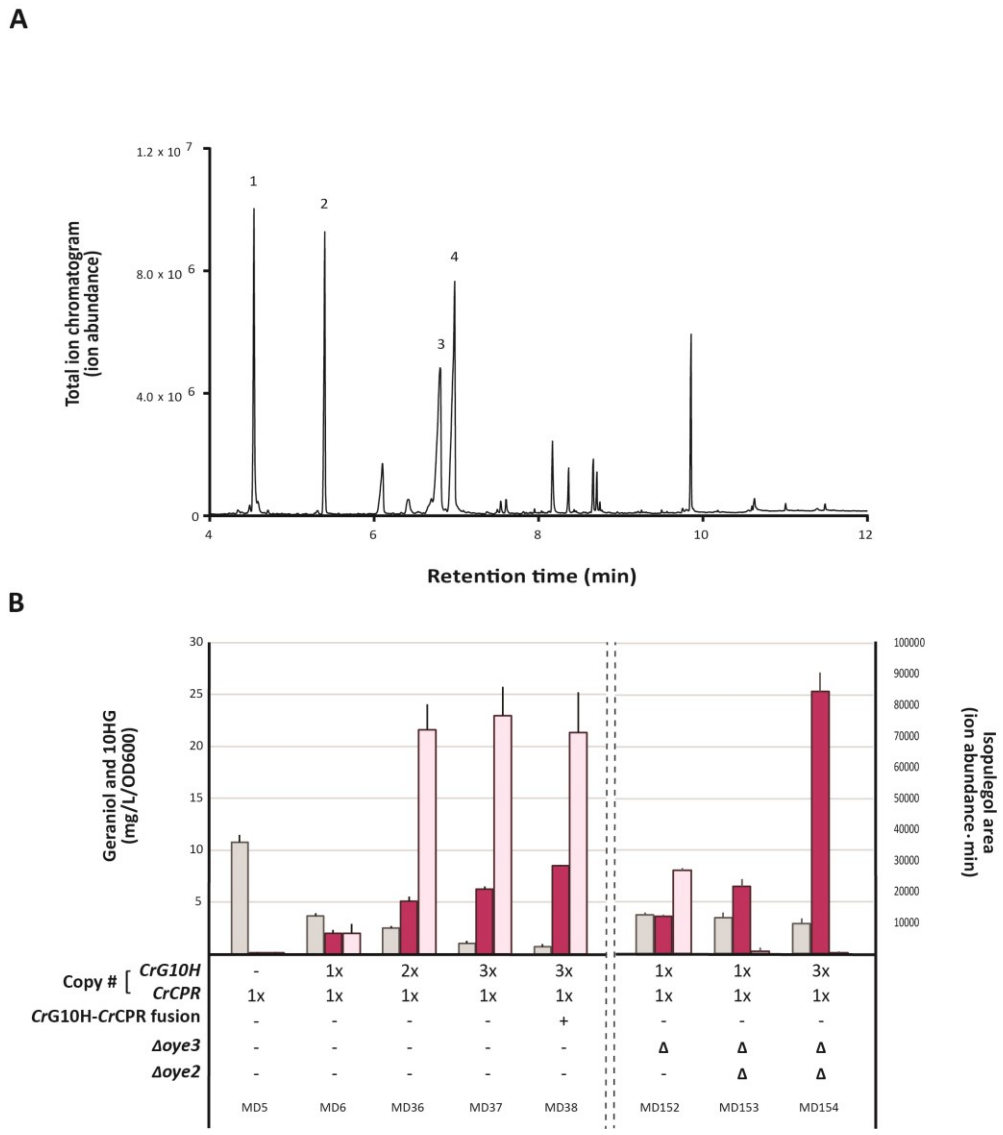

Figure 13. Identification of non-productive synthesis of isopulegol.

A) Total ion abundance for the 3x-*Cr*G10H strain extracted from culture broths. Peak identification using analytic standards are as follows: **1.** Geraniol; **2.** Eugenol (internal standard); **4.** 10HG. Peak **3** was predicted by the NIST08 MS Search as isopulegol. The remaining peaks are standard yeast metabolites. **B)** Monoterpene titers in the multi-*CrG10H* and *Cr*G10H-*Cr*CPR fusion background strains (left) followed by the subsequent elimination of isopulegol content by targeted deletions (right). Geraniol (grey), 10HG (red) and isopulegol (pink). Due to the lack of an appropriate standard, isopulegol is reported as area under the curve. Error bars represent the standard deviation from triplicate cultures.

isopulegol[155,156]. The native yeast old yellow enzymes (Oye2 and Oye3) are NADPH oxidoreductases that have previously been observed to reduce geraniol and citral to citronellol and citronellal, respectively[100,104]. These reductions could provide yeast with the appropriate substrate for isopulegol synthesis. This peak appeared after the integration of *CrG10H* and the presence of multiple *CrG10H* copies further increased isopulegol peak area (Figure 13B). The most notable increase coincided with the addition of the second *Cr*G10H copy, increasing peak area from 6,628.59 ± 3,061.08 ion abundance•min/OD_{600} to 72,037.61 ± 8,219.94 ion abundance•min/OD_{600}. Following this identification, the G10H variant library was analyzed for isopulegol content, however the results indicate that this product is relatively unique to *Cr*G10H with only trance abundances detected from *Gr*G10H, *Hr*G10H and *Rc*G10H (Appendix Figure A5A). Furthermore, the exchange of the *Cr*G10H N-terminus for the *St*G10H counterpart reduced accumulation (Appendix Figure A5B). Isopulegol content also varied depending on the paired CPR. In the 1x-*CrG10H* strain, *Sr1*CPR accumulated more isopulegol compared to the other backgrounds (Appendix Figure A6A), however, in the 3x-*CrG10H* strain, *Hc*CPR displayed the highest isopulegol content (Appendix Figure A6B).

To explore the possibility of Oye2 and Oye3's role in the formation of isopulegol, both homologs were deleted iteratively from the 1x-*CrG10H* strain. Since Oye2 can reduce geraniol and citral to citronellol and citronellal, respectively, *Oye3* was deleted first to determine its effect on isopulegol production. Deletion of *Oye3* was observed to increase both 10HG and isopulegol from 1.99 ± 0.31 mg/L/OD_{600} to 3.62 ± 0.13 mg/L/OD_{600} and 6,628.58 ± 3,061 ion abundance•min/OD_{600} to 26,955.54 ± 523.09 ion abundance•min/OD_{600}, respectively (Figure 13B). Further deletion of *Oye2* eliminated isopulegol production resulting in 6.47 ± 0.74 mg/L/OD_{600} and 25.33 ± 1.81 mg/L/OD_{600} of 10HG in the 1x-*CrG10H* and 3x-*CrG10H* backgrounds, a 3.3- and 4-fold increase, respectively. Corresponding to a total 10HG accumulation (non OD_{600} normalized) of 144.04 ± 5.10 mg/L in the 3x-*CrG10H* strain background.

The effect of deleting these reductases is not only apparent in the monoterpene profiles but an altered growth phenotype is observed upon *Oye3* deletion compared to the 1x-*CrG10H* strain (Figure 14A). This phenotype is synonymous to that observed from the CYP-CPR fusion

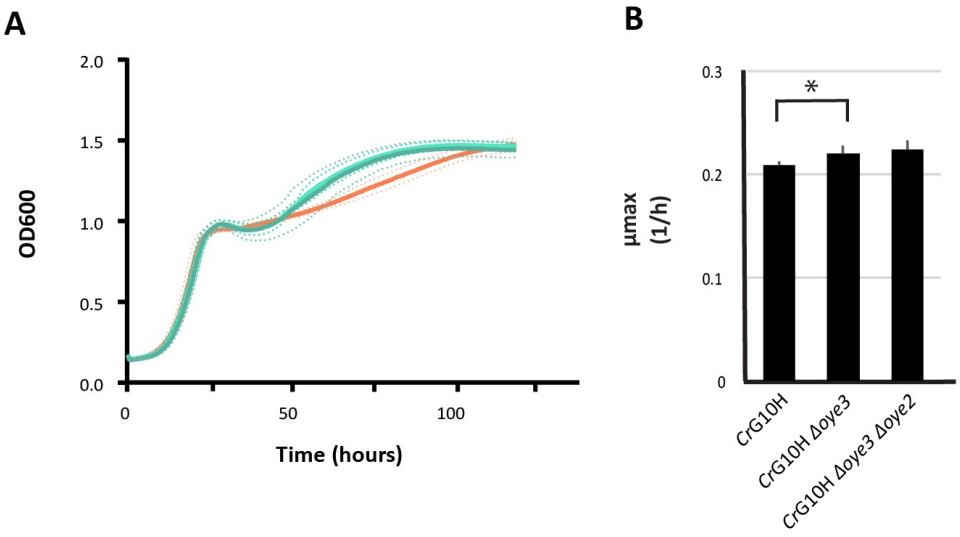

Figure 14. Growth profiles for deletion of *Oye3* and *Oye2* from engineered *S. cerevisiae*.
A) Comparison of growth profiles for *Cr*G10H (—), *Cr*G10H Δ*oye3* (—) and *Cr*G10H Δ*oye3* Δ*oye2* (—) strains. The dotted lines correspond to the standard deviation quadruplicate cultures. **B)** Specific growth rate (μmax) for engineered strains from quadruplicate cultures. Asterisk indicates statistically significant ($p<0.05$) growth rates compared to the strain expressing *Cr*G10H as determined by Student's two-sample *t*-test assuming unequal variances with two tailed distribution.

strain, where a slight increase in growth rate ($p<0.05$) is observed followed by a notable change post-diauxic shift as growth seem to increase much more rapidly in comparison to the 1x-*CrG10H* strain (Figure 14B). The growth phenotype remains consistent following *Oye2* deletion, possibly indicating Oye3 is responsible for the altered growth.

3.7.2 Overall increase in carbon flux towards 10-hydroxygeraniol

The overarching goal in optimizing G10H activity was to increase total carbon allocation towards 10HG biosynthesis by improving inefficient reactions and minimizing off-target products. These initial attempts in optimizing *Cr*G10H activity were successful in channeling more carbon towards 10HG, achieving a modest total increase (non-OD_{600} normalized) from 34.84 ± 3.64 mg/L (1x-*CrG10H*) to 37.47 ± 8.35 mg/L and 63.07 ± 4.66 mg/L for the 3x-*CrG10H* and CYP-CPR fusion strains respectively (Figure 15). Moreover, optimization of the 1x-*CrG10H* strain by mitigating off-target reactions via *OYE* deletions, further improved the total carbon directed towards 10HG. Successive deletions increased 10HG accumulation, and therefore total carbon, with total carbon content yielding 46.82 ± 4.27 mg/L (*Δoye3*) and 63.86 ± 2.36 mg/L (*Δoye2Δoye3*), a 1.3- and 1.8- fold improvement, respectively. Deletion of these reductases from the final 3x-*CrG10H* strain achieved a total carbon content at 160.72 ± 7.56 mg/L. Further deletion of these reductases from the CYP-CPR fusion strain has the potential to significantly increase 10HG titers.

4. DISCUSSION

Optimizing 10HG production in *S. cerevisiae* lays the foundation for a more sustainable and economical means to produce nepetalactone and MIAs. With the elucidation of the strictosidine biosynthetic pathway and the identification of several limiting steps, engineering of a high-producing strain is possible. Optimization of the initial heterologous bottleneck in the nepetalactol/strictosidine pathway in *S. cerevisiae* is dependent on several factors: redirecting carbon from FPP synthesis towards GPP via Erg20 mutation, increasing carbon flux through the MVA pathway bottleneck, optimizing heterologous expression of G10H and reducing off-target reactions. By addressing each area, 10HG production has been optimized for future nepetalactol and strictosidine synthesis.

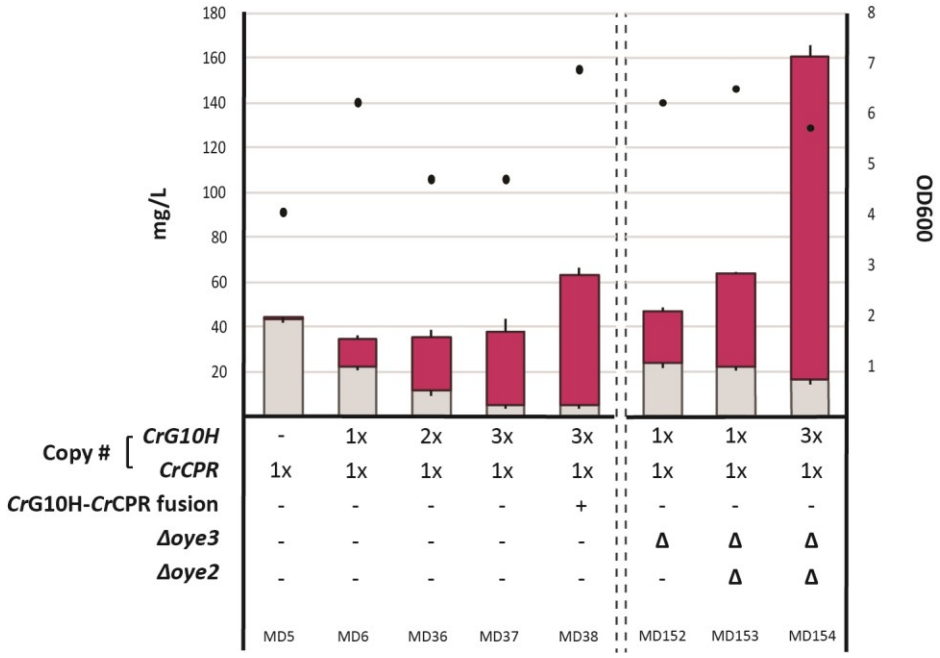

Figure 15. Total monoterpene accumulation for optimized *Cr*G10H strains.
Monoterpene titers in the multi-*CrG10H* and *Cr*G10H-*Cr*CPR fusion background strains (left) followed by the subsequent elimination of isopulegol content by targeted deletions (right). Geraniol (grey) and 10HG (red) and OD_{600} values (•). Error bars represent the standard deviation across experimental triplicates.

4.1 Redirecting carbon for GPP production

The identification of several Erg20 mutations allowing the release of GPP enables the first essential step towards monoterpene production in *S. cerevisiae* – available precursor supply. The Erg20^{K197E} mutation[59] was previously introduced into *S. cerevisiae* CEN.PK, in place of the native Erg20, by Campbell *et al.*[58]. The observable growth defect of this mutant strain can be explained by the essential nature of this enzyme (Figure 5B). With roles in sterol biosynthesis, contributing to membrane structure and cell-wall synthesis, as well as protein prenylation and ubiquinone synthesis, redirecting carbon away from this pathway negatively impacted cell growth[157]. In addition, cell culture growth was further hampered by the toxic nature of geraniol on *S. cerevisiae*. Geraniol is an essential plant oil with antifungal properties, wherein *S. cerevisiae* displays a minimum inhibitory concentration between 200-400 mg/L with 150 mg/L causing a slight inhibition[106,158]. With these concentrations in mind, no growth defects should be observed for MD5 since geraniol titers hovered around 43.45 ± 1.84 mg/L after 48 hours. However, the volatility of geraniol may result in underestimating actual yields, where a study reported an 87.8% loss of radiolabeled geraniol (99% purity) after 21 hours[159]. This volatility is corroborated by the extensive use of biphasic systems aimed at trapping geraniol, as well as other terpenoids, for metabolite analysis. These biphasic systems generally use the alkane hydrocarbon, dodecane, which is well suited for capturing terpenoids and is biocompatible with *S. cerevisiae*[106,109,160–162]. The observance of toxic effects pre- and post- diauxic shift for strains MD4 and MD5 suggest geraniol titers may hover in the 150 mg/L range, however a dodecane biphasic growth system should be optimized to identify exact titers. Regardless of toxicity, initial strain optimizations increased *in vivo* geraniol levels.

4.2 Assessing the activity of the G10H variant library

We hypothesized that geraniol toxicity could be alleviated by hydroxylating geraniol to 10HG by G10H. Since the mechanism of geraniol toxicity involves the permeabilization of cell membranes[158,163], addition of the hydroxyl group may decrease membrane disruption by preventing insertion of the hydrocarbon tail into lipid bilayers. Unfortunately, no 10-hydroxygeraniol was directly observed for the G10H variants aside from *Cr*G10H (Figure 6B). The decrease in geraniol titers for several of the variants suggests the occurrence of non-C10 hydroxylation reactions, where geraniol may be a substrate for secondary reactions involving

oxidations, rearrangements or reductions - activities not uncommon to CYPs[164]. Alternatively, these variants may catalyze sequential reactions initiated by C10 hydroxylation of geraniol, resulting in final products undetectable by GC-MS. A multi-functional CYP was documented by Liu *et al.* that catalyzed several consecutive reactions, beginning with C3 hydroxylation of the sesquiterpene substrate, followed by water elimination, cyclization and regioselective deprotonation[165]. Bifunctional CYPs have been identified in monoterpene biosynthesis that catalyzes sequential stereoselective hydroxylation and carbon–carbon bond cleavage[166]. Conversely, G10H variants that increase geraniol titers in comparison to MD5 may suggest geraniol synthase activity, utilizing GPP as a substrate. This is supported by the previous characterization of a CYP monooxygenase with a secondary terpene synthase active site able to synthesize farnesene isomers from FPP[167]. Analysis of the G10H variant library using the ExPASy ScanProsite tool[168] indicated that a characteristic, highly conserved terpene synthase Mg^{2+} binding motif (**DD***XX*(**D**/**E**))[169] was present in the *Ha*G10H and *Hr*G10H variants, possibly explaining the increase in geraniol accumulation in the *Ha*G10H background. Since the *Hr*G10H strain did not accumulate more geraniol compared to MD5, the lowered geraniol levels indicate the variant may have geraniol-based synthase activity.

After identifying the putative terpene synthase domains, the G10H candidates were systematically searched for conserved domains using NCBI delta-blast[133] which uses domain conservation to identify remote protein homologs. All but two variants (*Hr*G10H and *Rc*G10H) were predicted to contain an allene oxide synthase domain (averaging E scores of $1e^{-7}$), a CYP that catalyses the unusual dehydration of lipoxygenase products to unstable allene epoxides[170]. It is unclear whether this domain can influence geraniol accumulation. Although these variants may not possess C10 geraniol hydroxylation activity and/or may possess other activities, there is the possibility that their heterologous expression is influencing their activity.

The lack of geraniol hydroxylation activity may partially be explained by the nature of the ER membrane in *S. cerevisiae*. While the catalytic domain of a CYP is largely cytosolic, hydrophobic helices embedded the protein in the ER membrane where substrate channels can open both to the cytosol and the membrane adjacent to where these helices interact[171]. Several researchers have characterized the influence of CYP and membrane lipid interactions on enzyme catalytic function and efficiency[172–176]. For instance, the phospholipid composition of the

membrane interacting with the catalytic domain can influence substrate binding and enzymatic activity of CYPs[177]. Murtazina *et al.* observed variations in CYP substrate specificity and enzymatic activity depending on the phospholipid composition of the *in vitro* lipid expression vesicle[177]. Although the phospholipid content of the ER in *S. cerevisiae*[178] falls within the range of plant ER phospholipid content, variations in plants are observed across tissues and species[179]. Furthermore, CYP activity can be affected by the specific N-terminal anchor sequence; influencing CYP-membrane interactions and thus the interaction with membrane phospholipids[176].

4.3 Assessing differential activity from chimeric G10H variants

The CYP N-terminal transmembrane domain is important for protein orientation, stabilizing protein interactions with membrane phospholipids, thereby affecting the catalytic activity of heterologously expressed proteins[124,176,180]. To this effect, the N-termini of the CYP variants were exchanged for that of *Cr*G10H, an enzyme with 10HG activity in *S. cerevisiae*. Unfortunately, the chimeras displayed no new identified hydroxylation activities (Figure 7), although the *Cr*G10H(Nt)*Gr*G10H strain did start producing a compound predicted to be neric acid (Appendix Figure A3). Neric acid is derived from the monoterpene nerol following the dephosphorylation of neryl pyrosphsphate, the *cis*-isomer of GPP[181]. This peak was not detected in the *Gr*G10H chromatograms prior to the N-terminal exchange, suggesting an altered arrangement imparted nerol synthase activity. However, since *Gr*G10H does not have the characteristic terpene synthase domains, it may be acting as an isomerase using geraniol or geranial as the substrate. This altered activity may explain the differences in geraniol yields for several of the variants between their respective wild type and chimeric backgrounds. Changing the N-terminus may alter how the CYP interacts with the membrane, modifying active site accessibility or substrate specificity. This may further explain why the *St*G10H(Nt)*Cr*G10H chimera retains metabolite profiles from both *St*G10H and *Cr*G10H (Figure 6B and 7). In addition, *St*G10H(Nt)*Cr*G10H protein levels appear to be lower than *Cr*G10H, suggesting the N-terminus can modify expression or protein regulation (Figure 8). However, both the wild type and the chimera accumulate similar levels of 10HG (Figure 7), suggesting the chimera is more efficient, possibly due to increased substrate access. Not only does the transmembrane domain influence the

interaction of the protein with phospholipids, but the N-terminus may regulate the efficiency of CYP-CPR interactions and therefore may influence their catalytic efficiency[176,180,182,183].

4.4 CPR influence on CYP activity

Recent studies have postulated that the transmembrane domain composition and the membrane itself may have a role in CYP-CPR coupling efficiency, referring to the amount of transferred electrons committed to the monooxygenase activity versus "uncoupled" (or unproductive) pathways. The orientation of the CYP and the CPR, dictated by their respective N-termini, can influence unproductive reactions, namely, release of superoxide from the ferrous dioxygen, release of hydrogen peroxide, and the $4e^-$ reduction of heme to produce water[184]. The varying degrees of monoterpene profiles between different CYP-CPR pairs in Figure 10B may be explained by coupling efficiencies. Since no observable trend was identified within a single CPR or G10H lineage, the results suggest that coupling efficiencies are depended on the specific CYP-CPR pair. In the case of the strain expressing 1x-*CrG10H*, the non-cognate CPRs increased coupling efficiencies thereby increasing substrate turnover resulting in higher 10HG titers. A further increase in coupling may be attainted with the upregulation of *Ice2*, a type III membrane protein in *S. cerevisiae* involved in maintaining ER morphology[185,186]. Overexpression has been shown to stabilize CPR levels and activity over extended periods of bioconversion, which may benefit prolonged production of 10HG. Furthermore, it would be interesting to assess the effect of the native *S. cerevisiae* CPRs on G10H activity. Nonetheless, with the identification of several CPRs shown to improve efficiency of the CYP catalyzed conversion, further optimization strategies allowed for increased 10HG production.

The introduction of the more efficient, non-cognate CPRs into the 3x-*CrG10H* strain resulted in unexpected differences in molar conversion compared to the 1x-*CrG10H* background (Figure 11 and 12). It was expected that 10HG production would increase proportionally in accordance with the CPRs' influence in the 1x-*CrG10H* strain. However, introduction of these CPRs into the 3x-*CrG10H* strain resulted in a close competition between *Cr*CPR, *Hc*CPR and *Ns*CPR, suggesting a more efficient coupling between *Cr*G10H and *Cr*CPR at a 3:1 ratio. This relates back different optimal CYP:CPR ratios for different enzyme partners, wherein *Hc*CPR or *Ns*CPR may require a higher CYP:CPR ratio for optimal coupling. In this regard, delta-CRISPR could be used as a high-throughput means to assess optimal ratios, wherein multiple *G10Hs*

could be integrated into the Ty retrotransposon delta sites in the yeast genome[187–189]. However, an important aspect remains to be addressed for the 3x-*CrG10H* strain harboring *Ns*CPR, the low OD_{600} for this strain could influence its position as a better alternative to *Cr*CPR.

This observed growth defect may be explained by the N-end rule, relating the *in vivo* half-life of a protein to the N-terminal residue directly after the initial *Met* residue. A hierarchical structure has been identified characterizing the destabilising effect of the N-terminal residue and determining the nature of the N-degron (degradation signals targeted by the N-end rule pathway)[190–192]. For the CPRs tested in the strain expressing 3x-*CrG10H*, *Ns*CPR had an N-terminal residue denoting a protein half-life of >20h, with the remaining CPRs falling below 30 minutes[190]. A short *in vivo* lifetime could imply a less stringent control over the relative rates of synthesis, since proteins produced in excess would not accumulate to a significant level[193]. Since *Ns*CPR has a high *in vivo* half-life, degradation rates are slower compared to the other CPRs and may result in ER flooding, causing cell growth defects[194]. Alternatively, *Ns*CPR may donate electrons at a higher turnover rate but also have a higher degree of uncoupling, potentially resulting in increased H_2O_2 production and cell toxicity. It would be interesting to analyze the effect of different CPR N-end residues on cell growth and 10HG production. Even with the growth impairment, *Ns*CPR remains one of the best candidates for optimal 10HG production along with *Hc*CPR and *Cr*CPR.

4.5 Identification and reduction of off-target metabolism

The last requirement for optimal 10HG production in *S. cerevisiae* was to mitigate off-target products responsible for sequestering carbon away from 10HG biosynthesis. The putative identification of isopulegol suggested a carbon sink that could limit 10HG accumulation. A logical synthesis scheme can be derived involving native yeast enzymes combined with a potential heterologous cyclase, where geraniol or 10HG undergo a sequential oxidation/reduction combination to produce citronellal, followed by cyclisation to isopulegol (Figure 16). As explained earlier Oye2 has been shown to reduce geraniol and citral to citronellol and citronellal, respectively; citronellal has been used extensively as the starting material for direct chemical synthesis of isopulegol[152–154,195,196]. Furthermore, several *S. cerevisiae* alcohol dehydrogenases are capable of oxidizing a primary alcohol to an aldehyde, namely Adh6 and Adh7, which have displayed activity on 10HG, converting it to 10-hydroxgeranial[102]. A recent study developed an

enzyme cascade to selectively convert citral to isopulegol by combining a bacterial ene-reductase OYE homolog and a squalene hopene cyclase[197]. To this effect, three possible biosynthetic routes can be proposed, two of which begin with geraniol and one branching from 10HG (Figure 16). For route 1 geraniol is reduced to citronellol, followed by oxidation to citronellal. Route 2 involves geraniol oxidation to citral, followed by reduction to citronellal and similarly route 3 branches from 10HG and involves an oxidation to 10-hydroxygeranial and reduction to 10-hydroycitronellal. In the case of route 3, 10-hydroxycitronellal is used as the direct substrate for isopulegol synthesis or undergoes a dehydration reaction to an alkene followed by reduction to citronellal.

The final cyclisation of citronellal to isopulegol may be directly catalyzed by *Cr*G10H since this activity was not observed prior to *CrG10H* integration and isopulegol abundances increase with multiple *CrG10H* copies. Moreover, isopulegol accumulation is limited to the strains expressing *Cr*G10H, apart from two exceptions in which small accumulations can be observed (Appendix Figure A5). Additionally, N-terminal modification of *Cr*G10H influences isopulegol accumulation. With respect to terpene cyclase, *Cr*G10H does not have the characteristic aspartate-rich motifs conserved among this class of enzyme, however, Itoh *et al*. characterized a novel set of terpene cylases similarly lacking these conserved domains[198]. Furthermore, Lichman *et al.* characterized a set of non-oxidoreductive cyclases from *N. mussinii*, NEPS, that are involved in the cyclisation of 10-oxogeranial to nepetalactol and similarly lack the conserved terpene cyclase domains[89]. Therefore, *Cr*G10H may be responsible for the cyclisation, though further analyses will be required for conformation.

4.6 Potential effect of *OYE* deletions on cell growth

The OYE homologs have featured prominently in the non-productive reduction of monoterpene intermediates[100–104]. In this study, Oye2 was directly involved in producing the off-target metabolite, predicted to be isopulegol, as its removal abolished accumulation. In contrast, *Oye3* deletion increased 10HG and isopulegol levels, suggesting Oye3 diverts carbon by converting a pathway intermediate to an unknown product. The *Δoye3* strain displayed pronounced diauxic shifts (Figure 14A), similar to the CYP-CPR fusion, possibly implying Oye3 counteracts geraniol-induced growth defects, increasing cell survival.

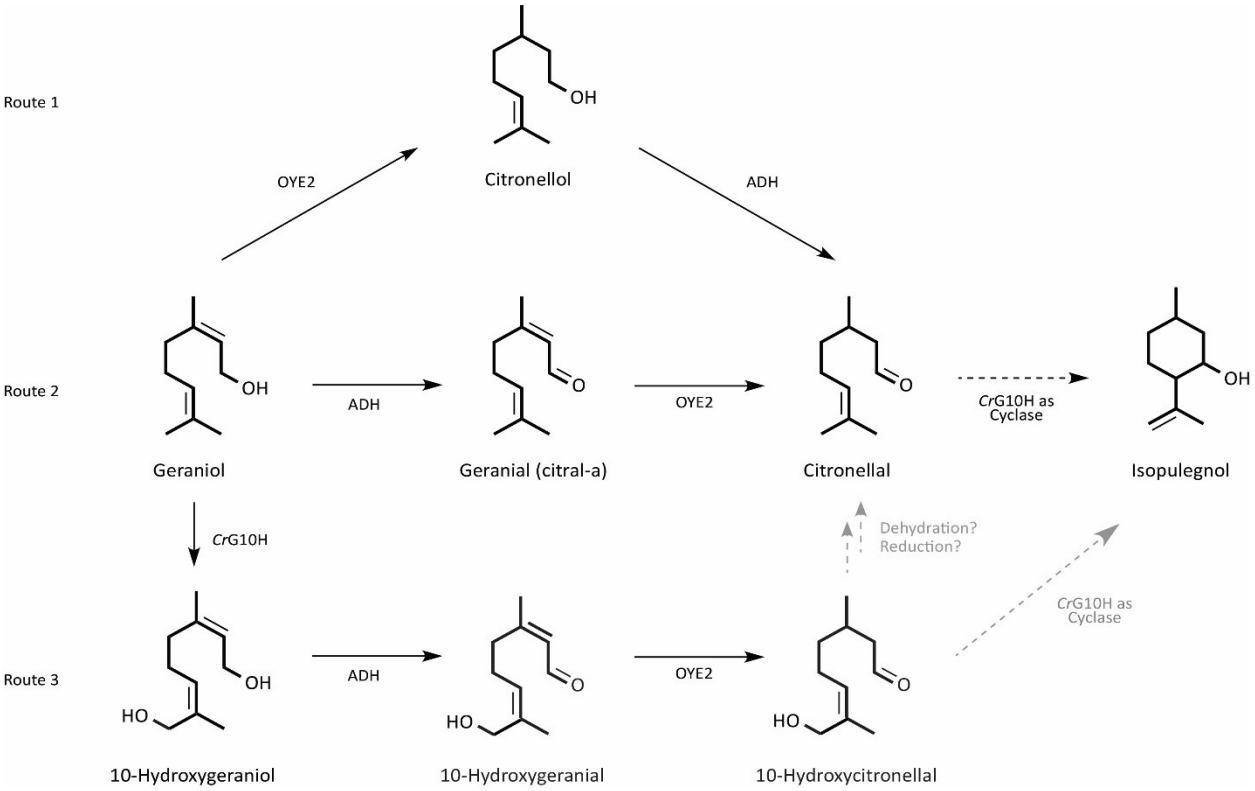

Figure 16. *In vivo* synthesis scheme of isopulegol in *S. cerevisiae*.
Based on the observed influence of the *oye2* deletion on isopulegol accumulation, three hypothetical *in vivo* biosynthetic routes can be proposed. These routes all proceed via differential combinations of the same native yeast enzymes, Oye2 and an Adh. Solid black lines indicate a prior observance of this exact or related activity. Hashed grey lines indicate unknown but possible steps and hashed black lines indicate the hypothesized step responsible for the final cyclisation to isopulegol.

Collectively, previous studies suggest a model for geraniol-based growth defects and rescue by *Δoye3*. Geraniol production likely causes loss of the diauxic shift by inhibiting yeast cells through cell membrane permeabilization, leading to K^+ efflux and membrane depolarization[158,163]. K^+ efflux drives ROS-induced programmed cell death (PCD) by the de-repression of PCD-specific hydrolytic enzymes (proteases and endonucleases), normally inhibited by high K^+ concentrations[199,200]. This decrease in osmolarity stimulates oxidative metabolism, leading to enhanced intracellular ROS production, which the cell integrity pathway addresses by redirecting protein kinase C to areas of compromised membrane integrity for ROS management[201]. Unfortunately for the cell, geraniol can inhibit protein kinase C[163], hampering the pathway involved in mitigating the ROS-induced PCD initiated by geraniol. Furthermore, aceto-acetyl CoA and HMG-CoA are subject to glucose repression so the majority of geraniol accumulation occurs post diauxic shift, during growth on ethanol[202]. This accumulation is in tandem with increased ROS formation from mitochondrial respiration, ultimately increasing the effect of osmotic stress and premature PCD, causing cell growth to flat-line post diauxic shift.

In response to ROS, the OYE homologs heterodimerize and influence H_2O_2-induced PCD[203,204]. Individually they have opposing roles whereby Oye2 alone is a potent antioxidant, mitigating ROS formation and suppressing PCD, however, Oye3 seems to be directly involved in H_2O_2-induced PCD, antagonizing the protective action of Oye2[203,205]. The activity of Oye3 is dependent on the presence of Oye2. Thus, deletion of *Oye3* prevents Oye2 sequestration and increases cell survival in the presence of H_2O_2[203]. Taking this into account, removal of *Oye3* may lessen the oxidative stress response that is responsible for initiating ROS mediated PCD, therefore increasing cell survival post diauxic shift, resulting in increased OD_{600} values during respiration. However, to validate this hypothesis, the *OYE* deletions should be performed individually and in tandem in the geraniol overexpressing MD5 to asses changes in growth phenotype.

Interestingly, this increase in cell growth post diauxic shift was mimicked by the CYP-CPR fusion strain which exhibited a similar growth phenotype to the *Δoye3* strain (Figure 9B and 14A). While the cause in unclear, it may be worth testing a strain with multiple CYP-CPR fusions to observe the effect on growth. It would also be interesting to observe the individual effects of a fourth and second copy of a *CrG10H* and a *CrCPR* respectively, as well as both

together (un-fused) into the 3x-*CrG10H* strain. This would aid in elucidating the reason for these altered growth phenotypes, as well as the CYP-CPR fusions' influence on 10HG synthesis.

Collectively, this study found multiple modifications that boost 10HG biosynthesis and made key observations that will inform future improvements for synthesis. Having identified optimal CPRs, appropriate CYP:CPR ratios can now be determined to enhance coupling for monooxygenase activity. Moreover, incorporation of known mutations that stabilize CYP-CPR interactions, improve enzyme turnover and promote cell growth should now be pursued. Finally, introduction of the CYP-CPR fusion and the *Δoye2*Δ*oye3* into the optimal strain backgrounds could further increase 10HG biosynthesis.

5. CONCLUSION

This study revealed multiple ways to improve the yield of heterologously produced 10HG in *S. cerevisiae*, ultimately for nepetalactol and strictosidine biosynthesis. The overexpression of the two key MVA genes paired with the Erg20^{K197E} mutant was successful in modulating the carbon flow for geraniol production. Though the G10H variants did not directly display novel C10 geraniol hydroxylation abilities, numerous CPRs beneficially influenced *Cr*G10H monooxygenase activity. Finally, the mitigation of off-target metabolite synthesis increased the amount of carbon allocated to 10HG biosynthesis. Together, the optimization strategies of this study add to the growing cohort of strictosidine pathway improvements and will accelerate efforts to replace the current inefficient and uneconomical production strategies of commercially valuable plant-derived monoterpenes.

REFERENCES

1. De Luca, V., Salim, V., Thamm, A., Masada, S. A. & Yu, F. Making iridoids/secoiridoids and monoterpenoid indole alkaloids: progress on pathway elucidation. *Curr. Opin. Plant Biol.* **19**, 35–42 (2014).
2. Wang, G., Tang, W. & Bidigare, R. R. Terpenoids As Therapeutic Drugs and Pharmaceutical Agents. in *Natural Products* (eds. Zhang, L. & Demain, A. L.) 197–227 (Humana Press, 2005). doi:10.1007/978-1-59259-976-9_9
3. Noble, R. ., Beer, C. T. & Cutts, J. H. Role of chance observations in chemotherapy: Vinca rosea. *Ann. N. Y. Acad. Sci.* **76**, 882–894 (1958).
4. Liscombe, D. K. & O'Connor, S. E. A virus-induced gene silencing approach to understanding alkaloid metabolism in Catharanthus roseus. *Phytochemistry* **72**, 1969–1977 (2011).
5. Zhou, X.-J. & Rahmani, R. Preclinical and Clinical Pharmacology of Vinca Alkaloids. *Drugs* **44**, 1–16 (1992).
6. van Der Heijden, R., Jacobs, D. I., Snoeijer, W., Hallard, D. & Verpoorte, R. The Catharanthus alkaloids: pharmacognosy and biotechnology. *Curr. Med. Chem.* **11**, 607–628 (2004).
7. Pyne, M. E. *et al.* Microbial Synthesis of Plant Alkaloids. in *Biotechnology of Natural Products* (eds. Schwab, W., Lange, B. & Wust, M.) 99–130 (Springer, 2018). doi:10.1007/978-3-319-67903-7_5
8. Zhu, J., Wang, M., Wen, W. & Yu, R. Biosynthesis and regulation of terpenoid indole alkaloids in Catharanthus roseus. *Pharmacogn. Rev.* **9**, 24–28 (2015).
9. Rischer, H. *et al.* Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in Catharanthus roseus cells. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5614–5619 (2006).
10. De Luca, V., Salim, V., Levac, D., Atsumi, S. M. & Yu, F. Discovery and Functional Analysis of Monoterpenoid Indole Alkaloid Pathways in Plants. in *Methods in Enzymology* **515**, 207–229 (Academic Press, 2012).
11. O'Connor, S. E. *et al.* Chemistry and biology of monoterpene indole alkaloid biosynthesis. *Nat. Prod. Rep.* **23**, 532–547 (2006).
12. Birkett, M. A., Hassanali, A., Hoglund, S., Pettersson, J. & Pickett, J. A. Repellent activity of catmint, Nepeta cataria, and iridoid nepetalactone isomers against Afro-tropical mosquitoes, ixodid ticks and red poultry mites. *Phytochemistry* **72**, 109–114 (2011).
13. Eisner, T. Catnip: Its Raison d'Être. *Science (80-.).* **146**, 1318–1320 (1964).
14. Schultz, G., Peterson, C. & Coats, J. Natural Insect Repellents: Activity against Mosquitoes and Cockroaches. *ACS Symp. Ser.* **927**, 168–181 (2006).
15. Bernier, U. R., Furman, K. D., Kline, D. L., Allan, S. A. & Barnard, D. R. Comparison of Contact and Spatial Repellency of Catnip Oil and N,N-Diethyl-3-methylbenzamide (Deet)

Against Mosquitoes. *J. Med. Entomol.* **42**, 306–311 (2005).

16. Stanczyk, N. M., Brookfield, J. F. Y., Ignell, R., Logan, J. G. & Field, L. M. Behavioral insensitivity to DEET in Aedes aegypti is a genetically determined trait residing in changes in sensillum function. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 8575–8580 (2010).

17. Murata, J., Roepke, J., Gordon, H. & De Luca, V. The Leaf Epidermome of Catharanthus roseus Reveals Its Biochemical Specialization. *Plant Cell* **20**, 524–542 (2008).

18. Yang, L. *et al.* Ultrasound-assisted extraction of the three terpenoid indole alkaloids vindoline, catharanthine and vinblastine from Catharanthus roseus using ionic liquid aqueous solutions. *Chem. Eng. J.* **172**, 705–712 (2011).

19. Choi, Y. H., Yoo, K.-P. & Kim, J. Supercritical Fluid Extraction and Liquid Chromatography-Electrospray Mass Analysis of Vinblastine from Catharanthus roseus. *Chem. Pharm. Bull. (Tokyo).* **50**, 1294–1296 (2002).

20. Mu, F. *et al.* Negative-Pressure Cavitation Extraction of Four Main Vinca Alkaloids from Catharanthus roseus Leaves. *Molecules* **17**, 8742–8752 (2012).

21. Glenn, W. S., Runguphan, W. & O'Connor, S. E. Recent progress in the metabolic engineering of alkaloids in plant systems. *Curr. Opin. Biotechnol.* **24**, 354–365 (2013).

22. Lipinski, C. & Hopkins, A. Navigating chemical space for biology and medicine. *Nature* **432**, 855–861 (2004).

23. Pollastro, F., Fontana, G. & Appendino, G. Natural Products Drug Discovery. in *Comprehensive Natural Products II Chemistry and Biology* (ed. Elsevier) 205–236 (2009).

24. Martin, D. B. C., Nguyen, L. Q. & Vanderwal, C. D. Syntheses of Strychnine, Norfluorocurarine, Dehydrodesacetylretuline, and Valparicine Enabled by Intramolecular Cycloadditions of Zincke Aldehydes. *J. Org. Chem.* **77**, 17–46 (2012).

25. Park, C.-H. *et al.* Catnip as a Source of Essential Oils. (2007).

26. Liblikas, I. *et al.* Simplified Isolation Procedure and Interconversion of the Diastereomers of Nepetalactone and Nepetalactol. *J. Nat. Prod.* **68**, 886–890 (2005).

27. Setzer, W. N. Catnip essential oil: There is more to it than making your cat go crazy. *Am. J. Essent. Oils Nat. Prod.* **4**, 12–15 (2016).

28. Zhu, J. *et al.* Adult Repellency and Larvicidal Activity of Five Plant Essential Oils Against Mosquitoes. *J. Am. Mosq. Control Assoc.* **22**, 515–522 (2006).

29. Chaturvedi, Y. Mosquito Repellent Industry, Insect Repellent, Market Analysis. (2016). Available at: https://www.alliedmarketresearch.com/mosquito-repellent-market. (Accessed: 11th June 2019)

30. Insect Repellent Market | Growth, Trends and Forecasts (2018-2023). (2018). Available at: https://www.mordorintelligence.com/industry-reports/insect-repellent-market. (Accessed: 11th June 2019)

31. North America Mosquito Repellent Market | Industry Report, 2018-2025. (2018). Available at: https://www.grandviewresearch.com/industry-analysis/north-america-mosquito-repellent-market. (Accessed: 11th June 2019)

32. Canada, S. *Agriculture Water Use in Canada*. (2013).

33. Pasquali, G., Porto, D. D. & Fett-Neto, A. G. Metabolic engineering of cell cultures versus whole plant complexity in production of bioactive monoterpene indole alkaloids: Recent progress related to old dilemma. *J. Biosci. Bioeng.* **101**, 287–296 (2006).

34. Heldt, H.-W. & Piechulla, B. Phenylpropanoids comprise a multitude of plant secondary metabolites and cell wall components. in *Plant Biochemistry* 431–449 (Elsevier, 2011). doi:10.1016/B978-0-12-384986-1.00018-1

35. Valluri, J. V. Bioreactor Production of Secondary Metabolites from Cell Cultures of Periwinkle and Sandalwood. in *Methods in Molecular Biology* 325–335 (Humana Press, Totowa, NJ, 2009). doi:10.1007/978-1-60327-287-2_26

36. Burlat, V., Oudin, A., Courtois, M., Rideau, M. & St-Pierre, B. Co-expression of three MEP pathway genes and *geraniol 10-hydroxylase* in internal phloem parenchyma of *Catharanthus roseus* implicates multicellular translocation of intermediates during the biosynthesis of monoterpene indole alkaloids and iso. *Plant J.* **38**, 131–141 (2004).

37. Ramani, S. & Jayabaskaran, C. Enhanced catharanthine and vindoline production in suspension cultures of Catharanthus roseus by ultraviolet-B light. *J. Mol. Signal.* **3**, 9 (2008).

38. Liu, D.-H. *et al.* Enhanced accumulation of catharanthine and vindoline in Catharanthus roseus hairy roots by overexpression of transcriptional factor ORCA2. *African J. Biotechnol.* **10**, 3260–3268 (2011).

39. Pomahačová, B. *et al.* Improved accumulation of ajmalicine and tetrahydroalstonine in Catharanthus cells expressing an ABC transporter. *J. Plant Physiol.* **166**, 1405–1412 (2009).

40. Pan, Q. *et al.* Overexpression of ORCA3 and G10H in Catharanthus roseus Plants Regulated Alkaloid Biosynthesis and Metabolism Revealed by NMR-Metabolomics. *PLoS One* **7**, e43038 (2012).

41. Sharma, A., Verma, P., Mathur, A. & Mathur, A. K. Genetic engineering approach using early Vinca alkaloid biosynthesis genes led to increased tryptamine and terpenoid indole alkaloids biosynthesis in differentiating cultures of Catharanthus roseus. *Protoplasma* **255**, 425–435 (2018).

42. Gorman, M., Neuss, N. & Biemann, K. Vinca Alkaloids. X1. The Structure of Vindoline. *J. Am. Chem. Soc.* **84**, 1058–1059 (1962).

43. Svoboda, G. H., Neuss, N. & Gorman, M. Alkaloids of Vinca rosea Linn. (Catharanthus roseus G. Don.) V.: Preparation and Characterization of Alkaloids. *J. Am. Pharm. Assoc. (Scientific ed.)* **48**, 659–666 (1959).

44. Hayato Ishikawa, David A. Colby, A. & Boger, D. L. Direct Coupling of Catharanthine

and Vindoline to Provide Vinblastine: Total Synthesis of (+)- and ent-(−)-Vinblastine. *J. Am. Chem. Soc.* **130**, 420–421 (2008).

45. Ishikawa, H. *et al.* Total Synthesis of Vinblastine, Vincristine, Related Natural Products, and Key Structural Analogues. *J. Am. Chem. Soc.* **131**, 4904–4916 (2009).

46. Yokoshima, S., Tokuyama, H. & Fukuyama, T. Total synthesis of (+)-vinblastine: Control of the stereochemistry at C18′. *Chem. Rec.* **10**, 101–118 (2010).

47. Kobayashi, S., Ueda, T. & Fukuyama, T. An efficient total synthesis of (-)-Vindoline. *Synlett* **6**, 883–886 (2000).

48. Satoshi Yokoshima *et al.* Stereocontrolled Total Synthesis of (+)-Vinblastine. *J. Am. Chem. Socitey* **124**, 2137–2139 (2002).

49. Vucurovic, V., Razmovski, R. & Rebic, M. A corn stem as biomaterial for Saccharomyces cerevisiae cells immobilization for the ethanol production. *Chem. Ind. Chem. Eng. Q.* **14**, 235–238 (2008).

50. Demirbas, M. F., Balat, M. & Balat, H. Potential contribution of biomass to the sustainable energy development. *Energy Convers. Manag.* **50**, 1746–1760 (2009).

51. Ragauskas, A. J. *et al.* The Path Forward for Biofuels and Biomaterials. *Science (80-.).* **311**, 484–489 (2006).

52. Westfall, P. J. *et al.* Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E111-8 (2012).

53. Asadollahi, M. A. *et al.* Production of plant sesquiterpenes in Saccharomyces cerevisiae: Effect of ERG9 repression on sesquiterpene biosynthesis. *Biotechnol. Bioeng.* **99**, 666–677 (2008).

54. DeJong, J. M. *et al.* Genetic engineering of taxol biosynthetic genes in Saccharomyces cerevisiae. *Biotechnol. Bioeng.* **93**, 212–224 (2006).

55. Engels, B., Dahm, P. & Jennewein, S. Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. *Metab. Eng.* **10**, 201–206 (2008).

56. Kirby, J. & Keasling, J. D. Biosynthesis of Plant Isoprenoids: Perspectives for Microbial Engineering. *Annu. Rev. Plant Biol.* **60**, 335–355 (2009).

57. Brown, S., Clastre, M., Courdavault, V. & O'Connor, S. E. De novo production of the plant-derived alkaloid strictosidine in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3205–3210 (2015).

58. Campbell, A. *et al.* Engineering of a Nepetalactol-Producing Platform Strain of Saccharomyces cerevisiae for the Production of Plant Seco-Iridoids. *ACS Synth. Biol.* **5**, 405–414 (2016).

59. Fischer, M. J. C., Meyer, S., Claudel, P., Bergdoll, M. & Karst, F. Metabolic engineering of monoterpene synthesis in yeast. *Biotechnol. Bioeng.* **108**, 1883–1892 (2011).

60. Ignea, C. *et al.* Improving yeast strains using recyclable integration cassettes, for the production of plant terpenoids. *Microb. Cell Fact.* **10**, 4 (2011).

61. Ro, D.-K. *et al.* Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **440**, 940–943 (2006).

62. Rohmer, M., Knani, M., Simonin, P., Sutter, B. & Sahm, H. Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate. *Biochem. J.* **295**, 517–524 (1993).

63. Eisenreich, W., Sagner, S., Zenk, M. H. & Bacher, A. Monoterpenoid essential oils are not of mevalonoid origin. *Tetrahedron Lett.* **38**, 3889–3892 (1997).

64. Rohmer, M. The discovery of a mevalonate-independent pathway for isoprenoid biosynthesis in bacteria, algae and higher plants. *Nat. Prod. Rep.* **16**, 565–574 (1999).

65. Lange, B. M., Rujan, T., Martin, W. & Croteau, R. Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways across genomes. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 13172–13177 (2000).

66. Chang, W., Song, H., Liu, H. & Liu, P. Current development in isoprenoid precursor biosynthesis and regulation. *Curr. Opin. Chem. Biol.* **17**, 571–579 (2013).

67. Brown, M. S., Dana, S. E. & Goldstein, J. L. Regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase activity in human fibroblasts by lipoproteins. *Proc. Natl. Acad. Sci. U. S. A.* **70**, 2162–2166 (1973).

68. Siperstein, M. D. & Guest, M. J. Studies on the site of the feedback control of cholesterol synthesis. *J. Clin. Invest.* **39**, 642–652 (1960).

69. Brown, M. S. & Goldstein, J. L. Multivalent feedback regulation of HMG CoA reductase, a control mechanism coordinating isoprenoid synthesis and cell growth. *J. Lipid Res.* **21**, 505–517 (1980).

70. Burg, J. S. & Espenshade, P. J. Regulation of HMG-CoA reductase in mammals and yeast. *Prog. Lipid Res.* **50**, 403–410 (2011).

71. Kizer, L., Pitera, D. J., Pfleger, B. F. & Keasling, J. D. Application of functional genomics to pathway optimization for increased isoprenoid production. *Appl. Environ. Microbiol.* **74**, 3229–3241 (2008).

72. Basson, M. E., Thorsness, M. & Rine, J. Saccharomyces cerevisiae contains two functional genes encoding 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Proc. Natl. Acad. Sci. U. S. A.* **83**, 5563–5567 (1986).

73. Hampton, R., Dimster-Denk, D. & Rine, J. The biology of HMG-CoA reductase: the pros of contra-regulation. *Trends Biochem. Sci.* **21**, 140–145 (1996).

74. Dorsey, J. K. & Porter, J. W. The inhibition of mevalonic kinase by geranyl and farnesyl pyrophosphates. *J. Biol. Chem.* **243**, 4667–4670 (1968).

75. Primak, Y. A. *et al.* Characterization of a feedback-resistant mevalonate kinase from the archaeon Methanosarcina mazei. *Appl. Environ. Microbiol.* **77**, 7772–7778 (2011).

76. Wu, Z., Wouters, J. & Poulter, C. D. Isopentenyl diphosphate isomerase. Mechanism-based inhibition by diene analogues of isopentenyl diphosphate and dimethylallyl diphosphate. *J. Am. Chem. Soc.* **127**, 17433–17438 (2005).

77. Liu, J., Zhang, W., Du, G., Chen, J. & Zhou, J. Overproduction of geraniol by enhanced precursor supply in Saccharomyces cerevisiae. *J. Biotechnol.* **168**, 446–451 (2013).

78. Rohdich, F. *et al.* Studies on the nonmevalonate terpene biosynthetic pathway: Metabolic role of IspH (LytB) protein. *Proc. Natl. Acad. Sci.* **99**, 1158–1163 (2002).

79. Oswald, M., Fischer, M., Dirninger, N. & Karst, F. Monoterpenoid biosynthesis in Saccharomyces cerevisiae. *FEMS Yeast Res.* **7**, 413–421 (2007).

80. Herrero, Ó., Ramón, D. & Orejas, M. Engineering the Saccharomyces cerevisiae isoprenoid pathway for de novo production of aromatic monoterpenes in wine. *Metab. Eng.* **10**, 78–86 (2008).

81. Yang, L. *et al.* A homomeric geranyl diphosphate synthase-encoding gene from Camptotheca acuminata and its combinatorial optimization for production of geraniol in Escherichia coli. *J. Ind. Microbiol. Biotechnol.* **44**, 1431–1441 (2017).

82. Burke, C. & Croteau, R. Geranyl diphosphate synthase from Abies grandis: cDNA isolation, functional expression, and characterization. *Arch. Biochem. Biophys.* **405**, 130–136 (2002).

83. Mahmoud, S. S. & Croteau, R. B. Strategies for transgenic manipulation of monoterpene biosynthesis in plants. *Trends Plant Sci.* **7**, 366–373 (2002).

84. Kleinig, H. The Role of Plastids in Isoprenoid Biosynthesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 39–59 (1989).

85. Iijima, Y., Gang, D. R., Fridman, E., Lewinsohn, E. & Pichersky, E. Characterization of geraniol synthase from the peltate glands of sweet basil. *Plant Physiol.* **134**, 370–9 (2004).

86. Collu, G. *et al.* Geraniol 10-hydroxylase 1 , a cytochrome P450 enzyme involved in terpenoid indole alkaloid biosynthesis. *FEBS Lett.* **508**, 215–220 (2001).

87. Krithika, R. *et al.* Characterization of 10-hydroxygeraniol dehydrogenase from Catharanthus roseus reveals cascaded enzymatic activity in iridoid biosynthesis. *Sci. Rep.* **5**, 8258 (2015).

88. Geu-Flores, F. *et al.* An alternative route to cyclic terpenes by reductive cyclization in iridoid biosynthesis. *Nature* **492**, 138–142 (2012).

89. Lichman, B. R. *et al.* Uncoupled activation and cyclization in catmint reductive terpenoid biosynthesis. *Nat. Chem. Biol.* **15**, 71–79 (2019).

90. Courdavault, V. *et al.* A look inside an alkaloid multisite plant: the Catharanthus logistics. *Curr. Opin. Plant Biol.* **19**, 43–50 (2014).

91. Carrau, F. M. *et al.* De novo synthesis of monoterpenes by *Saccharomyces cerevisiae* wine yeasts. *FEMS Microbiol. Lett.* **243**, 107–115 (2005).

92. Blanchard, L. & Karst, F. Characterization of a lysine-to-glutamic acid mutation in a conservative sequence of farnesyl diphosphate synthase from Saccharomyces cerevisiae. *Gene* **125**, 185–189 (1993).

93. Chambon, C. *et al.* Sterol pathway in yeast. Identification and properties of mutant strains defective in mevalonate diphosphate decarboxylase and farnesyl diphosphate synthetase. *Lipids* **26**, 633–636 (1991).

94. Chambon, C., Ladeveze, V., Oulmouden, A., Servouse, M. & Karst, F. Isolation and properties of yeast mutants affected in farnesyl diphosphate synthetase. *Curr. Genet.* **18**, 41–46 (1990).

95. Ignea, C., Pontini, M., Maffei, M. E., Makris, A. M. & Kampranis, S. C. Engineering Monoterpene Production in Yeast Using a Synthetic Dominant Negative Geranyl Diphosphate Synthase. *ACS Synth. Biol.* **3**, 298–306 (2014).

96. Geelen, M. J. H., Gibson, D. M. & Rodwell, V. W. *Hydroxymethylglutaryl-CoA reductase-the rate-limiting enzyme of cholesterol biosynthesis*. **201**, (1986).

97. Martin, V. J. J., Pitera, D. J., Withers, S. T., Newman, J. D. & Keasling, J. D. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat. Biotechnol.* **21**, 796–802 (2003).

98. Polakowski, T., Stahl, U. & Lang, C. Overexpression of a cytosolic hydroxymethylglutaryl-CoA reductase leads to squalene accumulation in yeast. *Appl. Microbiol. Biotechnol.* **49**, 66–71 (1998).

99. Gil, G., Faust, J. R., Chin, D. J., Goldstein, J. L. & Brown, M. S. Membrane-bound domain of HMG CoA reductase is required for sterol-enhanced degradation of the enzyme. *Cell* **41**, 249–258 (1985).

100. Steyer, D. *et al.* Genetic analysis of geraniol metabolism during fermentation. *Food Microbiol.* **33**, 228–234 (2013).

101. 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. *Microb. Cell Fact.* **16**, 17 (2017).

102. Billingsley, J. M. *et al.* Engineering the biocatalytic selectivity of iridoid production in Saccharomyces cerevisiae. *Metab. Eng.* **44**, 117–125 (2017).

103. Trotter, E. W., Collinson, E. J., Dawes, I. W. & Grant, C. M. Old yellow enzymes protect against acrolein toxicity in the yeast Saccharomyces cerevisiae. *Appl. Environ. Microbiol.* **72**, 4885–4892 (2006).

104. Zheng, L., Lin, J., Zhang, B., Kuang, Y. & Wei, D. Identification of a yeast old yellow enzyme for highly enantioselective reduction of citral isomers to (R)-citronellal. *Bioresour. Bioprocess.* **5**, 9 (2018).

105. Bruce, N. C. & Williams, R. E. 'New uses for an Old Enzyme' – the Old Yellow Enzyme family of flavoenzymes. *Microbiology* **148**, 1607–1614 (2002).

106. Zhao, J., Bao, X., Li, C., Shen, Y. & Hou, J. Improving monoterpene geraniol production through geranyl diphosphate synthesis regulation in Saccharomyces cerevisiae. *Appl. Microbiol. Biotechnol.* **100**, 4561–4571 (2016).

107. Qin, L. *et al.* Structure of iridoid synthase in complex with NADP + /8-oxogeranial reveals the structural basis of its substrate specificity. *J. Struct. Biol.* **194**, 224–230 (2016).

108. Young, E. T. & Pilgrim, D. Isolation and DNA sequence of ADH3, a nuclear gene encoding the mitochondrial isozyme of alcohol dehydrogenase in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **5**, 3024–3034 (1985).

109. Jiang, G.-Z. *et al.* Manipulation of GES and ERG20 for geraniol overproduction in Saccharomyces cerevisiae. *Metab. Eng.* **41**, 57–66 (2017).

110. Guirimand, G. *et al.* Optimization of the transient transformation of Catharanthus roseus cells by particle bombardment and its application to the subcellular localization of hydroxymethylbutenyl 4-diphosphate synthase and geraniol 10-hydroxylase. *Plant Cell Rep.* **28**, 1215–1234 (2009).

111. De Lemos-Chiarandini, C., Frey, A. B., Sabatini, D. D. & Kreibich, G. Determination of the membrane topology of the phenobarbital-inducible rat liver cytochrome P-450 isoenzyme PB-4 using site-specific antibodies. *J. Cell Biol.* **104**, 209–219 (1987).

112. Szczesna-Skorupa, E. & Kemper, B. An N-terminal glycosylation signal on cytochrome P450 is restricted to the endoplasmic reticulum in a luminal orientation. *J. Biol. Chem.* **268**, 1757–1762 (1993).

113. Szczesna-Skorupa, E., Ahn, K., Chen, C. D., Doray, B. & Kemper, B. The cytoplasmic and N-terminal transmembrane domains of cytochrome P450 contain independent signals for retention in the endoplasmic reticulum. *J. Biol. Chem.* **270**, 24327–24333 (1995).

114. Ro, D.-K., Ehlting, J. & Douglas, C. J. Cloning, Functional Expression, and Subcellular Localization of Multiple NADPH-Cytochrome P450 Reductases from Hybrid Poplar. *Plant Physiol.* **130**, 1837–1851 (2002).

115. Shephard, E. A., Phillips, I. R., Bayney, R. M., Pike, S. F. & Rabin, B. R. Quantification of NADPH: cytochrome P-450 reductase in liver microsomes by a specific radioimmunoassay technique. *Biochem. J.* **211**, 333–340 (1983).

116. Walters Biggs, B. *et al.* Overcoming heterologous protein interdependency to optimize P450-mediated Taxol precursor synthesis in Escherichia coli. *Proc. Natl. Acad. Sci.* **113**, 3209–3214 (2016).

117. Jensen, K. & Møller, B. L. Plant NADPH-cytochrome P450 oxidoreductases. *Phytochemistry* **71**, 132–141 (2010).

118. Manoj, K. M., Gade, S. K. & Mathew, L. Cytochrome P450 Reductase: A Harbinger of Diffusible Reduced Oxygen Species. *PLoS One* **5**, e13272 (2010).

119. Gold, N. D. *et al.* A Combinatorial Approach To Study Cytochrome P450 Enzymes for *De Novo* Production of Steviol Glucosides in Baker's Yeast. *ACS Synth. Biol.* **7**, 2918–2929 (2018).

120. Barnes, H. J., Arlotto, M. P. & Waterman, M. R. Expression and enzymatic activity of recombinant cytochrome P450 17 alpha-hydroxylase in Escherichia coli. *Proc. Natl. Acad. Sci.* **88**, 5597–5601 (1991).

121. Jenkins, C. M. & Waterman, M. R. Flavodoxin and NADPH-flavodoxin reductase from Escherichia coli support bovine cytochrome P450c17 hydroxylase activities. *J. Biol. Chem.* **269**, 27401–27408 (1994).

122. Urban, P., Mignotte, C., Kazmaier, M., Delorme, F. & Pompon, D. Cloning, Yeast Expression, and Characterization of the Coupling of Two Distantly Related *Arabidopsis thaliana* NADPH-Cytochrome P450 Reductases with P450 CYP73A5. *J. Biol. Chem.* **272**, 19176–19186 (1997).

123. Eberle, D., Ullmann, P., Werck-Reichhart, D. & Petersen, M. cDNA cloning and functional characterisation of CYP98A14 and NADPH:cytochrome P450 reductase from Coleus blumei involved in rosmarinic acid biosynthesis. *Plant Mol. Biol.* **69**, 239–253 (2009).

124. Larbat, R. *et al.* Molecular Cloning and Functional Characterization of Psoralen Synthase, the First Committed Monooxygenase of Furanocoumarin Biosynthesis. *J. Biol. Chem.* **282**, 542–554 (2007).

125. Reed, J. R. & Backes, W. L. Physical Studies of P450–P450 Interactions: Predicting Quaternary Structures of P450 Complexes in Membranes from Their X-ray Crystal Structures. *Front. Pharmacol.* **8**, 28 (2017).

126. Galanie, S., Thodey, K., Trenchard, I. J., Filsinger Interrante, M. & Smolke, C. D. Complete biosynthesis of opioids in yeast. *Science (80-.).* **349**, 1095–1100 (2015).

127. Nedelkina, S. *et al.* Novel characteristics and regulation of a divergent cinnamate 4-hydroxylase (CYP73A15) from French bean: engineering expression in yeast. *Plant Mol. Biol.* **39**, 1079–1090 (1999).

128. McLean, K. J., Girvan, H. M. & Munro, A. W. Cytochrome P450/redox partner fusion enzymes: biotechnological and toxicological prospects. *Expert Opin. Drug Metab. Toxicol.* **3**, 847–863 (2007).

129. Noble, M. A. *et al.* Roles of key active-site residues in flavocytochrome P450 BM3. *Biochem. J.* **339**, 371–379 (1999).

130. Leonard, E. & Koffas, M. A. G. Engineering of Artificial Plant Cytochrome P450 Enzymes for Synthesis of Isoflavones by Escherichia coli. *Appl. Environ. Microbiol.* **73**, 7246–7251 (2007).

131. Schläpfer, P. *et al.* Genome-Wide Prediction of Metabolic Enzymes, Pathways, and Gene Clusters in Plants. *Plant Physiol.* **173**, 2041–2059 (2017).

132. Matasci, N. *et al.* Data access for the 1,000 Plants (1KP) project. *Gigascience* **3**, 17 (2014).

133. Boratyn, G. M. *et al.* Domain enhanced lookup time accelerated BLAST. *Biol. Direct* **7**, 12 (2012).

134. Huang, Y., Niu, B., Gao, Y., Fu, L. & Li, W. CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**, 680–682 (2010).

135. Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).

136. Emanuelsson, O., Nielsen, H. & Heijne, G. Von. ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**, 978–984 (1999).

137. Sonnhammer, E. L., von Heijne, G. & Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proceedings. Int. Conf. Intell. Syst. Mol. Biol.* **6**, 175–82 (1998).

138. Jinek, M. *et al.* A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (80-.).* **337**, 816–821 (2012).

139. Oldenburg, K., Vo, K. T., Michaelis, S. & Paddon, C. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic Acids Res.* **25**, 451–452 (1997).

140. Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. J. Yeast transformation: a model system for the study of recombination. *Proc. Natl. Acad. Sci.* **78**, 6354–6358 (1981).

141. Mikkelsen, M. D. *et al.* Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab. Eng.* **14**, 104–11 (2012).

142. Flagfeldt, D. B., Siewers, V., Huang, L. & Nielsen, J. Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. *Yeast* **26**, 545–51 (2009).

143. Stemmer, M., Thumberger, T., del Sol Keyer, M., Wittbrodt, J. & Mateo, J. L. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One* **10**, e0124633 (2015).

144. Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2**, 31–4 (2007).

145. DiCarlo, J. E. *et al.* Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Res.* **41**, 4336–43 (2013).

146. Helfman, D. M., Fiddes, J. C. & Hanahan, D. Directional cDNA cloning in plasmid vectors by sequential addition of oligonucleotide linkers. in *Methods in Enzymology* (eds. Berger, S. L. & Kimmel, A. R.) **152**, 349–359 (Academic Press, 1987).

147. NIST Standard Reference Database Number 173. in *NIST Standard Reference Simulation Website* (ed. Shen, V.K., Siderius, D.W., Krekelberg, W.P., and Hatch, H. W.) 20899 (National Institute of Standards and Technology). doi:http://doi.org/10.18434/T4M88Q

148. Werck-Reichhart, D. & Feyereisen, R. Cytochromes P450: a success story. *Genome Biol.* **1**, reviews3003.1 (2000).

149. Shitiz, K., Sharma, N., Pal, T., Sood, H. & Chauhan, R. S. NGS Transcriptomes and

Enzyme Inhibitors Unravel Complexity of Picrosides Biosynthesis in Picrorhiza kurroa Royle ex. Benth. *PLoS One* **10**, e0144546 (2015).

150. Zhang, X., Allan, A. C., Li, C., Wang, Y. & Yao, Q. De Novo Assembly and Characterization of the Transcriptome of the Chinese Medicinal Herb, Gentiana rigescens. *Int. J. Mol. Sci.* **16**, 11550–11573 (2015).

151. Bourgeois, L., Pyne, M. E. & Martin, V. J. J. A Highly Characterized Synthetic Landing Pad System for Precise Multicopy Gene Integration in Yeast. *ACS Synth. Biol.* **7**, 2675–2685 (2018).

152. Jacob, R. G., Perin, G., Loi, L. N., Pinno, C. S. & Lenardão, E. J. Green synthesis of (−)-isopulegol from (+)-citronellal: application to essential oil of citronella. *Tetrahedron Lett.* **44**, 3605–3608 (2003).

153. Chuah, G. ., Liu, S. ., Jaenicke, S. & Harrison, L. . Cyclisation of Citronellal to Isopulegol Catalysed by Hydrous Zirconia and Other Solid Acids. *J. Catal.* **200**, 352–359 (2001).

154. Jensen, B. L., Malkawi, A. & McGowan, V. Cyclization of the Monoterpene Citronellal to Isopulegol: A Biomimetic Natural Product Synthesis. *J. Chem. Educ.* **77**, 1474 (2000).

155. Croteau, R. B., Davis, E. M., Ringer, K. L. & Wildung, M. R. (−)-Menthol biosynthesis and molecular genetics. *Naturwissenschaften* **92**, 562–577 (2005).

156. Siedenburg, G. *et al.* Activation-independent cyclization of monoterpenoids. *Appl. Environ. Microbiol.* **78**, 1055–62 (2012).

157. Szkopińska, A. *et al.* Polyprenol formation in the yeast Saccharomyces cerevisiae: effect of farnesyl diphosphate synthase overexpression. *J. Lipid Res.* **38**, 962–968 (1997).

158. Bard, M., Albrecht, M. R., Gupta, N., Guynn, C. J. & Stillwell, W. Geraniol interferes with membrane functions in strains of Candida and Saccharomyces. *Lipids* **23**, 534–538 (1988).

159. Gilpin, S. J., Hui, X. & Maibach, H. I. Volatility of fragrance chemicals: patch testing implications. *Dermat. contact, atopic, Occup. drug* **20**, 200–207

160. Zhang, Y., Nielsen, J. & Liu, Z. Engineering yeast metabolism for production of terpenoids for use as perfume ingredients, pharmaceuticals and biofuels. *FEMS Yeast Res.* **17**, (2017).

161. Jongedijk, E. *et al.* Capturing of the monoterpene olefin limonene produced in *Saccharomyces cerevisiae*. *Yeast* **32**, n/a-n/a (2014).

162. Tippmann, S., Scalcinati, G., Siewers, V. & Nielsen, J. Production of farnesene and santalene by *Saccharomyces cerevisiae* using fed-batch cultivations with *RQ* -controlled feed. *Biotechnol. Bioeng.* **113**, 72–81 (2016).

163. Carnesecchi, S. *et al.* Perturbation by Geraniol of Cell Membrane Permeability and Signal Transduction Pathways in Human Colon Cancer Cells. *J. Pharmacol. Exp. Ther.* **303**, 711–715 (2002).

164. Guengerich, F. P. & Munro, A. W. Unusual Cytochrome P450 Enzymes and Reactions. *J.*

Biol. Chem. **288**, 17065–17073 (2013).

165. Liu, Q. *et al.* Kauniolide synthase is a P450 with unusual hydroxylation and cyclization-elimination activity. *Nat. Commun.* **9**, 4657 (2018).

166. Yang, Y. *et al.* Bifunctional Cytochrome P450 Enzymes Involved in Camptothecin Biosynthesis. *ACS Chem. Biol.* acschembio.8b01124 (2019). doi:10.1021/acschembio.8b01124

167. Zhao, B. *et al.* Crystal structure of albaflavenone monooxygenase containing a moonlighting terpene synthase active site. *J. Biol. Chem.* **284**, 36711–36719 (2009).

168. Sigrist, C. J. A. *et al.* New and continuing developments at PROSITE. *Nucleic Acids Res.* **41**, D344–D347 (2012).

169. Jia, Q. Computational Identification of Terpene Synthase Genes and Their Evolutionary Analysis. (University of Tennessee, 2016).

170. Tijet, N. & Brash, A. R. Allene oxide synthases and allene oxides. *Prostaglandins Other Lipid Mediat.* **68–69**, 423–431 (2002).

171. Johnson, E. F. & Stout, C. D. Structural diversity of eukaryotic membrane cytochrome p450s. *J. Biol. Chem.* **288**, 17082–90 (2013).

172. Shaw, P. M., Hosea, N. A., Thompson, D. V., Lenius, J. M. & Guengerich, F. P. Reconstitution Premixes for Assays Using Purified Recombinant Human Cytochrome P450, NADPH-Cytochrome P450 Reductase, and Cytochromeb5. *Arch. Biochem. Biophys.* **348**, 107–115 (1997).

173. Denisov, I. G. & Sligar, S. G. Cytochromes P450 in Nanodiscs. *Biochim. Biophys. Acta - Proteins Proteomics* **1814**, 223–229 (2011).

174. Denisov, I. G. & Sligar, S. G. Nanodiscs in Membrane Biochemistry and Biophysics. *Chem. Rev.* **117**, 4669–4713 (2017).

175. Barnaba, C., Martinez, M. J., Taylor, E., Barden, A. O. & Brozik, J. A. Single-Protein Tracking Reveals That NADPH Mediates the Insertion of Cytochrome P450 Reductase into a Biomimetic of the Endoplasmic Reticulum. *J. Am. Chem. Soc.* **139**, 5420–5430 (2017).

176. Barnaba, C., Gentry, K., Sumangala, N. & Ramamoorthy, A. The catalytic function of cytochrome P450 is entwined with its membrane-bound nature. *F1000Research* **6**, 662 (2017).

177. Murtazina, D. A. *et al.* Phospholipids modify substrate binding and enzyme activity of human cytochrome P450 27A1. *J. Lipid Res.* **45**, 2345–2353 (2004).

178. Klug, L. & Daum, G. Yeast lipid metabolism at a glance. *FEMS Yeast Res.* **14**, 369–388 (2014).

179. Fouillen, L., Maneta-Peyret, L. & Moreau, P. ER Membrane Lipid Composition and Metabolism: Lipidomic Analysis. in 125–137 (2018). doi:10.1007/978-1-4939-7389-7_10

180. Kim, H.-J., Lee, S.-B., Guengerich, F. P., Park, Y. I. & Dong, M.-S. Effects of N-terminal modification of recombinant human cytochrome P450 1A2 on catalytic activity. *Xenobiotica* **37**, 356–365 (2007).

181. Schilmiller, A. L. *et al.* Monoterpenes in the glandular trichomes of tomato are synthesized from a neryl diphosphate precursor rather than geranyl diphosphate. *Proc. Natl. Acad. Sci.* **106**, 10865–10870 (2009).

182. Hanna, I. H., Kim, M.-S. & Guengerich, F. P. Heterologous Expression of Cytochrome P450 2D6 Mutants, Electron Transfer, and Catalysis of Bufuralol Hydroxylation: The Role of Aspartate 301 in Structural Integrity. *Arch. Biochem. Biophys.* **393**, 255–261 (2001).

183. Dong, M.-S. *et al.* Identification of Retained *N* -Formylmethionine in Bacterial Recombinant Mammalian Cytochrome P450 Proteins with the N-Terminal Sequence MALLLAVFL...: Roles of Residues 3−5 in Retention and Membrane Topology. *Biochemistry* **35**, 10031–10040 (1996).

184. Ilia G. Denisov, Thomas M. Makris, Stephen G. Sligar, and & Ilme Schlichting. Structure and Chemistry of Cytochrome P450. *Chermical Rev.* **105**, 2253–2278 (2005).

185. Emmerstorfer, A. *et al.* Over-expression of *ICE2* stabilizes cytochrome P450 reductase in *Saccharomyces cerevisiae* and *Pichia pastoris*. *Biotechnol. J.* **10**, 623–635 (2015).

186. Estrada de Martin, P., Du, Y., Novick, P. & Ferro-Novick, S. Ice2p is important for the distribution and structure of the cortical ER network in Saccharomyces cerevisiae. *J. Cell Sci.* **118**, 65–77 (2005).

187. Yamada, R. *et al.* Cocktail δ-integration: a novel method to construct cellulolytic enzyme expression ratio-optimized yeast strains. *Microb. Cell Fact.* **9**, 32 (2010).

188. Lee, F. W. & Da Silva, N. A. Improved efficiency and stability of multiple cloned gene insertions at the delta sequences of Saccharomyces cerevisiae. *Appl. Microbiol. Biotechnol.* **48**, 339–45 (1997).

189. Shi, S., Liang, Y., Zhang, M. M., Ang, E. L. & Zhao, H. A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in Saccharomyces cerevisiae. *Metab. Eng.* **33**, 19–27 (2016).

190. Gonda, D. K. *et al.* Universality and structure of the N-end rule. *J. Biol. Chem.* **264**, 16700–16712 (1989).

191. Mogk, A., Schmidt, R. & Bukau, B. The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends Cell Biol.* **17**, 165–172 (2007).

192. Varshavsky, A. The N-end rule pathway and regulation by proteolysis. *Protein Sci.* **20**, 1298–1345 (2011).

193. Varshavsky, A. The N-end rule pathway of protein degradation. *Genes to Cells* **2**, 13–28 (1997).

194. Osterberg, M. *et al.* Phenotypic effects of membrane protein overexpression in

Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci.* **103**, 11148–11153 (2006).

195. Hart-Cooper, W. M., Clary, K. N., Toste, F. D., Bergman, R. G. & Raymond, K. N. Selective Monoterpene-like Cyclization Reactions Achieved by Water Exclusion from Reactive Intermediates in a Supramolecular Catalyst. *J. Am. Chem. Soc.* **134**, 17873–17876 (2012).

196. da Silva, K. A. *et al.* Cyclization of (+)-citronellal to (−)-isopulegol catalyzed by H3PW12O40/SiO2. *Catal. Commun.* **5**, 425–429 (2004).

197. Peters, C. & Buller, R. Linear enzyme cascade for the production of (-)-iso-isopulegol. *Z. Naturforsch* **74**, 63–70 (2019).

198. Itoh, T. *et al.* Reconstitution of a fungal meroterpenoid biosynthesis reveals the involvement of a novel family of terpene cyclases. *Nat. Chem.* **2**, 858–864 (2010).

199. Demidchik, V. *et al.* Stress-induced electrolyte leakage: the role of K+-permeable channels and involvement in programmed cell death and metabolic adjustment. *J. Exp. Bot.* **65**, 1259–1270 (2014).

200. Hoeberichts, F. A. *et al.* The role of K+ and H+ transport systems during glucose- and H2O2-induced cell death in Saccharomyces cerevisiae. *Yeast* **27**, 713–725 (2010).

201. Hohmann, S. Osmotic stress signaling and osmoadaptation in yeasts. *Microbiol. Mol. Biol. Rev.* **66**, 300–72 (2002).

202. Seker, T., Moller, K. & Nielsen, J. Analysis of acyl CoA ester intermediates of the mevalonate pathway in Saccharomyces cerevisiae. *Appl. Microbiol. Biotechnol.* **67**, 119–124 (2005).

203. Odat, O. *et al.* Old Yellow Enzymes, Highly Homologous FMN Oxidoreductases with Modulating Roles in Oxidative Stress and Programmed Cell Death in Yeast. *J. Biol. Chem.* **282**, 36010–36023 (2007).

204. Stott, K., Saito, K., Thiele, D. J. & Massey, V. Old Yellow Enzyme. The discovery of multiple isozymes and a family of related proteins. *J. Biol. Chem.* **268**, 6097–6106 (1993).

205. Reekmans, R., De Smet, K., Chen, C., Van Hummelen, P. & Contreras, R. Old yellow enzyme interferes with Bax-induced NADPH loss and lipid peroxidation in yeast. *FEMS Yeast Res.* **5**, 711–725 (2005).

APPENDIX

LIST OF APPENDIX TABLES

Table A1. Strains used throughout this thesis....68
Table A2. Plasmids used throughout this thesis. 71
Table A3. Nucleotide and amino acid sequences of genes and inserts used throughout this thesis....71
Table A4. Primers for gRNA amplification....72
Table A5. Primers to amplify genes for geraniol optimization....72
Table A6. Primers to amplify *G10H* variants. 73
Table A7. Primers to amplify N-terminally modified *G10H* variants. 74
Table A8. Primers to amplify *CPR*s 75
Table A9. Primers for *CrG10H* optimization 76
Table A10. Primers to amplify UP and DOWN regions of integration loci 76
Table A11. Averages and standard deviations of geraniol for CYP-CPR combinatorial analysis 81
Table A12. Averages and standard deviations of 10HG for CYP-CPR combinatorial analysis....82

LIST OF APPENDIX FIGURES

Figure A1. Initial phylogenetic tree of G10H variants 78
Figure A2. Identification of conserved cytochrome p450 structural domains for the G10H variants. 79
Figure A3. Identification of neric acid in *S. cerevisiae* strains expressing *Cr*G10H(Nt)-*Gr*G10H. 80
Figure A4. Monoterpene production from *S. cerevisiae* strains expressing 3x-*Cr*G10H with different CPRs. 83
Figure A5. Isopulegol production from engineered *S. cerevisiae* strains expressing G10H variants. 84
Figure A6. Isopulegol titers from *S. cerevisiae* strains expressing 1x- and 3x- *CrG10H* with different CPRs. 85

Table A1. Strains used throughout this thesis

Group	Strain Name	Integration		Background	Source
		Integrant	Locus		
WT	CEN.PK113	-	-	*MATα; leu2-3, 112; ura3-*	EUROPHINS

				52; MAL2-8C; SUC2	
Geraniol optimization	MD1	P_{TPI1}-***erg20^{K197E}***-loxP-kanMX	ERG20	CEN.PK113	[61]
	MD2	P_{PGK}-***tHMGR***-T_{TDH2}-P_{TEF1}-***IDI***-T_{CYC1}	USERXII-2	MD1	This study
	MD3	P_{TEF2}-***ObGES***-T_{TDH2}	FgF20	MD2	This study
	MD4	P_{TEF2}-***t63ObGES***-T_{TDH2}	FgF20	MD2	This study
	MD5	P_{PYK1}-***AgGPPS***-T_{ENO2}-P_{PDC1}-***CrCPR***-T_{ADH2}	USERXII-3	MD4	This study
G10H variants	MD6	P_{TEF1}-***CrG10H***-T_{ADH1}	USERXII-5	MD5	This study
	MD7	P_{TEF1}-***BsG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD8	P_{TEF1}-***EgG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD9	P_{TEF1}-***GrG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD10	P_{TEF1}-***GriG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD11	P_{TEF1}-***HaG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD12	P_{TEF1}-***HrG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD13	P_{TEF1}-***InG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD14	P_{TEF1}-***JrG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD15	P_{TEF1}-***PkG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD16	P_{TEF1}-***PpG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD17	P_{TEF1}-***RcG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD18	P_{TEF1}-***SaG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD19	P_{TEF1}-***StG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD20	P_{TEF1}-***VvG10H***- T_{ADH1}	USERXII-5	MD5	This study
N-terminal exchange	MD21	P_{TEF1}-$_{StG10H(Nt)}$***CrG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD22	P_{TEF1}-$_{CrG10H(Nt)}$***BsG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD23	P_{TEF1}-$_{CrG10H(Nt)}$***GrG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD24	P_{TEF1}-$_{CrG10H(Nt)}$***GriG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD25	P_{TEF1}-$_{CrG10H(Nt)}$***HaG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD26	P_{TEF1}-$_{CrG10H(Nt)}$***HrG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD27	P_{TEF1}-$_{CrG10H(Nt)}$***JrG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD28	P_{TEF1}-$_{CrG10H(Nt)}$***PkG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD29	P_{TEF1}-$_{CrG10H(Nt)}$***RcG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD30	P_{TEF1}-$_{CrG10H(Nt)}$***SaG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD31	P_{TEF1}-$_{CrG10H(Nt)}$***StG10H***- T_{ADH1}	USERXII-5	MD5	This study
	MD32	P_{TEF1}-$_{CrG10H(Nt)}$***VvG10H***- T_{ADH1}	USERXII-5	MD5	This study
GFP strains	MD33	P_{TDH3}-***CrG10H-ENVY***-T_{TDH1}	FgF16	BY4741	This study
	MD34	P_{TDH3}-$_{StG10H(Nt)}$***CrG10H- ENVY***-T_{TDH1}	FgF16	BY4741	This study
	MD35	P_{TDH3}-***CYP719A21- ENVY***-T_{TDH1}	FgF16	BY4741	This study
CrG10H optimization	MD36	P_{TEF1}-***CrG10H***- T_{ADH1}	Fgf16	MD5	This study
	MD37	P_{TEF1}-***CrG10H***- T_{ADH1}	Fgf24	MD36	This study
	MD38	P_{PGK}-***CrG10H-CrCPR***- T_{ADH1}	106a	MD37	This study
CrG10H - CPR swap	MD39	*AtCPR*	*CrCPR*	MD6	This study
	MD40	*HcCPR*	*CrCPR*	MD6	This study
	MD41	*NdCPR*	*CrCPR*	MD6	This study
	MD42	*NsCPR*	*CrCPR*	MD6	This study
	MD43	*PkCPR*	*CrCPR*	MD6	This study
	MD44	*Sr1CPR*	*CrCPR*	MD6	This study
	MD45	*Sr8CPR*	*CrCPR*	MD6	This study
	MD46	*TcCPR*	*CrCPR*	MD6	This study
	MD47	*VvCPR*	*CrCPR*	MD6	This study
BsG10H - CPR swap	MD48	*HcCPR*	*CrCPR*	MD7	This study
	MD49	*NdCPR*	*CrCPR*	MD7	This study
	MD50	*NsCPR*	*CrCPR*	MD7	This study
	MD51	*PkCPR*	*CrCPR*	MD7	This study
	MD52	*Sr1CPR*	*CrCPR*	MD7	This study
	MD53	*Sr8CPR*	*CrCPR*	MD7	This study
	MD54	*TcCPR*	*CrCPR*	MD7	This study
	MD55	*VvCPR*	*CrCPR*	MD7	This study
GrG10H - CPR swap	MD56	*AtCPR*	*CrCPR*	MD9	This study
	MD57	*HcCPR*	*CrCPR*	MD9	This study
	MD58	*NdCPR*	*CrCPR*	MD9	This study
	MD59	*NsCPR*	*CrCPR*	MD9	This study
	MD60	*Sr1CPR*	*CrCPR*	MD9	This study
	MD61	*Sr8CPR*	*CrCPR*	MD9	This study
	MD62	*TcCPR*	*CrCPR*	MD9	This study
	MD63	*VvCPR*	*CrCPR*	MD9	This study
GriG10H - CPR swap	MD64	*HcCPR*	*CrCPR*	MD10	This study
	MD65	*NdCPR*	*CrCPR*	MD10	This study
	MD66	*NsCPR*	*CrCPR*	MD10	This study
	MD67	*PkCPR*	*CrCPR*	MD10	This study
	MD68	*Sr8CPR*	*CrCPR*	MD10	This study

	MD69	*VvCPR*	*CrCPR*	MD10	This study
HaG10H - CPR swap	MD70	*AtCPR*	*CrCPR*	MD11	This study
	MD71	*HcCPR*	*CrCPR*	MD11	This study
	MD72	*NdCPR*	*CrCPR*	MD11	This study
	MD73	*NsCPR*	*CrCPR*	MD11	This study
	MD74	*PkCPR*	*CrCPR*	MD11	This study
	MD75	*Sr1CPR*	*CrCPR*	MD11	This study
	MD76	*Sr8CPR*	*CrCPR*	MD11	This study
	MD77	*TcCPR*	*CrCPR*	MD11	This study
	MD78	*VvCPR*	*CrCPR*	MD11	This study
HrG10H-CPR swap	MD79	*AtCPR*	*CrCPR*	MD12	This study
	MD80	*HcCPR*	*CrCPR*	MD12	This study
	MD81	*NdCPR*	*CrCPR*	MD12	This study
	MD82	*NsCPR*	*CrCPR*	MD12	This study
	MD83	*PkCPR*	*CrCPR*	MD12	This study
	MD84	*Sr1CPR*	*CrCPR*	MD12	This study
	MD85	*Sr8CPR*	*CrCPR*	MD12	This study
	MD86	*TcCPR*	*CrCPR*	MD12	This study
	MD87	*VvCPR*	*CrCPR*	MD12	This study
JrG10H-CPR swap	MD88	*AtCPR*	*CrCPR*	MD14	This study
	MD89	*HcCPR*	*CrCPR*	MD14	This study
	MD90	*NdCPR*	*CrCPR*	MD14	This study
	MD91	*NsCPR*	*CrCPR*	MD14	This study
	MD92	*PkCPR*	*CrCPR*	MD14	This study
	MD93	*Sr1CPR*	*CrCPR*	MD14	This study
	MD94	*Sr8CPR*	*CrCPR*	MD14	This study
	MD95	*TcCPR*	*CrCPR*	MD14	This study
	MD96	*VvCPR*	*CrCPR*	MD14	This study
PkG10H-CPR swap	MD97	*AtCPR*	*CrCPR*	MD15	This study
	MD98	*HcCPR*	*CrCPR*	MD15	This study
	MD99	*NdCPR*	*CrCPR*	MD15	This study
	MD100	*NsCPR*	*CrCPR*	MD15	This study
	MD101	*PkCPR*	*CrCPR*	MD15	This study
	MD102	*Sr1CPR*	*CrCPR*	MD15	This study
	MD103	*Sr8CPR*	*CrCPR*	MD15	This study
	MD104	*TcCPR*	*CrCPR*	MD15	This study
	MD105	*VvCPR*	*CrCPR*	MD15	This study
PpG10H-CPR swap	MD106	*AtCPR*	*CrCPR*	MD16	This study
	MD107	*HcCPR*	*CrCPR*	MD16	This study
	MD108	*NdCPR*	*CrCPR*	MD16	This study
	MD109	*NsCPR*	*CrCPR*	MD16	This study
	MD110	*PkCPR*	*CrCPR*	MD16	This study
	MD111	*Sr1CPR*	*CrCPR*	MD16	This study
	MD112	*Sr8CPR*	*CrCPR*	MD16	This study
	MD113	*TcCPR*	*CrCPR*	MD16	This study
	MD114	*VvCPR*	*CrCPR*	MD16	This study
RcG10H-CPR swap	MD115	*AtCPR*	*CrCPR*	MD17	This study
	MD116	*HcCPR*	*CrCPR*	MD17	This study
	MD117	*NdCPR*	*CrCPR*	MD17	This study
	MD118	*NsCPR*	*CrCPR*	MD17	This study
	MD119	*PkCPR*	*CrCPR*	MD17	This study
	MD120	*Sr1CPR*	*CrCPR*	MD17	This study
	MD121	*Sr8CPR*	*CrCPR*	MD17	This study
	MD122	*TcCPR*	*CrCPR*	MD17	This study
	MD123	*VvCPR*	*CrCPR*	MD17	This study
SaG10H-CPR swap	MD124	*AtCPR*	*CrCPR*	MD18	This study
	MD125	*HcCPR*	*CrCPR*	MD18	This study
	MD126	*NdCPR*	*CrCPR*	MD18	This study
	MD127	*NsCPR*	*CrCPR*	MD18	This study
	MD128	*PkCPR*	*CrCPR*	MD18	This study
	MD129	*Sr1CPR*	*CrCPR*	MD18	This study
	MD130	*Sr8CPR*	*CrCPR*	MD18	This study
	MD131	*TcCPR*	*CrCPR*	MD18	This study
	MD132	*VvCPR*	*CrCPR*	MD18	This study
StG10H-CPR swap	MD133	*AtCPR*	*CrCPR*	MD19	This study
	MD134	*HcCPR*	*CrCPR*	MD19	This study
	MD135	*NdCPR*	*CrCPR*	MD19	This study
	MD136	*NsCPR*	*CrCPR*	MD19	This study
	MD137	*PkCPR*	*CrCPR*	MD19	This study
	MD138	*Sr1CPR*	*CrCPR*	MD19	This study

	MD139	*TcCPR*	*CrCPR*	MD19	This study
	MD140	*VvCPR*	*CrCPR*	MD19	This study
VvG10H-CPR swap	MD141	*AtCPR*	*CrCPR*	MD20	This study
	MD142	*NdCPR*	*CrCPR*	MD20	This study
	MD143	*NsCPR*	*CrCPR*	MD20	This study
	MD144	*Sr1CPR*	*CrCPR*	MD20	This study
	MD145	*TcCPR*	*CrCPR*	MD20	This study
	MD146	*VvCPR*	*CrCPR*	MD20	This study
3x-*Cr*G10H - CPR swap	MD147	*HcCPR*	*CrCPR*	MD37	This study
	MD148	*NsCPR*	*CrCPR*	MD37	This study
	MD149	*PkCPR*	*CrCPR*	MD37	This study
	MD150	*Sr1CPR*	*CrCPR*	MD37	This study
	MD151	*TcCPR*	*CrCPR*	MD37	This study
OYE deletions	MD152	LP2.T10	*OYE3*	MD6	This study
	MD153	LP2.T10	*OYE2*	MD152	This study
	MD154	LP2.T10	*OYE2/OYE3*	MD37	This study

Table A2. Plasmids used throughout this thesis.

Plasmid	Relevant Genotype	Source
pCas-Tyr	P_{RNR2}-*Cas9_NLS*-T_{CYC1}-*pUC-2µOri-tRNA_Tyr-3'HDV-gRNAxScaffold*-T_{SNR52}-P_{TEF1}-*hphNT1*-T_{TEF1}	Modified from Ryan et al. (2014) Addgene #60847
pCas-Tyr-gRNA1	gRNA1 is targeted to USERXII_2	This study
pCas-Tyr-gRNA2	gRNA2 is targeted to FgF20	This study
pCas-Tyr-gRNA3	gRNA3 is targeted to USERXII_5	This study
pCas-Tyr-gRNA4	gRNA4 is targeted to FgF16	This study
pCas-Tyr-gRNA5	gRNA5 is targeted to FgF18	This study
pCas-Tyr-gRNA6	gRNA6 is targeted to 106a	This study
pCas-Tyr-gRNA7	gRNA7 is targeted to *Cr*CPR	This study
pCas-Tyr-gRNA8	gRNA8 is targeted to OYE3	This study
pCas-Tyr-gRNA9	gRNA9 is targeted to OYE2	This study
pYES_pGerOH	*2µori, URA3,* P_{PYK1}-***AgGPPS2***-T_{ENO2} -P_{PDC1}-***CrCPR***-T_{ADH2}-P_{TEF2}-***ObGES***-T_{TDH2} -P_{FBA1}-***CrG10H***-T_{PGI1}	Campbell et al. (2016) Thermofisher
pJet_BsG10H	*BsG10H*	This study
pJet_HaG10H	*HaG10H*	This study
pJet_HrG10H	*HrG10H*	This study
pJet_InG10H	*InG10H*	This study
pJet_JrG10H	*JrG10H*	This study
pJet_PkG10H	*PkG10H*	This study
pJet_PpG10H	*PpG10H*	This study
pJet_RcG10H	*RcG10H*	This study
pJet_SaG10H	*SaG10H*	This study
pJet_StG10H	*StG10H*	This study

Table A3. Nucleotide and amino acid sequences of genes and inserts used throughout this thesis

NOTE: See attached excel file

Table A4. Primers for gRNA amplification

Target site	Primer ID	Sequence (5' -> 3')	Description	5'-N20gRNA-3'
USERXII_2	MD134	TCGAGAGAGTCGCCGATAGTGTTTTAGAGCTAGAAATAGCAAGT	USERXII2_F with 5' homology to scaffold	TCGAGAGAGTCGCCGATAGT
	MD135	ACTATCGGCGACTCTCTCGAAAAGTCCCATTCGCC	USERXII2_F with 3' homology to HDV	
FgF20	FF154	GTTAGAGCTGTTACAAGTTAGTTTTAGAGCTAGAAATAGCAAGT	FgF20_F with 5' homology to scaffold	GTTAGAGCTGTTACAAGTTA
	FF155	TAACTTGTAACAGCTCTAACAAAGTCCCATTCGCC	FgF20_F with 3' homology to HDV	
USERXII_5	MD3	TTGTCACAGTGTCACATCAGGTTTTAGAGCTAGAAATAGCAAGT	USERXII5_F with 5' homology to scaffold	TTGTCACAGTGTCACATCAG
	MD4	CTGATGTGACACTGTGACAAAAAGTCCCATTCGCC	USERXII5_F with 3' homology to HDV	
FgF16	LB944	TGTACCAAAAGTTATCCTGTGTTTTAGAGCTAGAAATAGCAAGT	FgF16_F with 5' homology to scaffold	TGTACCAAAAGTTATCCTGT
	LB945	ACAGGATAACTTTTGGTACAAAAGTCCCATTCGCCACCCGAA	Fgf16_F with 3' homology to HDV	
FgF24	FF166	CCTATTGGACAAGATTTACGGTTTTAGAGCTAGAAATAGCAAGT	Fgf24_F with 5' homology to scaffold	CCTATTGGACAAGATTTACG
	FF167	CGTAAATCTTGTCCAATAGGAAAGTCCCATTCGCC	Fgf24a_F with 3' homology to HDV	
106a	LB1549	atacggtcagggtagcgcccGTTTTAGAGCTAGAAATAGCAAGT	106a_F with 5' homology to scaffold	ATACGGTCAGGGTAGCGCCC
	LB1550	gggcgctaccctgaccgtatAAAGTCCCATTCGCCACCCGAA	106a_F with 3' homology to HDV	
*Cr*CPR	MD357	TTTGTTGAAGGCAACGACAGGTTTTAGAGCTAGAAATAGCAAGT	CrCPR_F with 5' homology to scaffold	TTTGTTGAAGGCAACGACAG
	MD358	CTGTCGTTGCCTTCAACAAAAAAAGTCCCATTCGCC	CrCPR_F with 3' homology to HDV	
OYE3	MD305	CTCATCAGACCAAATCCCAGGTTTTAGAGCTAGAAATAGCAAGT	OYE3_F with 5' homology to scaffold	CTCATCAGACCAAATCCCAG
	MD306	CTGGGATTTGGTCTGATGAGAAAAGTCCCATTCGCC	OYE3_F with 3' homology to HDV	
OYE2	MD303	GTAACCCCCAGATTGTGGAGGTTTTAGAGCTAGAAATAGCAAGT	OYE2_F with 5' homology to scaffold	GTAACCCCCAGATTGTGGAG
	MD304	CTCCACAATCTGGGGGTTACAAAAGTCCCATTCGCC	OYE2_F with 3' homology to HDV	
pCas universal gRNA primers	LB33	CACCTATATCTGCGTGTTGC	pCas_HDV_F	-
	LB34	GTCAAGACTGTCAAGGAGG	pCas-scaffold_R	

Table A5. Primers to amplify genes for geraniol optimization.

Amplicon	Primer ID	5'-Sequence-3'	Description
tHMGR - IDI cassette	MD139	AACGAAAAAGAAAAGAAAGACCATGTCATGTACGGGCAATCAGAATCTGTAACAAGCGCCacgcacagatattataacatctgcataatag	tHMGR cassette_F with 5' overhang
	MD168	gtaaggattcgcggtcctcgaaaattaaaagtccaacgcgcctgttgcttGCGAAAAGCCAATTAGTGTG	tHMGR cassette_R with 3' homology to ISI cassette
	MD171	ATCACGGATTTTCGATAAAGCACTTAGTATCACACTAATTGGCTTTTCGCaagcaacaggcgcgttg	IDI cassette_F with 5' homology to tHMGR cassette
	MD140	CTATTTCCTATAATAGAAATCCAAGTGGCAAAAGCGTTAGACGCAGTACAAGGACGCGTTAAGGCAAATTAAAGCCTTCGAGCG	IDI cassette_R with 3' overhang

ObGES/ tObGES	payge122	AGGAATACTCTGAATAAAACAACTTATATAATAAAAATG Cctatatggggccgtatact	full length ObGES cassette_F with 5' homology to LV3
	payge130	TGTTATAATTATTTTCTTATTTTGATGTAATATAAAGAGG gcgaaaagccaattagt	full length ObGES cassette_R with 5' homology to LV5
	MD146	GTACTTGTTTTTAGAATATACGGTCAACGAACTATAATT AAAAAACAATGCAACACATGGAGGAGAGCAGC	Truncation of ObGES at position 63_F with 5' homology to TEF2p
AgGPPS - CrCPR cassette	MD158	aggaatactctgaataaaacaacttatataataaaaatgcAATGCTACTATTTTG GAGATTAATCTCAG	AgGPPS cassette_F with 5' homology to LV3
	MD169	TTTATCTTGCACATCACATCAGCGGAACATATGCTCACC CAGTCGCATGTAGGTATCATCTCCATCTCCCATATG	AgGPPS cassette_R with 3' homology to CrCPR cassette
	MD170	TGCGGGCCACGACCACAGTGATATGCATATGGGAGATG GAGATGATACCTACATGCGACTGGGTGAG	CrCPR cassette_F with 5' homology to GPPS cassette
	payge126	TGTTATAATTATTTTCTTATTTTGATGTAATATAAAGAG Gtagaattatataacttgatgagatgag	CrCPR cassette_R with 3' homology to LV5

Table A6. Primers to amplify G10H variants.

Amplicon	Primer ID	5'-Sequence-3'	Description
BsG10H	MD178	Aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaat AAAACAATGGACATCTTAAGCTC	BsG10H_F with 5' homology to TEF1p
	MD180	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATTATATTG CGATCGGAACGG	BsG10H_R with 3' homology to ADH1t
CrG10H	MD187	Aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaat AAAACAATGGATTACTTAACTATCATATTGAC	CrG10H_F with 5' homology to TEF1p
	MD189	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATCACAGGG TAGAAGGCAC	CrG10H_R with 3' homology to ADH1t
EgG10H	Md191	aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaatA AAACAATGGATGTTTGGGTGTC	EgG10H_F with 5' homology to TEF1p
	MD193	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATTACAACG GTTTAGGAATAACCC	EgG10H_R with 3' homology to ADH1t
GrG10H	MD195	aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaatA AAACAATGAGAGAAATGGATCTGC	GrG10H_F with 5' homology to TEF1p
	MD197	AGGTAGACAAGCCGACAACCTTGATTGGAGACTTG ACCAAACCTCTGGCGAAGAAGTCCATTATATCACA ACTGGGATAGCCTG	GrG10H_R with 3' homology to ADH1t
GriG10H	MD199	Aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaat AAAACAATGTTGGGCTCAAATAAGTC	GriG10H_F with 5' homology to TEF1p
	MD201	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATTAAAGGG ATGTTGGGATGGC	GriG10H_R with 3' homology to ADH1t
HaG10H	MD203	aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaatA AAACAATGGGTTTTGTCATCGTTG	HaG10H_F with 5' homology to TEF1p
	MD205	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATTAATTTA GGGGTATCGGCACG	HaG10H_R with 3' homology to ADH1t

HrG10H	MD207	aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaatA AAACAATGGACTTCTTGGTATTCG	HrG10H_F with 5' homology to TEF1p
	MD209	AGGTAGACAAGCCGACAACCTTGATTGGAGACTTG ACCAAACCTCTGGCGAAGAAGTCCATTAAACGGGT ACTATTAATAGTGGGTCA	HrG10H_R with 3' homology to ADH1t
InG10H	MD211	aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaatA AAACAATGGATATGGACTACCAGAG	InG10H_F with 5' homology to TEF1p
	MD213	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATTAGAAGG GAATCGGTATAGCC	InG10H_R with 3' homology to ADH1t
JrG10H	MD215	aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaatA AAACAATGGACTTTTTGGGGTTG	JrG10H_F with 5' homology to TEF1p
	MD217	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATTATCTCA GAATTGGAACAGCACG	JrG10H_R with 3' homology to ADH1t
PkG10H	MD219	aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaatA AAACAATGGGCTTTTTGTCTGC	PkG10H_F with 5' homology to TEF1p
	MD221	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATTATACAT GAAAACTTGGTATCGCTAG	PkG10H_R with 3' homology to ADH1t
PpG10H	MD223	aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaatA AAACAATGGATTTTTTAACTATAATTCTGGG	PpG10H_F with 5' homology to TEF1p
	MD225	AGGTAGACAAGCCGACAACCTTGATTGGAGACTTG ACCAAACCTCTGGCGAAGAAGTCCATTACAGTGGT ATTGGTACAGCC	PpG10H_R with 3' homology to ADH1t
RcG10H	MD227	aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaatA AAACAATGATGGACCTTCTTGTAG	RcG10H_F with 5' homology to TEF1p
	MD229	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATTAGACCT GATTGGGTATGGC	RcG10H_R with 3' homology to ADH1t
SaG10H	MD231	aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaatA AAACAATGGATTTCCTTGGTTTTATAC	SaG10H_F with 5' homology to TEF1p
	MD233	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATTAACGAA TGGGTACAGCACATAATG	SaG10H_R with 3' homology to ADH1t
StG10H	MD235	aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaatA AAACAATGGAGTATGTAAACATCCTTC	StG10H_F with 5' homology to TEF1p
	MD237	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATTAGTATC CCAGAAGCTGTAGGG	StG10H_R with 3' homology to ADH1t
VvG10H	MD239	Aacttttttacttcttgctcattagaaagaaagcatagcaatctaatctaagttttaat AAAACAATGGATTACACTCCGC	VvG10H_F with 5' homology to TEF1p
	MD241	AGGTAGACAAGCCGACAACCTTGATTGGAGACTT GACCAAACCTCTGGCGAAGAAGTCCATTAGGGTT TAGTCGGTACGG	VvG10H_R with 3' homology to ADH1t

Table A7. Primers to amplify N-terminally modified G10H variants.

Amplicon	Primer ID	5'-Sequence-3'	Description
CrG10H(Nt)-BsG10H	MD398	GTTCGCCTTGACATTGTATGAAGCCTTTAGTT ACTTGTCAAAAGGACGTAAACCCACCTC	Removal of BsG10H(Nt) with 5' homolgy to CrG10H(Nt)
StG10H(Nt)-CrG10H	MD400	ATTCGCCTGTTTCCTGGTTCGTGTCGTATTTAT AAGCTTGTACTTGTCACGTAGAACCAAAAAC	Removal of CrG10H(Nt) with 5' homolgy to StG10H(Nt)
CrG10H(Nt)-GrG10H	MD403	GTTCGCCTTGACATTGTATGAAGCCTTTAGT TACTTGTCAATCAAAAACGGGAACAAAAGC TC	Removal of GRG10H(Nt) with 5' homolgy to CrG10H(Nt)
CrG10H(Nt)-GriG10H	MD404	GTTCGCCTTGACATTGTATGAAGCCTTTAGTT ACTTGTCATCCTTCAGTCGTAAAAGTAAGAA CC	Removal of GriG10H(Nt) with 5' homolgy to CrG10H(Nt)
CrG10H(Nt)-HaG10H	MD405	GTTCGCCTTGACATTGTATGAAGCCTTTAGT TACTTGTCAGTCGGACGTCCCAAGAAC	Removal of HaG10H(Nt) with 5' homolgy to CrG10H(Nt)
CrG10H(Nt)-HrG10H	MD406	GTTCGCCTTGACATTGTATGAAGCCTTTAGTT ACTTGTCAACGTCCCTTAGTAGAAGGAAG	Removal of HrG10H(Nt) with 5' homolgy to CrG10H(Nt)
CrG10H(Nt)-JrG10H	MD408	GTTCGCCTTGACATTGTATGAAGCCTTTAGTT ACTTGTCAAGCTTACTGGGGTGGAAAAAAAG	Removal of JrG10H(Nt) with 5' homolgy to CrG10H(Nt)
CrG10H(Nt)-PkG10H	MD409	GTTCGCCTTGACATTGTATGAAGCCTTTAGTT ACTTGTCATCTTCCAAAGGAAAGCGTTTG	Removal of PkG10H(Nt) with 5' homolgy to CrG10H(Nt)
CrG10H(Nt)-RcG10H	MD411	GTTCGCCTTGACATTGTATGAAGCCTTTAGTT ACTTGTCAGCCCTGACTAGCATTAGTAAAC	Removal of RcG10H(Nt) with 5' homolgy to CrG10H(Nt)
CrG10H(Nt)-SaG10H	MD412	GTTCGCCTTGACATTGTATGAAGCCTTTAGTT ACTTGTCAGCGCTACGTTCCCTGTCTAG	Removal of SaG10H(Nt) with 5' homolgy to CrG10H(Nt)
CrG10H(Nt)-StG10H	MD413	GTTCGCCTTGACATTGTATGAAGCCTTTAGTT ACTTGTCAAGCTTGCGTAGGTCCAAAAG	Removal of StG10H(Nt) with 5' homolgy to CrG10H(Nt)
CrG10H(Nt)-VvG10H	MD414	GTTCGCCTTGACATTGTATGAAGCCTTTAGT TACTTGTCAATGGATTACACTCCGCTTGTTC	Removal of VvG10H(Nt) with 5' homolgy to CrG10H(Nt)

Table A8. Primers to amplify CPRs

Amplicon	Primer ID	5'-Sequence-3'	Description
AtCPR	MD335	TATCTTCTACTCATAACCTCACGCAAAATAACACA GTCAAATCAAAAACAATGAGTTCATCCAGTTCTTC TTC	AtCPR_F with 5' homology to PDC1p
	MD336	ATGCTTGATAATGAAAACTATAAATCGTAAAGAC ATAAGATCCGCCTAGCTCTTCAGCCCCAGAC	AtCPR_R with 3' homology to ADH2t
HcCPR	MD339	TATCTTCTACTCATAACCTCACGCAAAATAACACA GTCAAATCAAAAACAATGGAGAGCAGTAGCGTAA AG	HcCPR_F with 5' homology to PDC1p
	MD340	GCATGCTTGATAATGAAAACTATAAATCGTAAAGA CATAAGATCCGCCTAGCTCTTCAGCCCCATACG	HcCPR_R with 3' homology to ADH2t
NdCPR	MD343	TATCTTCTACTCATAACCTCACGCAAAATAACACAG TCAAATCAAAAACAATGCAAGAAGCGTCCAGC	NdCPR_F with 5' homology to PDC1p
	MD344	GCATGCTTGATAATGAAAACTATAAATCGTAA AGACATAAGATCCGCCTAGCTCTTCAGCCCCAA ACATC	NdCPR_R with 3' homology to ADH2t

NsCPR	MD345	TATCTTCTACTCATAACCTCACGCAAAATAACA CAGTCAAATCAAAAACAATGGCAAGTAATCTT GTAAATTCTTTAG	NsCPR_F with 5' homology to PDC1p
	MD346	TAGGCATGCTTGATAATGAAAACTATAAATCGT AAAGACATAAGATCCGCCTAGCTCTTCAGCCCC ACAC	NsCPR_R with 3' homology to ADH2t
Sr1CPR	MD349	TATCTTCTACTCATAACCTCACGCAAAATAACA CAGTCAAATCAAAAACAATGCAATCAGATTCA GTCAAAGTC	Sr1CPR_F with 5' homology to PDC1p
	MD350	TAGGCATGCTTGATAATGAAAACTATAAATCG TAAAGACATAAGATCCGCCTACCAAACATCTC TTAAGTATCTTCCAG	Sr1CPR_R with 3' homology to ADH2t
Sr8CPR	MD351	TATCTTCTACTCATAACCTCACGCAAAATAACA CAGTCAAATCAAAAACAATGCAATCTAACTCCG TGAAG	Sr8CPR_F with 5' homology to PDC1p
	MD352	TAGGCATGCTTGATAATGAAAACTATAAATCGT AAAGACATAAGATCCGCCTATTACCAAACGTCA CGGAGG	Sr8CPR_R with 3' homology to ADH2t
TcCPR	MD353	TATCTTCTACTCATAACCTCACGCAAAATAACAC AGTCAAATCAAAAACAATGCAGACTTCCGAAGT CAAAATATC	TcCPR_F with 5' homology to PDC1p
	MD354	TAGGCATGCTTGATAATGAAAACTATAAATCGT AAAGACATAAGATCCGCCTAGCTCTTCAGCCCC ATACG	TcCPR_R with 3' homology to ADH2t
PkCPR	MD375	AATTATTATCTTCTACTCATAACCTCACGCAAA ATAACACAGTCAAATCAAAAACAATGCAATCC ACGC	PkCPR_F with 5' homology to PDC1p
	MD376	TAGGCATGCTTGATAATGAAAACTATAAATCGT AAAGACATAAGATCCGCTTACCAAACGTCTCTT AAATACCTTC	PkCPR_R with 3' homology to ADH2t
VvCPR	MD377	AATTATTATCTTCTACTCATAACCTCACGCAAAA TAACACAGTCAAATCAAAAACAATGCAATCCTC CTCTG	VvCPR_F with 5' homology to PDC1p
	MD378	TAGGCATGCTTGATAATGAAAACTATAAATCGT AAAGACATAAGATCCGCTTACCATACATCACGC AAGTATC	VvCPR_R with 3' homology to ADH2t

Table A9. Primers for *Cr*G10H optimization

Amplicon	Primer ID	5'-Sequence-3'	Description
OYE2	MD295	TCCAGATATAGAATAAATCATCATATTAAGCTAAA TATAGACGATAATATAGTATCGATATGCTCCAAGTG TGTGACTC	LP2.T10_F with 5' homology to OYE2 loci
	MD296	ATATATTCATTAATTATATAAATTAGAAGAAAAAGA AATGGTGCTACAAAGTACGGTTAAGCTTCAAAGACG TGAGTGTG	LP2.T10_R with 3' homology to OYE2 loci
OYE3	MD299	TAATTAAAAATATGGCAGGAATATGAAAAATACATAA CATCAATGTCTTTATTCATGATTTGCTCCAAGTGTGTG ACTC	LP2.T10_F with 5' homology to OYE23loci
	MD300	TTCAGAGATTCTACTCTTGACCACTGTTTCGTGTAGCC GCTCAAGGTTTATTTCTTTCTTGCTTCAAAGACGTGA GTGTG	LP2.T10_R with 3' homology to OYE3 loci
CrG10H-CrCPR fusion	MD384	GTTCGACAACCTTTTTCCCAGAACCTGAACTTGTAG AGCCCAGGGTAGAAGGCACAGC	CrG10H_R with GST linker and 5' homology to tCrCPR
	MD383	TTAAGAGCTGTGCCTTCTACCCTGGGCTCTACAAGT TCAGGTTCTGGGAAAAAGGTTG	tCrCPR_F with GST linker and 3' homology to CrG10H

Table A10. Primers to amplify UP and DOWN regions of integration loci

Integration loci	Primer ID	5'-Sequence-3'	Description
USERXII-3	MD5	GAAACTAACCCGATGGGACAATTAC	UP region_F
	MD6	gcatttttattatataagttgttttattcagagtattcctTACCCCTTAT TATAATGATTAATACTTACATCATAG	UP region_R with 3' homology to LV3
	MD7	cctctttatattacatcaaaataagaaaataattataacaGGAAGTTT TGCAGATGAAGTGC	DOWN region_F with 5' homology to LV5
	MD8	CCAACGCATTTACAAACCACG	DOWN region_R
USERXII-5	MD159	caatctggcggcttgagttc	UP region_F
	MD160	Gcatttttattatataagttgttttattcagagtattccttaccggttctgccac ctc	UP region_R with 3' homology to LV3
	MD161	Cctctttatattacatcaaaataagaaaataattataacactcagaagtttga cagcaagc	DOWN region_F with 5' homology to LV5
	MD162	atactagagttaactgatggtcttaaacag	DOWN region_R
FgF16	MD249	gcatttttattatataagttgttttattcagagtattcctTCCGTTAATT CGGGTTTCAATC	UP region_R with 3' homology to LV3
	MD250	TTCGTGAAACACGTGGGATA	UP region_F
	MD251	TTGTTGGGATTCCATTGTGATTAAG	DOWN region_R
	MD252	cctctttatattacatcaaaataagaaaataattataacaTGCCTACG CAACACTTTAGC	DOWN region_F with 5' homology to LV5
FgF24	MD257	gcatttttattatataagttgttttattcagagtattcctGGATCACCT CGCCCTG	UP region_R with 3' homology to LV3
	MD258	GGCTGAACAACAGTCTCTCC	UP region_F
	MD259	GGCGGTAGTGATAACCATTCTC	DOWN region_R
	MD260	cctctttatattacatcaaaataagaaaataattataacaGTTGAA GTCGCCTGGTAGC	DOWN region_F with 5' homology to LV5
106a	LB1543	ccctgttcatggcaacgtcacc	UP region_F
	LB1544	Gcatttttattatataagttgttttattcagagtattcctcgtatcacaacc gacgatccg	UP region_R with 3' homology to LV3
	LB1545	Cctctttatattacatcaaaataagaaaataattataacacctggtcaa acttcagaactaa	DOWN region_F with 5' homology to LV5
	LB1546	ggattaagtaacagatacagacatcac	DOWN region_R

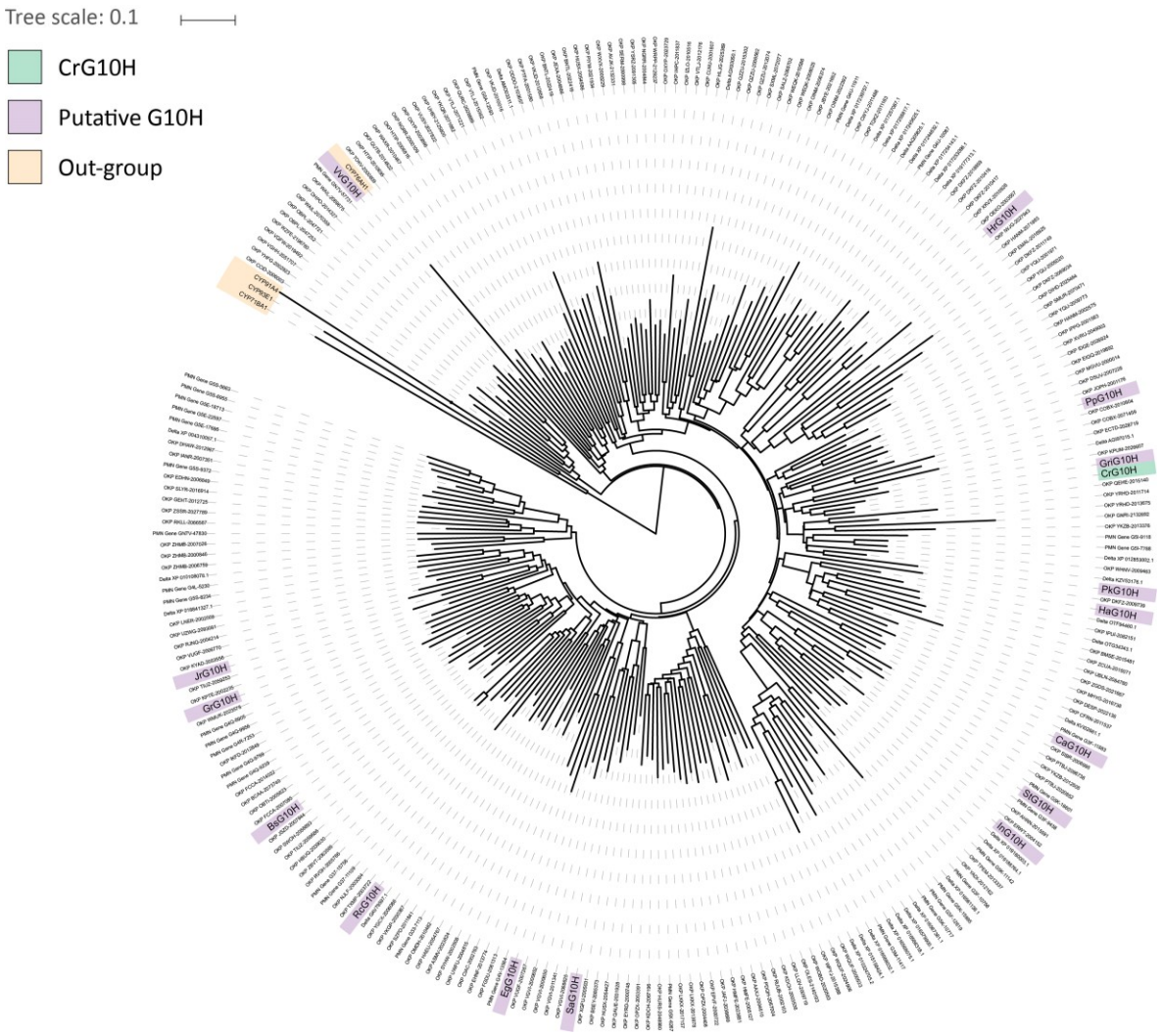

Figure A1. Initial phylogenetic tree of G10H variants
Initial bootstrapped maximum likelihood phylogenetic tree of potential G10H candidates using the *Cr*G10H as the search query (green), from which the final candidates were chosen (purple). The tree was rooted using various CYP outgroups (yellow).

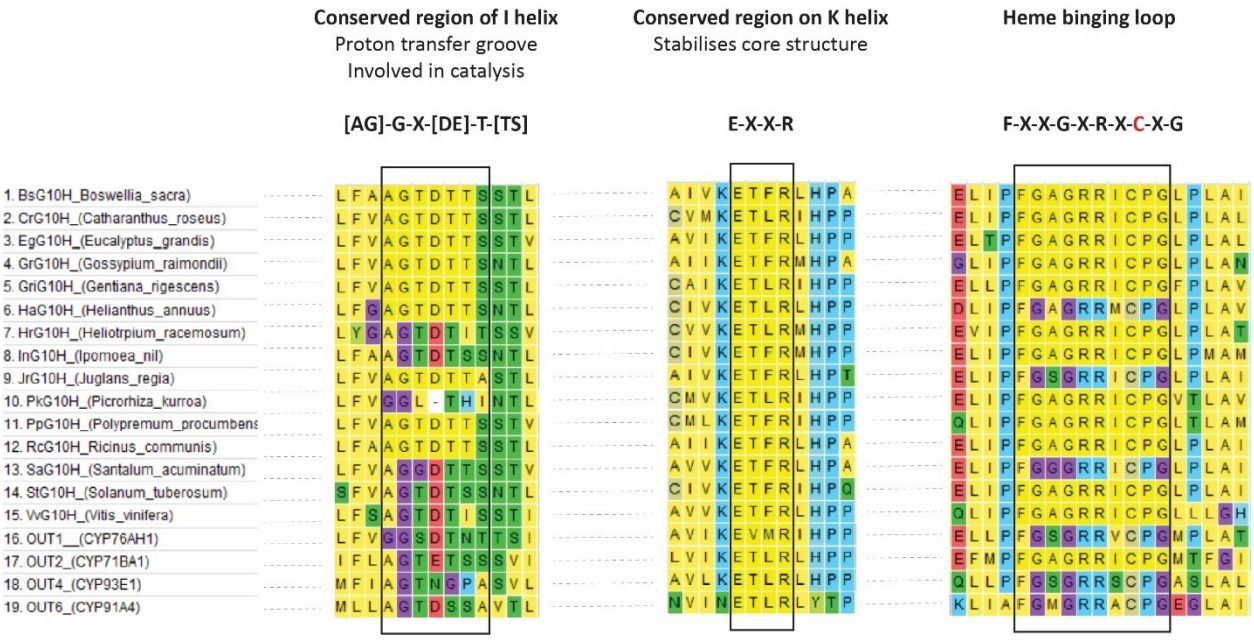

Figure A2. Identification of conserved cytochrome p450 structural domains for the G10H variants.
The G10H candidate list was assessed for their inclusion of the structural conservation normally found in the core of the protein around the heme. The absolutely conserved cysteine residue of the heme binding loop is shown in red in the conserved motif above the figure.

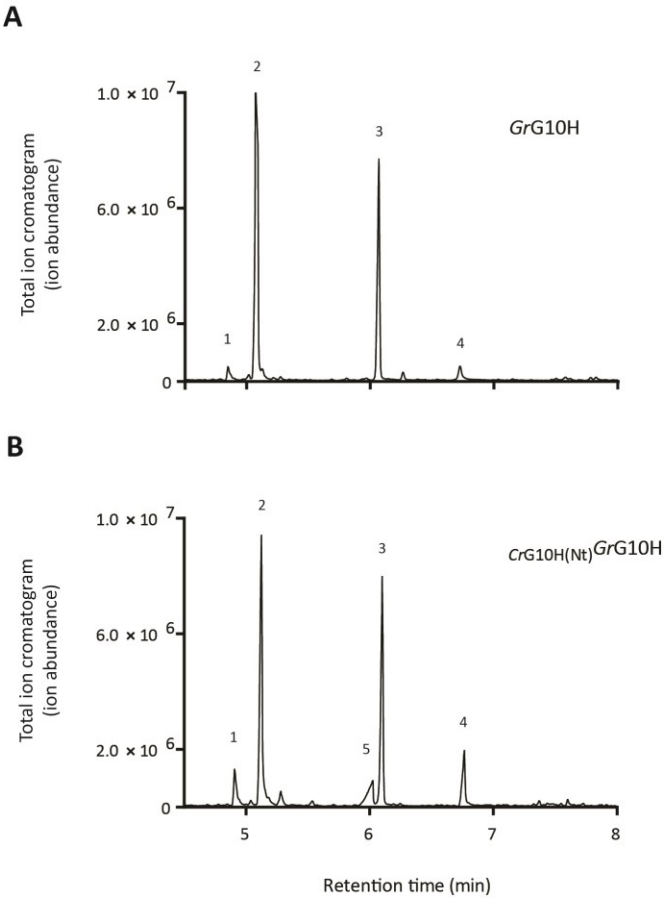

Figure A3. Identification of neric acid in *S. cerevisiae* strains expressing *Cr*G10H(Nt)-*Gr*G10H.
Total ion abundance for **A)** full length *Gr*G10H and **B)** $_{Cr\text{G10H(Nt)}}$-*Gr*G10H strains extracted from culture broths. Peak identification using analytic standards are as follows: **1.** Citronellol; **2.** Geraniol; **3.** Eugenol (internal standard); **4.** 10HG. Peak **5** was identified by the NIST08 MS Search as neric acid.

Table A11. Averages and standard deviations of geraniol for CYP-CPR combinatorial analysis

Average of Geraniol (mg/L)
Normalized to OD600 of 1

		CPR									
		Cr	*At*	*Hc*	*Nd*	*Ns*	*Pk*	*Sr1*	*Sr8*	*Tc*	*Vv*
G10H	*Br*	2.28		0.31	7.68	4.39	7.86	4.41	3.83	2.65	3.29
	Cr	1.71	3.35	0.40	3.71	0.96	0.44	3.15	3.85	1.84	2.86
	Gr	4.87	8.34	2.32	5.43	3.48		3.97	1.78	3.61	3.54
	Gri	3.82		0.14	9.77	9.45	0.11		2.43		2.12
	Ha	4.84	6.98	2.81	2.58	2.43	5.14	3.38	3.95	2.94	4.59
	Hr	3.90	2.47	1.47	1.57	1.26	1.40	1.85	2.99	1.31	1.15
	Jr	4.19	3.62	1.85		6.43	5.03	3.95	7.20	3.72	5.14
	Pk	0.94	0.16	1.17	0.18	1.03	0.93	0.47	0.70	0.34	0.09
	Pp	0.22	0.10	0.07	0.06	0.00	0.02	0.02	0.01	0.06	4.44
	Rc	3.56	2.87	2.80	2.16	3.99	3.31	3.97	3.16	3.76	4.51
	Sa	4.38	5.93	6.78	6.27	7.30	6.11	5.14	3.63	6.52	7.57
	St	0.19	0.15	0.07	0.07	0.08	0.00	0.04		0.27	0.00
	Vv	3.77	0.01		0.03	0.01		3.49		0.03	0.03

Standard Deviation of Geraniol titers (mg/L)
Normalized to OD600 of 1

		CPR									
		Cr	*At*	*Hc*	*Nd*	*Ns*	*Pk*	*Sr1*	*Sr8*	*Tc*	*Vv*
G10H	*Br*	0.54		0.22	0.97	0.50	0.96	0.08	0.29	0.22	0.28
	Cr	0.15	0.34	0.11	0.48	0.17	0.03	0.48	0.40	0.18	0.48
	Gr	0.10	0.45	0.09	2.41	0.13		0.87	1.08	1.05	0.22
	Gri	0.10		0.07	0.44	0.32	0.02		0.22		0.40
	Ha	0.51	0.48	0.35	0.16	0.11	2.26	0.40	0.15	0.11	0.29
	Hr	0.22	0.16	0.02	0.09	0.04	0.26	0.21	0.14	0.13	0.02
	Jr	0.35	0.37	0.21		0.80	0.24	0.46	0.17	0.31	0.93
	Pk	0.07	0.06	0.16	0.02	0.20	0.11	0.08	0.19	0.16	0.10
	Pp	0.04	0.08	0.01	0.01	0.00	0.02	0.02	0.02	0.10	0.45
	Rc	0.20	0.94	0.32	0.28	0.45	0.52	0.13	0.20	0.09	0.75
	Sa	0.21	1.10	0.83	0.13	0.55	0.50	0.55	1.42	0.08	0.34
	St	0.11	0.10	0.06	0.03	0.14	0.00	0.05		0.22	0.00
	Vv	0.26	0.01		0.06	0.02		0.42		0.03	0.01

Table A12. Averages and standard deviations of 10HG for CYP-CPR combinatorial analysis

Average of 10-Hydroxygeraniol titers (mg/L)
Normalized to OD600 of 1

		CPR									
		Cr	*At*	*Hc*	*Nd*	*Ns*	*Pk*	*Sr1*	*Sr8*	*Tc*	*Vv*
G10H	*Br*	0.00		0.08	0.12	0.20	0.11	0.24	0.19	0.14	0.15
	Cr	1.03	1.41	2.29	1.81	2.37	1.54	1.04	1.10	1.28	1.56
	Gr	0.09	0.15	0.12	0.09	0.06		0.04	0.07	0.05	0.37
	Gri	0.00		0.06	0.12	0.15	0.02		0.06		0.14
	Ha	0.01	0.00	0.10	0.18	0.11	0.00	0.18	0.17	0.17	0.20
	Hr	0.00	0.06	0.11	0.03	0.00	0.05	0.11	0.01	0.02	0.00
	Jr	0.00	0.18	0.00		0.11	0.33	0.14	0.09	0.23	0.03
	Pk	0.10	0.17	0.53	0.09	0.55	0.24	0.48	0.02	0.19	0.04
	Pp	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
	Rc	0.00	0.14	0.22	0.17	0.00	0.00	0.25	0.08	0.12	0.08
	Sa	0.00	0.39	0.39	0.09	0.01	0.11	0.01	0.02	0.03	0.26
	St	0.00	0.09	0.04	0.10	0.08	0.07	0.01		0.06	0.02
	Vv	0.00	0.02		0.06	0.00		0.08		0.00	0.01

Standard Deviation of 10-Hydroxygeraniol titers (mg/L)
Normalized to OD600 of 1

		CPR									
		Cr	*At*	*Hc*	*Nd*	*Ns*	*Pk*	*Sr1*	*Sr8*	*Tc*	*Vv*
G10H	*Br*	0.00		0.03	0.02	0.02	0.01	0.03	0.01	0.10	0.13
	Cr	0.16	0.31	0.23	0.13	0.23	0.26	0.09	0.10	0.21	0.21
	Gr	0.10	0.01	0.10	0.08	0.06		0.04	0.06	0.04	0.05
	Gri	0.00		0.05	0.02	0.00	0.02		0.04		0.11
	Ha	0.03	0.00	0.07	0.07	0.12	0.00	0.03	0.03	0.06	0.03
	Hr	0.00	0.02	0.06	0.06	0.00	0.09	0.05	0.02	0.03	0.00
	Jr	0.00	0.04	0.00		0.02	0.04	0.01	0.01	0.10	0.04
	Pk	0.07	0.06	0.13	0.00	0.05	0.13	0.11	0.03	0.08	0.01
	Pp	0.00	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00
	Rc	0.00	0.03	0.07	0.09	0.00	0.00	0.03	0.07	0.09	0.01
	Sa	0.00	0.21	0.10	0.01	0.01	0.01	0.01	0.01	0.04	0.04
	St	0.00	0.07	0.04	0.04	0.08	0.01	0.01		0.05	0.01
	Vv	0.00	0.01		0.04	0.00		0.03		0.00	0.02

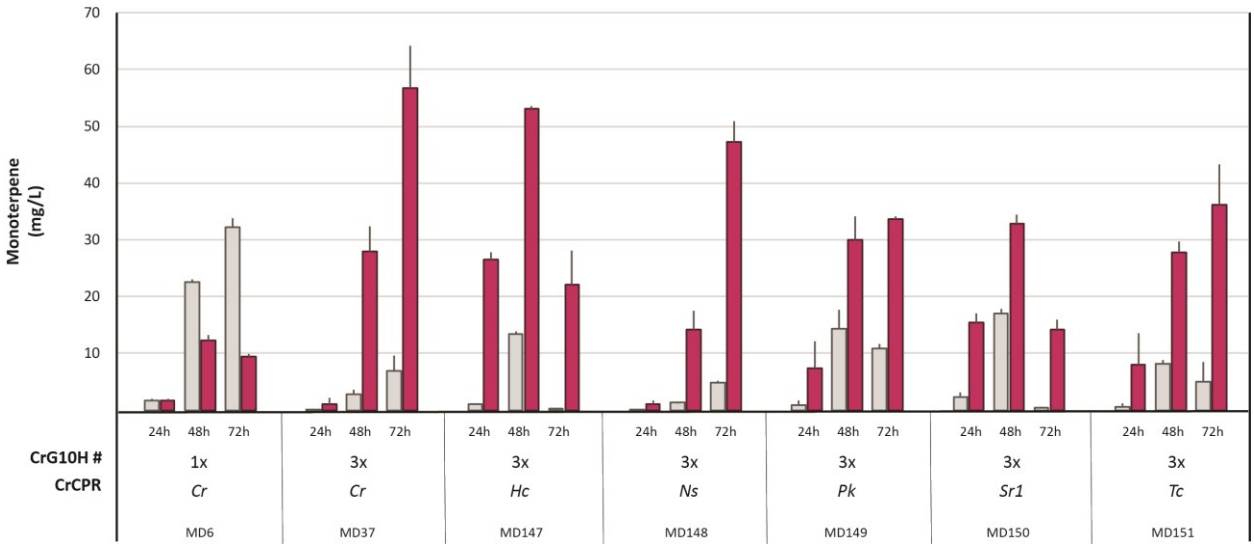

Figure A4. Monoterpene production from *S. cerevisiae* strains expressing 3x-*CrG10H* with different CPRs. *Sr1*CPR was included as it had a similar percent molar conversion as the *Cr*CPR strain. Geraniol (grey) and 10HG (pink) titers are reported non-OD600 normalized and were quantified at 24,48 and 72 hours. Error bars represent the standard deviation from triplicate cultures.

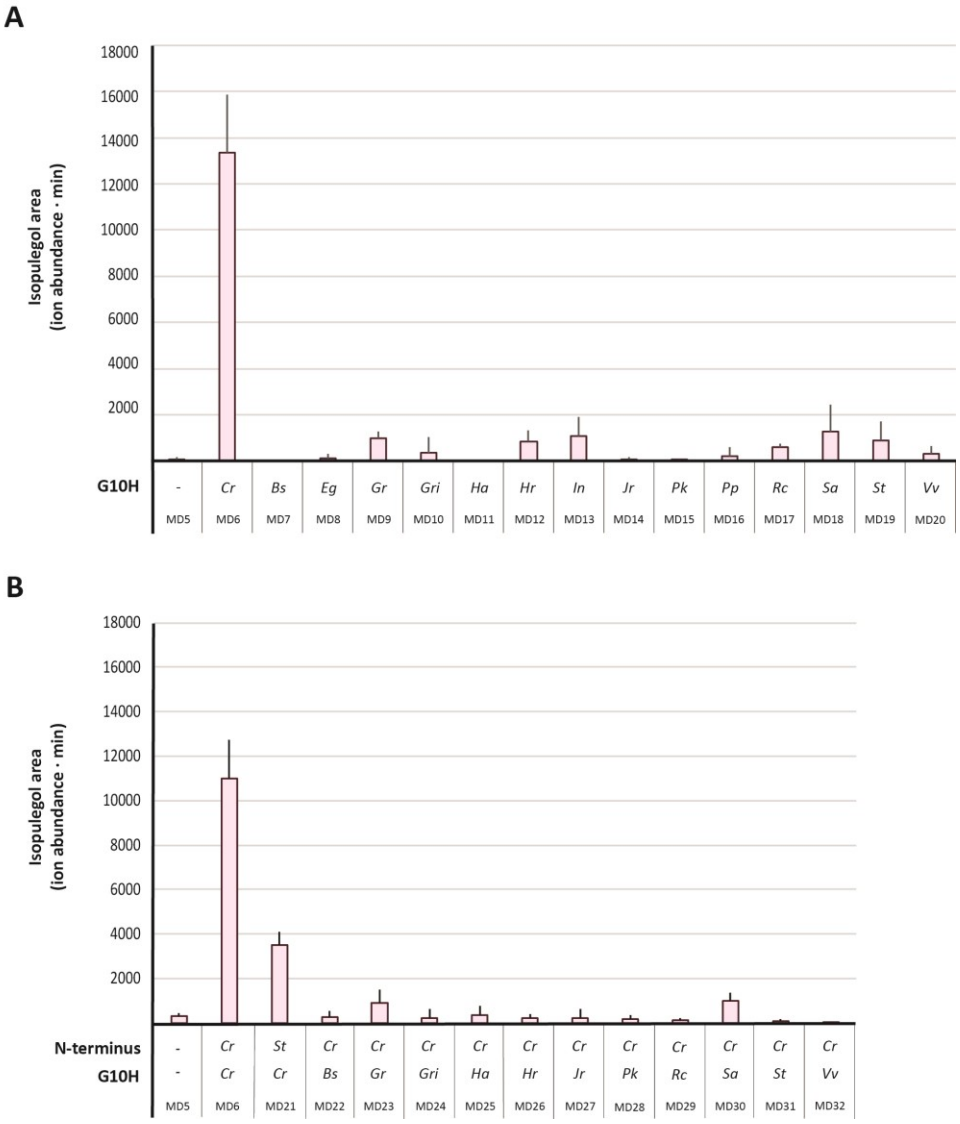

Figure A5. Isopulegol production from engineered *S. cerevisiae* strains expressing G10H variants.
A) Isopulegol titers from the full-length G10H variant library. **B)** Isopulegol titers from the N-terminally modified G10H variant library. Error bars represent the standard deviation from triplicate cultures.

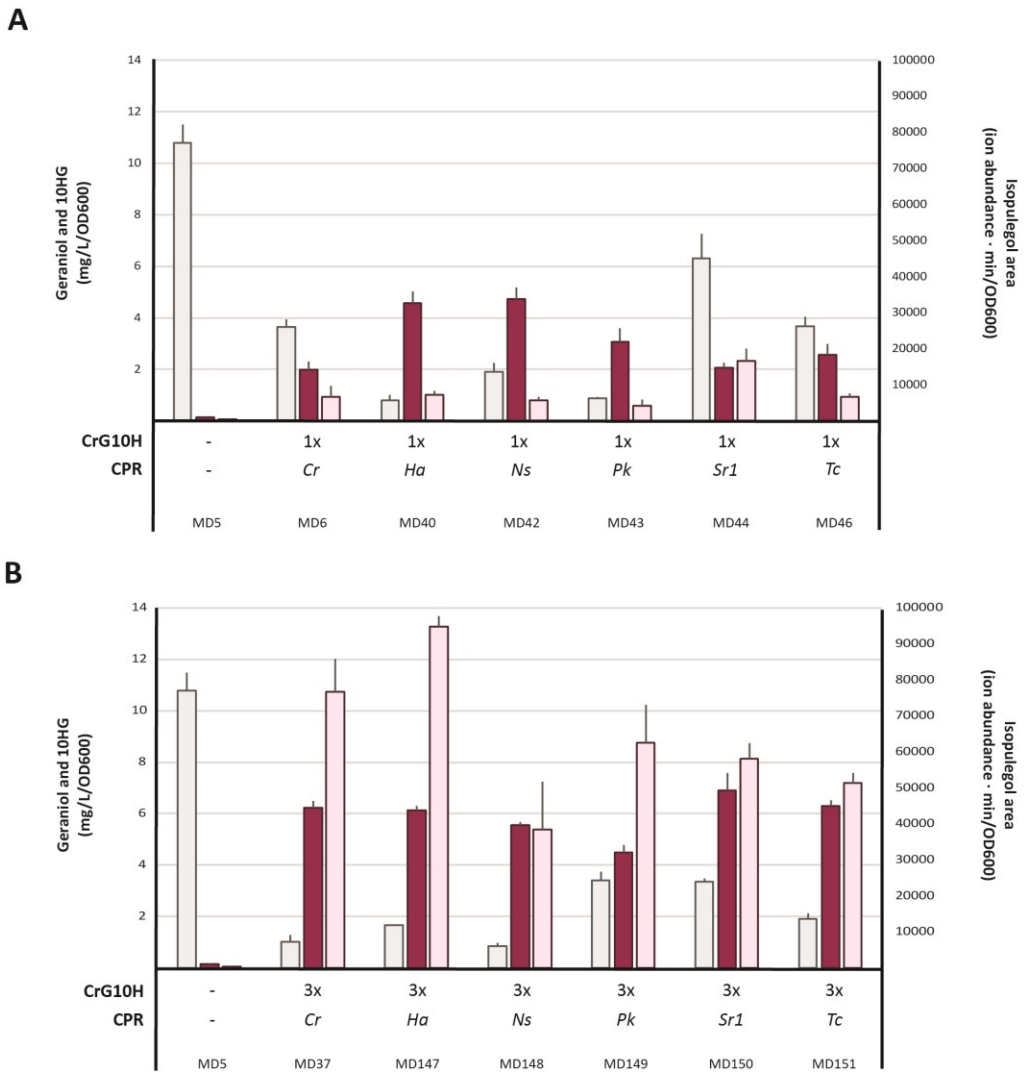

Figure A6. Isopulegol titers from *S. cerevisiae* strains expressing 1x- and 3x- *CrG10H* with different CPRs. Geraniol (grey), 10HG (pink) and isopulegol (light pink) production in **A)** 1x-*CrG10H* and **B)** 3x-*CrG10H* strains. Due to lack of an appropriate standard, isopulegol is reported as area under the curve on the right axis and normalized to OD_{600} of 1. Error bars represent the standard deviation across experimental triplicates.