EFFICIENT STRAIN DEVELOPMENT STRATEGIES FOR PRODUCTION OF TERPENES IN *Saccharomyces cerevisiae*

A Dissertation

by

AVINASH GODARA

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Katy C. Kao
Committee Members,	Qing Sun
	Hung Jen Wu
	Shuhua (Joshua) Yuan
Head of Department,	Arul Jayaraman

May 2021

Major Subject: Chemical Engineering

ABSTRACT

Adaptive laboratory evolution (ALE) is a powerful tool used to increase strain fitness in the presence of environmental stressors. Apart from growth improvement, ALE can also be used for improving productivity if production and fitness is coupled. However, extensive metabolic engineering is generally required to ensure growth-coupling of product formation based on computational predictions; however, such *in silico* predictions typically result in cells with lower than expected growth and/or production. In this work, we develop a generalized ALE strategy to improve production of compounds with antioxidant potential. Terpene are the biggest class of natural products produced by plants with various uses such as flavors, fragrances, colorants, vitamins, commodity chemicals and pharmaceuticals. We used ALE for improving production of two terpenes with antioxidant potential, one produced intracellularly and another extracellularly in yeast. Productivity of terpenes in yeast was improved using oxidative stress over short evolution experiments. Evolved population resulting from evolution experiments were further screened to select for hyperproducers with improved production. Using these hyperproducers, we also aim to gain a deeper understanding of the genotype-to-phenotype correlation of casual mutations responsible for improved production of terpenes in order to develop chassis strains with high productivity of terpenes. Next generation sequencing (NGS) analysis of hyperproducers revealed multiple unique mutations in genes which are not previously known to be related to terpene biosynthesis in yeast.

DEDICATION

I dedicate this work to my family who always pushed me to be better at everything I do.

ACKNOWLEDGEMENTS

First, I would like to thank my advisor Dr. Katy Kao for her guidance, support and understanding throughout my PhD. Second, I would like to thank Dr. Qing Sun and Dr. Jayaraman for their input and support during the research to ensure our research does not stop. Third, I would like to thank my committee members Dr. Hung Jen Wu and Dr. Joshua Yuan for their input on my research.

I would like to thank my colleagues, Jibran, Tianyu, Michelle, Quint for all their help in research and outside lab. I would like to thank Dr. Luis Reyes for the foundation of the work. I would like to thank my friends for making college station feel like home.

Lastly, I want to thank my family for helping me get through tough times and being there for me.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis (or) dissertation committee consisting of Dr. Katy C. Kao, Dr. Hung-Jen Wu and Dr. Qing Sun of Artie McFerrin department of Chemical Engineering and Dr. Joshua Yuan of the Department of Plant and Microbiology.

For Chapter 2 Raman experiments were conducted by Joshua Weatherston from Dr. Hung-Jen Wu's group.

All other work conducted for the thesis (or) dissertation was completed by the student independently.

Funding Sources

This work was also made possible in part by National Science Foundation, Grant/Award Number: CBET-1605347.

NOMENCLATURE

ALE	Adaptive laboratory evolution
NGS	Next-Gen Sequencing
FBA	Flux Balance Analysis
EMA	Elementary Mode Analysis
E.coli	*Escherichia coli*
S. cerevisiae	*Saccharomyces cerevisiae*
GFP	Green Fluorescent Protein
PCR	Polymerase Chain Reaction
DNA	Deoxy-Ribo Nucleaic Acid
RNA	Ribo Nucleic Acid
OD600	Optical Density at 600 nano-meters

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION

Microbial fermentation has been in existence for centuries in the food and beverage industry. Sustainable production of wide variety of products using microbes has been a focus for the past few decades. Products like biofuels (Buijs et al. 2013), antibodies (Robinson et al. 2015; Spadiut et al. 2014), active pharmaceutical ingredients (Waters et al. 2010), nutraceuticals (Wang et al. 2016), and platform chemicals (Zeng and Sabra 2011), can now be produced using microbial hosts using metabolic and pathway engineering. In order to improve production, metabolic engineering leverages pathway optimization and computational algorithms to improve flux toward the product. Engineered strains using these strategies generally exhibit slower growth rate and higher metabolic burden (Glick et al. 1986; Heyland et al. 2011; Hong et al. 1991). Complementary techniques like adaptive laboratory evolution (ALE) can be used to improve slow growth rate but may lead to loss of production (Conrad et al. 2011). Coupling cellular growth or survival with production can reduce the loss of production and has been used by metabolic engineering and computational strain engineering (Shabestary and Hudson 2016; von Kamp and Klamt 2017; Yim et al. 2011).

With increasing availability of annotated genomes, genome-scale metabolic networks have been constructed for various microorganisms (Edwards and Palsson 2000; Forster et al. 2003; Schilling et al. 2000). Genomic-scale models enable the use of

computational tools such as flux balance analysis (FBA) (Varma et al. 1993) to optimize certain cellular function (Segre et al. 2002; Varma and Palsson 1994). FBA-based method OptKnock (Burgard et al. 2003) was the first method to introduce reaction deletions to form growth-coupling of desired product using mixed integer linear programing, and has been successfully used to metabolically engineer strains for production of compounds including 2-3 butanediol (Ng et al. 2012) and lactic acid (Fong et al. 2005). Subsequent algorithms like OptStrain improved upon OptKnock and can predict insertions as well as deletion for increased production (Pharkya et al. 2004). OptCouple is a recently introduced algorithm that can identify insertions, deletions as well as modifications in growth media to facilitate growth-coupled product formation (Jensen et al. 2019). The possibility for growth-coupled production of almost all metabolites for several microorganisms has also been predicted theoretically (Feist et al. 2010; Klamt and Mahadevan 2015; von Kamp and Klamt 2017). However, experimental observation of increased productivity is usually less than observed in silico (Lee et al. 2005; Ng et al. 2012; Yim et al. 2011).

Advances in high throughput sequencing, omic technologies, automation and genetic engineering has enabled an efficient workflow for identifying fundamental mechanisms of microbial adaptation and beneficial mutations for phenotypes-of-interest. Adaptive laboratory evolution has been successfully used over the past few decades to help develop microbes with desired phenotypes. By using an appropriately designed selective pressure, mutants with desired phenotypes are selected for in the population

over time, making ALE a potentially effective but simple strategy for strain development. ALE has been widely used to develop strains with phenotypes-of-interest that are coupled with growth, such as improved substrate utilization (Cadiere et al. 2011; de Kok et al. 2012; Garcia Sanchez et al. 2010; Morais and Bogel-Lukasik 2013) and improved tolerance to inhibitors (Dhar et al. 2011; Fletcher et al. 2017; González-Ramos et al. 2016; Mundhada et al. 2017). However, as product formation is generally not coupled with growth, the use of ALE to improve productivity requires genetically engineered coupling between growth and product formation or specially designed selective pressure that gives fitness advantage to higher producers (Basso et al. 2011; Mahr et al. 2015a; Reyes et al. 2014).

Computational framework for growth-coupling of product synthesis

Genome-scale metabolic models are built from known and predicted cellular metabolic reactions and provide the necessary framework for computational design strategies to couple production of a product-of-interest to growth. By defining an objective function, genome-scale models can be used to predict genetic interventions (*e.g.* gene deletion, insertion, overexpression) that can potentially lead to desired cellular phenotypes such as increased product formation (Burgard et al. 2003; Feist et al. 2010; O'Brien et al. 2015; Ohno et al. 2014). These computational algorithms use stoichiometric metabolic reaction models and can be classified into two groups of methods, Flux Balance Analysis (FBA) and Elementary Modes Analysis (EMA). In

FBA, a desired metabolic phenotype is generally optimized under steady-state mass balance constraints (Savinell and Palsson 1992). Growth coupling designs using FBA-based methods aim to ensure production coupling with growth at distinct metabolic states, which in most cases is maximal biomass formation. OptKnock was one of the earliest computational algorithms developed to identify genetic interventions that lead to simultaneous product formation with biomass formation *in silico* (Burgard et al. 2003). In OptKnock, stoichiometric changes resulting from gene deletion were identified which should theoretically lead to the production of compounds-of-interest at user specified growth rates. Expansions beyond gene deletions were additionally introduced as new algorithms emerged. OptStrain was subsequently developed to improve upon OptKnock and can identify the effect of adding new genes as well as gene knockouts, but cannot predict the combined effects of gene knock-outs and the addition of new genes (Pharkya et al. 2004). SimOptStrain further improved upon that to allow the identification of both gene insertions and deletions simultaneously (Kim et al. 2011). While these methods focus on ensuring optimal biomass and product formations simultaneously, they have not focused on cellular designs that guarantees the dependence of biomass formation on product formation in a way that ensures optimal cellular growth only with high product yield. Such dependence between biomass and product formation is required if ALE is to be used for further strain improvement. The recently developed OptCouple was specifically developed to achieve forced coupling between the production of desired compounds and biomass via simultaneous identification of gene knockouts, insertions, and modifications in growth media (Jensen et al. 2019); the method was able to predict

the same strategy for growth-coupled product methylation that was experimentally validated and used in ALE to improve methyltransferase and acetyltransferase activities in another work (Luo et al. 2019).

In EMA, complex metabolic network is decomposed into a set of reaction sequences (pathways), called elementary modes (EMs), which must satisfy mass balance of metabolic intermediates and cofactors. Any possible flux distribution of the metabolic network can be expressed in the form of linear combinations of EMs (Gagneur and Klamt 2004; Schuster et al. 1999); and the desired set of EMs can be identified as ones that lead to the desired objectives (*e.g.* highest product yield). Methods using EMA have also been developed for a deeper assessment of the metabolic network by identifying the minimal cut sets (MCSs), which is the set of reactions that when deleted will lead to failure of specified cellular functions (Klamt and Gilles 2004). MCSs can be used to identify the minimum set of reactions that need to be deleted to remove the formation of unwanted side products or eliminating cellular ability to use various substrates. Since MCSs focus only on EMs that disable specified cellular functions, which can lead to the disruption of desired functionalities, methods such as constrained minimal cut sets (cMCS) have been developed to preserve EMs that enable desired cellular functions by setting additional constraints to the objectives to ensure both cellular growth and product formation (Hadicke and Klamt 2011; von Kamp and Klamt 2017). EMA-based methods are computationally expensive because of the necessity to establish all EMs, therefore use of these methods are limited to mid- to small-scale metabolic networks. The

introduction of MCSEnumarator, which sequentially enumerate smallest MCSs, helps to reduce the computation costs required (von Kamp and Klamt 2014).

The capability of *in silico* metabolic modeling to predict strategies that allow growth-coupling of product formation has been assessed theoretically (Alter and Ebert 2019; Klamt and Mahadevan 2015). Growth-coupling is classified as weak or strong as shown in **Figure 1**. Products with weak growth-coupling allows for strains with less than optimal yields as long as biomass yields are high whereas strong growth-coupling obligatorily produces high product yields at any biomass yields (even without cellular growth) (Klamt and Mahadevan 2015). Examples of weak growth-coupling products include overflow metabolic products (*e.g.* ethanol in *Saccharomyces cerevisiae* and acetate in *E. coli* (Alter and Ebert 2019)). If a set of deletions exist such that the reduced network shows weak or strong coupling of growth and production, then the formation of the metabolite can made to be growth-coupled. Theoretical analyses predicted that growth-coupling of all metabolites within the central metabolism (Klamt and Mahadevan 2015) and 90% of all metabolites in *E. coli* under aerobic conditions is possible when the flux limits for ATP and cofactors were modified (Alter and Ebert 2019). Recent work by von Kamp and Klamt (von Kamp and Klamt 2017) showed that using cMCS with whole genome-scale models, growth-coupled product synthesis can be achieved *in silico* for almost all metabolites in five different microorganisms including the cyanobacterium *Synechocystis sp*. PCC 6803. Even though a number of computational methods for growth-coupled strain designs exist, due to the large number

of required genetic alterations often predicted *in silico*, the number of *in vivo* applications is still few.

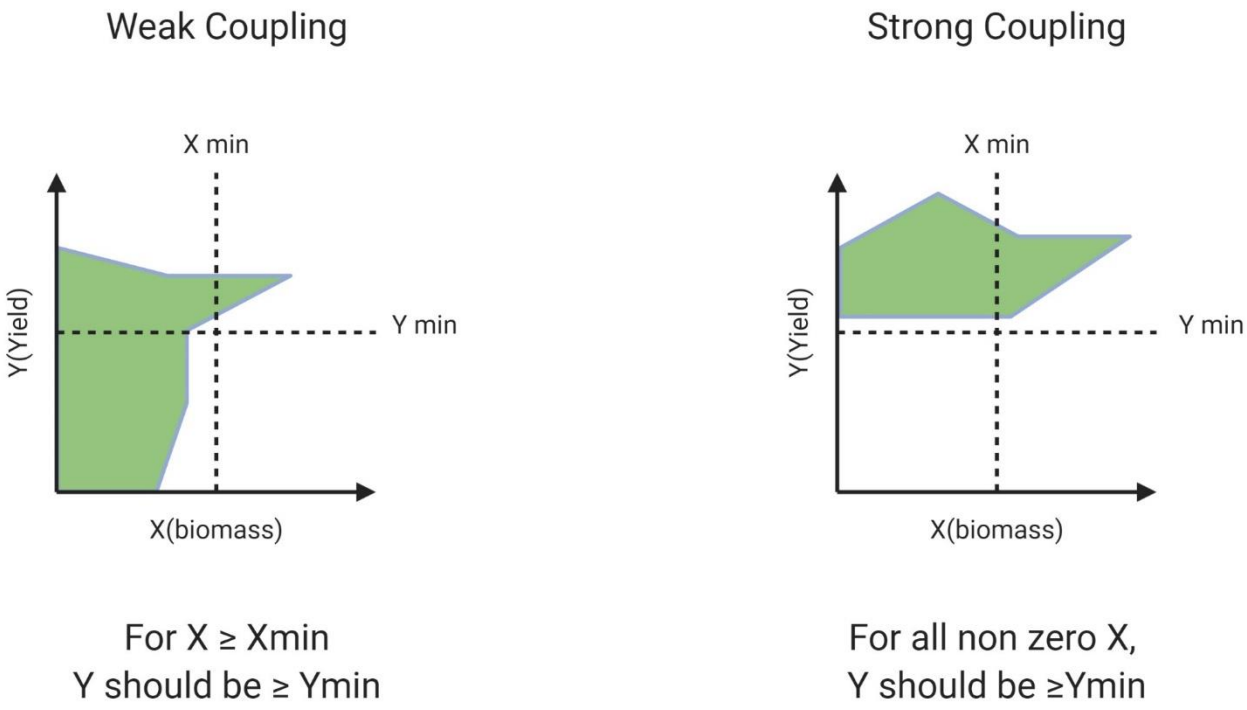


Figure 1 Yield space for weak and strong coupling with their respective constraints as described by (taken from (Klamt and Mahadevan 2015))

Adaptive laboratory evolution

Adaptive laboratory evolution works on the basis of natural selection, whereby a selective pressure is imposed on the evolving population, and mutants better adapted to the selective pressure are enriched in the population. ALE complements metabolic engineering efforts as it does not require *a priori* knowledge of the genotype-phenotype relationship needed in rational strain engineering. Different modes of operation can be

used in ALE, with chemostat and serial batch transfers being the most common (Ekkers et al. 2020; Holwerda et al. 2020; Wright et al. 2011). Microbes are generally evolved for several hundred or thousands of generations under the chosen selective pressure and the best performing strains are isolated for further analyses. With advances in sequencing and omics technology and genetic engineering tools, causal mutations for the desired phenotypes in isolated evolved mutants can be identified and characterized (Dragosits and Mattanovich 2013; LaCroix et al. 2015; Lee and Kim 2020; Phaneuf et al. 2020).

Combining computational approaches with ALE

Adaptive laboratory evolution (ALE) has been used extensively to generate phenotypes such as improving growth rate (Elena and Lenski 2003; Sandberg et al. 2014), improve substrate utilization (Cadiere et al. 2011; de Kok et al. 2012; Garcia Sanchez et al. 2010; Morais and Bogel-Lukasik 2013) and improved tolerance (Dhar et al. 2011; Fletcher et al. 2017; González-Ramos et al. 2016; Mundhada et al. 2017). If product formation can be coupled with growth using metabolic modeling, then the use of ALE should theoretically be able to further improve cell growth and product formation of mutants designed *in silico*. Fong *et al.* was the first to successfully demonstrated the feasibility of this approach; mutant *E. coli* strains designed by OptKnock for growth-coupled production of lactate was subjected to ALE and increase in both growth rate and lactate production were observed in evolved strains (Fong et al. 2005). As

computationally predicted strain designs often lead to mutants with sub-optimal growth and/or productivity *in vivo*, ALE can potentially be used to help the mutant reach or exceed computationally predicted performance. This concept of a combined strategy of computational design followed by ALE was demonstrated by Tokuyama et. al (Tokuyama et al. 2018). Mutants predicted for growth-coupled production of succinate on glycerol as the substrate was designed using an *E. coli* genome scale model; however, the *in vivo* deletion of the four gene targets predicted by FBA (*adhE-pykAF-gldA-pflB*) led to sub-optimal 0.08 C-mol succinate/C-mol glycerol, significantly below the predicted 0.32 C-mol/C-mol (Tokuyama et al. 2018). To improve strain performance, five parallel populations were evolved using adaptive laboratory evolution in glycerol minimal medium for 100 generations. Isolated evolved mutants from all five populations showed significant improved succinate yield between 0.24-0.33 C-mol/C-mol, near the predicted optimal. Subsequent genome sequencing analysis revealed that evolved mutants acquired mutations in phosphoenolpyruvate carboxylase (*ppc*) that reduced feedback inhibition by L-aspartate, leading to overproduction of succinate (Tokuyama et al. 2018).

A potential use of ALE that has not be broadly explored is for improving enzyme activity. Hao *et al* recently devised and constructed a growth-coupled selection design coupling S-adenosylmethionine (SAM)-dependent methylation to growth by rerouting the endogenous cysteine biosynthesis with heterologous pathway converting homocysteine produced by a methyltransferase to produce the cysteine required for

cellular survival, making SAM cycle essential for cellular growth (Luo et al. 2019). This ALE-based strategy was successfully used to improve activity of phenylethanolamine N-methyltransferase (Pnmt), aralkylamine N-acetyltransferase (Aanat), and acetylserotonine O-methyltransferase (Asmt) nearly 2 fold (Luo et al. 2019).

Environmental engineering for growth-coupled production

Adaptive laboratory evolution has been shown to be a powerful strain development tool for growth-coupled phenotypes. If specific products provide protection to cells in a stressful environment, the use of environmental engineering can be used to improve production using ALE as shown in **Figure 2**. Here we briefly describe a few cases where this strategy has been successful in improving product formation. Polyunsaturated fatty acids like docosahexaenoic acid (DHA) is an important nutraceutical that can be produced using *Schizochytrium* sp. Oxygen plays an important role in cell growth as well as lipid production in marine fungi such as *Schizochytrium* sp (Jakobsen et al. 2008). Based on the hypothesis that increased polyunsaturated fatty acid production such as DHA can increase resistance of the cell to oxidative stress in high oxygen conditions, Sun *et al* used increased aeration (high oxygen supply) as a selective stress to increase DHA production in *Schizochytrium* sp using ALE (Sun et al. 2016). While small gains in performance were made during batch cultivation, the evolved strains showed >30% increase in DHA titer (26.4 g L^{-1}) and productivity (220 mg L^{-1} h^{-1}) in fed-batch growth (Sun et al. 2016).

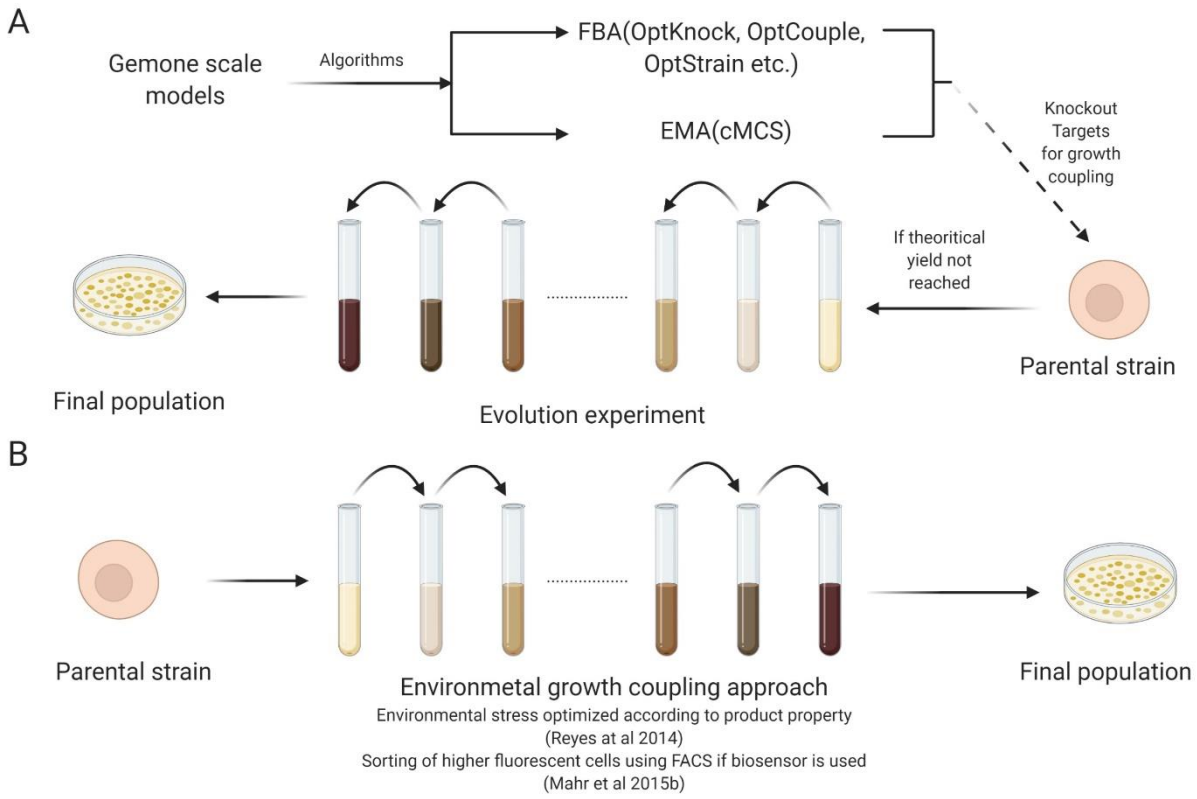


Figure 2 Strategies used for growth coupling production a) flowchart for product improvement using combination of computation algorithms and ALE b) use of environmental engineering to couple growth with production using ALE either by stress or fluorescence intensity. Darker culture represents higher productivity.(taken from (Godara and Kao 2020))

Carotenoid are known antioxidants produced in nature by plants and microbes as well as non-native producers using metabolic engineering (Bhosale and Gadre 2001; Chen et al. 2016; Nishizaki et al. 2007; Yang and Guo 2014). Earlier work on ALE-based improvement of carotenoid production in an engineered yeast strain took advantage of the antioxidant potential of carotenoid to couple cell survival with production (Reyes et al. 2014). Hydrogen peroxide was used as the selective pressure to

evolve a catalase deletion (Δ*ctt1*) strain for increased carotenoids production. Results showed that evolved mutants had up to 3-fold improvement in product yield (Reyes et al. 2014).

Computational strain design for growth-coupled product formation may not be feasible or practical for all metabolites. Environmental engineering approach to growth-coupling of desired compounds is also only feasible for specific classes of compounds. Thus, new innovations in synthetic circuits using biosensors can be leveraged to couple growth with production, which can then be combined with ALE to improve productivity of metabolites. These approaches generally use a transcription factor-based biosensor that detects the metabolite-of-interest in a concentration-dependent manner using a fluorescent readout, such that the higher producers generate a higher fluorescence signal. *Corynebacterium glutamicum* is extensively used in industrial applications for amino acid production. Biosensors have been developed based on the transcription factor Lrp of *Corynebacterium glutamicum* to detect L-methionine as well as branched chain amino acids L-leucine, L-isoleucine and L-valine, with eYFP fluorescence as output (Lange et al. 2012; Mustafi et al. 2012). With Lrp, Mahr *et al* used a biosensor-driven evolution approach to improve L-valine production by sorting cells showing top 10% of FACS sensor output during iterative fluorescent activated cell sorting; using this artificial evolution, a 25% increase was observed in L-valine production in *C. glutamicum* after five evolution steps (Mahr et al. 2015b). Using an anti-metabolite for phenylalanine (4-fluorophenylalanine) combined with a transcription factor-based sensor (Aro9) for aromatic amino acids (AAA), ALE was used to increase muconic acid production in in

S. cerevisiae (Leavitt et al. 2017). Since muconic acid biosynthesis branches off from the AAA biosynthesis pathway, AAA production was used as a surrogate to improve muconic acid production. A hybrid aromatic amino acid (AAA) biosensor was used to drive the expression of geneticin resistance gene to allow growth-coupled selection for aromatic amino acid biosynthesis; this growth-coupling was strengthened with the addition of anti-metabolite 4-fluorophenylalanine, whose toxicity can be overcome by increased production of phenylalanine (Leavitt et al. 2017). Using this strategy, a titer of 2.1 g L^{-1} of muconic acid was eventually achieved in bioreactors (Leavitt et al. 2017).

Conclusion

Development of microbial producers has been enabled by the development in silico techniques to couple product formation with cellular growth. Advances are being made in reducing the computational costs to run these models. Newer algorithms are incorporating not only gene knockout targets and insertions but modifications in the growth environment to achieve coupling of growth with production. However, these computational models are not without limitations. Quality of predictions depend on quality of genome-scale models. The predictions often involve multiple genetic disruptions and can be difficult to implement in vivo without detrimental effects to the cell. Complementing these techniques with ALE can help to overcome impairments to the cell to allow the microbes to reach the computationally predicted theoretical maximum.

Experimental evolution experiments are relatively easy to setup, but selecting effective selective pressures is one of the most difficult steps in achieving the desired increase in biosynthesis of product-of-interest. A few studies have explored the chemical properties of desired compounds to design an environment in which product formation leads to higher cellular growth or survival. However, not all products have properties that can be exploited in this manner, making the selective pressure design strategy difficult and limited. This limitation in ALE can potentially be alleviated by leveraging advances in computational predictions coupling production of desired compounds with growth. In addition, use of biosensors to artificially couple product formation with a reporter (fluorescence or cell growth) is a new and innovative strategy that does not require extensive genetic modifications to couple growth with product formation. However, biosensors for compound-of-interest may not be readily available and are difficult to develop. An iterative approach using a combination of computational predictions with ALE (with or without the use of biosensors) provide a framework for developing an effective strategy to improve the performance of the production host.

Aims of the study

To understand how adaptive laboratory evolution can impact the production in microbial systems we used terpenes as the model product and used *Saccharomyces cerevisiae* as our target microbe. We leveraged the antioxidant potential of the few terpenes to create an experimental strategy to growth couple the production. In earlier

work in our lab the production of a terpene β-carotene was improved using ALE and few hyperproducers were selected from the evolved population. For our first aim we targeted to understand the mechanism behind this improved production of carotenoid in yeast, and gain deeper understanding of genotype-phenotype changes occurred during evolution because of mutations accumulated by the cells. For our second aim we targeted another terpene which was extracellular unlike carotenoids which are intracellular. An optimized selective pressure scheme was constructed for production improvement using ALE. Hyperproducer mutant strains were screened out and sequenced to capture the mutations. Subsequently these mutations were reconstructed to understand the genotype-phenotype relation for another class of terpene.

CHAPTER II

GENOTYPE-PHENOTYPE RELATIONSHIP OF CAUSAL MUTATION IN EVOLUVED MUTANT OF CAROTENOID PRODUCING *Saccharomyces Cerevisiae*

Background

Microbial systems are being used over chemical synthesis in various industries to develop a plethora of products because of their advantage of sustainability and ease of engineering to produce variety of chemicals. As with any process, microbial production is also filled with challenges of optimization, rate yield and titer of the product. With advances in next gen sequencing and well characterized pathways of many microbial systems there are more available tools to address the challenges such as rate and yield of the production with strain engineering.

Terpene are the biggest class of natural products produced by plants with various uses as flavors, fragrances, colorants, vitamins, commodity chemicals and pharmaceuticals. However, low abundance of these chemicals in natural sources makes it very hard to extract them from plants for industrial application. Chemical synthesis and microbial production are two alternates for production of these compounds other than natural extraction. With increase in structural complexity of the terpene molecule the chemical synthesis becomes increasing cost inefficient and in such cases production of terpenes can be done with using biosynthesis using microbes. All terpenes or isoprenoids are secondary metabolites that are derived from C_5 precursor isopentenyl diphosphate (IPP) or dimethylallyl diphosphate (DMAPP) (Arigoni et al. 1997;

Asadollahi et al. 2008). Two distinct pathways have been reported for the production of terpenes, mevalonate pathway and 2C-methyl-D-erythritol-4-phosphate (MEP) pathway. Mevalonate pathway is present in eukaryotes (fungi, archaea, plants) while non-mevalonate pathway is present in bacteria, algae and chloroplast of plants.

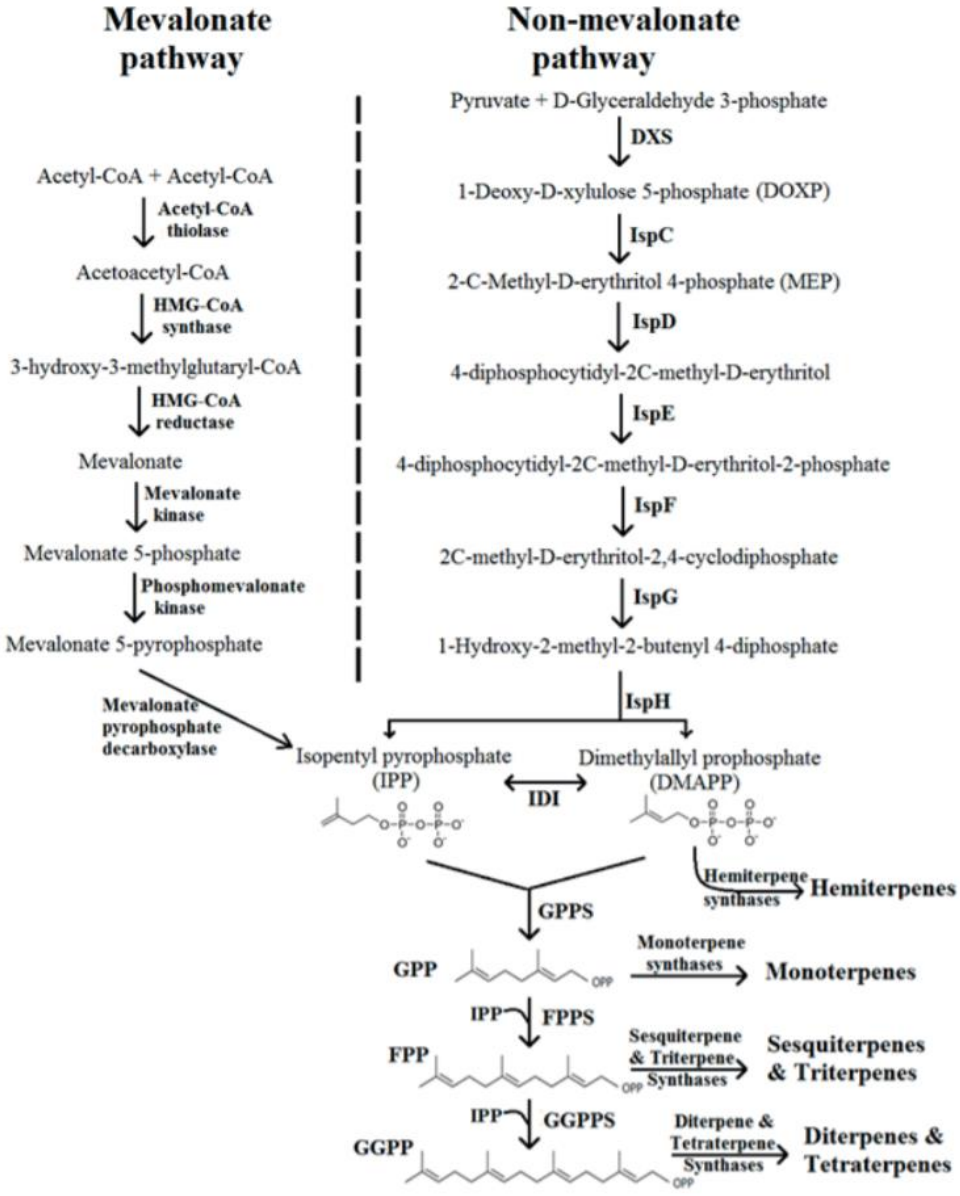


Figure 3 Biosynthesis pathway for production of isoprenoids (taken from (Ingy and Wim 2017))

However, several studies have shown that genetic determinants outside the known biosynthetic pathways can also play a role in production(Alper et al. 2005; de Ruijter et al. 2017). Due to the complexity of biological systems, these genetic determinant outside the known pathway are usually unknown. Thus, complementary genome-scale techniques, such as screening deletion or overexpression libraries (Arlt et al. 2011; Prelich 2012; Stevenson et al. 2001) and adaptive laboratory evolution (ALE)-based strategies (Portnoy et al. 2011; Reyes et al. 2014) can help to identify additional gene and/or pathway targets related with increasing product formation. Advances in next gen sequencing, omic technologies, automation and genetic engineering has enabled a deeper understanding of evolutionary processes and trajectories. Adaptive laboratory evolution has been successfully used over the past few decades to develop strains with desired industrially relevant phenotypes. Growth over numerous generations under an appropriately designed selective pressure to direct the accumulation of mutations responsible for desired phenotype has been the simple but effective strategy for ALE. However, due to the required coupling of the phenotype-of-interest with growth, there are numerous examples of ALE being used to improve substrate utilization(Cadiere et al. 2011; de Kok et al. 2012; Garcia Sanchez et al. 2010; Morais and Bogel-Lukasik 2013) and improved tolerance(Dhar et al. 2011; Fletcher et al. 2017; González-Ramos et al. 2016; Mundhada et al. 2017), but few examples for increased productivity(Basso et al. 2011; Mahr et al. 2015a; Reyes et al. 2014).

Carotenoids belong to an important class of isoprenoids naturally synthesized by plants and some microbes (An et al. 2001; Bhosale and Bernstein 2004; Chen et al. 2006; Jeon et al. 2006; Nanou et al. 2007; P. and R.V. 2001; Raja et al. 2007). Due to their pigmented and antioxidant properties, carotenoids have applications in food and nutraceutical industries. In addition to native microbial producers such as *Rhodotrula glutinis* (Bhosale and Gadre 2001; Kot et al. 2016; Malisorn and Suntornsuk 2008), non-native producers such as *Escherichia coli* and *Saccharomyces cerevisiae* have been engineered to produce these carotenoids via metabolic engineering (Chen et al. 2016; Mata-Gómez et al. 2014; Nishizaki et al. 2007; Verwaal et al. 2007; Wu et al. 2017; Yang and Guo 2014).

In our prior work, we leveraged the antioxidant properties of carotenoids to develop an ALE-based strategy to evolve strains for increased productivity. We evolved an engineered *S. cerevisiae* strain (YLH2) for improved β-carotene production using hydrogen peroxide as the selective pressure (Reyes et al. 2014). Adaptive mutants with enhanced carotenoids yields were isolated from two parallel populations at various time points throughout the evolution (Reyes et al. 2014).

Materials and methods

Strains, plasmids and growth conditions

All yeast strains used in this work are derivatives of S288c and listed in **Table 7**. *S. cerevisiae* strain YLH2 (Reyes et al. 2014) and FY2 (MATα, ura3-52, isogenic to S288C) (Fred et al. 1995) are used as the base strains in this study. Unless otherwise specified, yeast strains are cultured in Yeast Peptone Dextrose (YPD) media at 30 ºC and

E. coli strains used for subcloning were cultured at 37 ºC in Luria Broth (LB) supplemented with appropriate antibiotics. Cytosolic catalase T (*CTT1*) gene was deleted in FY2 strain as described in our earlier work (Reyes et al. 2014). For carotenoid quantification, yeast strains were cultured in YPD at 30 °C and 200 rpm for 72 h. For Raman Spectroscopy the strains were grown for 36 hours in YPD at 30 °C and 200 rpm.

Whole genome resequencing

Genomic DNA of isolated mutants and parental strain YLH2 were extracted using YeaStar DNA extraction kit (Zymo Research). Library preparations and NGS sequencing were performed by the Texas A&M Genomics Center for sequencing on the Illumina HiSeq 2500 platform using 100-bp single-end reads. An average coverage of >75-fold was obtained for each clone. The raw sequencing data were deposited in SRA database (https://www.ncbi.nlm.nih.gov/sra) with accession number SRP147744. The sequencing data was aligned to *S. cerevisiae* S288c reference genome with breseq v0.29 (Deatherage and Barrick 2014). *De novo* mutations in isolated mutants were identified by comparing against the YLH2 parental strain, and verified via Sanger sequencing (mutations listed in **Table 1**).

Reconstructing mutations into YLH2 and FY2 strain

CRISPR-Cas9 using the one plasmid-system with pCRCT developed by (Bao et al. 2015) was used for site-directed mutagenesis to reconstruct identified mutations into YLH2. Briefly, 120 bp sequence including donor sequence and guide sequence was

chemically synthesized by Integrated DNA Technologies, USA (**Table 8**), then introduced into the pCRCT plasmid using Golden Gate reaction using BsaI restriction sites (Engler et al. 2008), and transformed into *E. coli* cells, and plated on LB+Xgal for blue/white screening and incubated overnight at 37 °C. The bacterial colony with correct plasmid construction was verified by restriction digestion. The constructed CRISPR-Cas9 plasmid was then transformed into FY2 and YLH2 strains and selected on SC - uracil plates. Colonies were picked and target mutations were first verified using PCR amplification refractory mutation system (ARMS) **(**see **Table 9** for primer sequences) as described by (Little 2001) followed by confirmation via Sanger sequencing. Cas9 plasmid was cured by serially passaging the strain in YPD 3 times and verified by PCR amplification using primers curing_f and curing_r (**Table 9**).

Quantification of carotenoid production

Quantification of β-carotene production was performed as described by (Reyes et al. 2014). All samples were blank corrected using a dodecane blank. The same procedure was used to quantify lycopene using OD_{476}. Standard curves were generated using compounds purchased from Sigma. At least 3 biological replicates per strain were used for analysis.

Oxidative stress tolerance using hydrogen peroxide

Reconstructed mutant in FY2 were treated with 1 M hydrogen peroxide for 1 h in a shaking incubator at 30ºC. After the shock treatment, 10-fold serial dilutions were

performed in phosphate-buffered saline (PBS) and spotted on YPD plates. Three biological replicates per strain were used for analysis.

Quantifying relative frequency of *HIS7* mutation in population samples

The kinetic quantitative-PCR method described by (Germer et al. 2000) was used to estimate the relative frequency of the *HIS7 389* mutation from evolving population samples revived from frozen stocks using mutation-specific and wild-type-specific primers (his7_f_mut, his7_f_wt, and his7_r in **Table 9**).

Growth kinetics measurements

Microplate reader (TECAN Infinite ® M Nano) was used to measure growth curve for reconstructed mutant strains. Cells were grown for 24 h in YPD and normalized to OD_{600} ~0.05 in 200 µL final volume in YPD. Cells were grown in 96-well plates for 48 h with orbital shaking at constant intervals with 245 rpm. To obtain specific growth rate, the OD_{600} observed was plotted on natural logarithmic scale versus time (h). The slope of linear portions (specific growth rate µ) for each sample was calculated for each replicate. Three biological replicates per strain were used for analysis.

Quantification of carotenoid production

Single colony was isolated from a plate streaked from frozen stock and used to inoculate 5 mL of culture in YPD. This was grown overnight at 30 ºC and 200 rpm, then subcultured into 5 mL of fresh YPD at an initial OD_{600} ~0.05 and cultured at 30 °C and

200 rpm for 72 h. After ~72 h, 200 μL of culture was centrifuged and the supernatant was vacuum aspirated. 200 μL of acid washed beads 425-600 μm (Sigma) and 1 mL of dodecane (TCI America) were added to the cell pellet for disruption using the Genie Cell Disruptor (Scientific Industries) for 2 cycles of 6 minutes to ensure complete disruption. The mixture was then centrifuged and 200 μL of the supernatant was transferred to 96 well plate for absorbance measurements at OD_{454}, which corresponds to β-carotene peak (Verwaal et al. 2007), using a microplate reader (TECAN Infinite ® M Nano).

Validation of spectroscopy data using high performance liquid chromatography

The accuracy of using the spectrophotometric method to quantify β-carotene concentration was verified using high performance liquid chromatography (HPLC, Agilent Technologies, 1260 Infinity System) and C18 column ZORBAX Eclipse Plus (4.6 x 100 mm, 3.5 micron) using a method previously described (Xie et al. 2015).

Raman spectroscopy analysis

To prepare samples for Raman spectroscopy, strains were first grown in YPD from single colonies isolated from plates struck from frozen stock for 36 hours at 30 ºC and the yield of β-carotene or lycopene was calculated as described earlier. Three biological replicates per strain were used for analysis. One mL of the grown culture was then centrifuged and the supernatant was vacuum aspirated. The cell pellet was washed 3 times with 1 mL of PBS. 2 μL of each sample was pipetted and gently spread onto an aluminum-covered microscope slide, yielding a thin film of intact cells. The thin film

was allowed to dry fully before Raman spectral acquisition. 16 spectra were collected for each biological replicate (48 spectra per strain) using a DXR Raman microscope (Thermo Scientific) equipped with a high-brightness 780 nm diode laser, a 50x objective lens (N.A. 0.5, spot size 1.6 μm), a 25 μm confocal slit, a diffraction grating (400 line/mm, spectral resolution ~2.4-4.4 cm^{-1}), and a CCD detector. Laser power was set at 20 mW and spectra were collected over the spectral range 150 – 3200 cm^{-1}, with 10 accumulations of 12 s each (120 s total laser exposure per spectrum). Each spectrum was processed post-collection for artifact removal, baseline correction and noise reduction. After processing, the spectral intensity was tabulated (**Table 10**) for several characteristic peaks corresponding to the carotenoids, lipids, and proteins within the cells. Absolute peak intensities varied widely within samples due to inhomogeneity of the cell films. Therefore, ratios of the characteristic peaks were used to compare the relative abundance of chemical compounds across strains and mutations. The ratio of total fatty acid peak and carotenoid peak was multiplied by the total carotenoid content, which is obtained via quantification, to get a relative lipid indicator.

$$Lipid\ indicator = \frac{Total\ fatty\ acid}{Carotenoid} * absolute\ carotenoid\ content(\frac{mg}{gdcw})$$

Quantification of lipid using classical Bligh and Dyer method

Yeast cells were streaked from frozen stock and an individual colony was grown in 5 ml YPD overnight at 30 °C and 250 rpm. The resulting culture was normalized to initial OD_{600} ~0.05 in 100 ml YPD for further growth for an additional 36 h. Cells were harvested via centrifugation and the cell pellet was dried for 48 h at room temperature. 1

gram of cell pellet was mixed with 1:2:0.8 chloroform:methanol:water (v/v/v) and homogenized for 10 min using Branson Digital Sonifier® 250 cell disruptor with 40% amplitude and pulse on and off every 5 sec. Water containing 0.8% KCl and chloroform was added to bring the solvent ratio to 2:2:1.8 chloroform:methanol:water (v/v/v) and homogenized for another 10 min. The two liquid phases were separated for 2 hours at room temperature, and the total volume of organic layer (bottom phase) was measured. 1 ml of organic layer was spread on pre-weighed aluminum pans, and the sample was left to evaporate in chemical hood overnight. Weight was recorded after overnight evaporation using 3 technical replicates for each of 3 biological replicates per strain (Bligh and Dyer 1959; Breil et al. 2017).

Oxidative stress tolerance using hydrogen peroxide

Individual colonies of each reconstructed mutant in FY2 were cultured in 5 mL YPD for 72 h at 200 rpm. Cultures were then normalized to OD_{600} ~1.0, and 500 µL of each normalized sample was treated with 1 M hydrogen peroxide for 1 h in a shaking incubator at 30 ˚C. After the shock treatment, 10-fold serial dilutions were performed in phosphate-buffered saline (PBS). Five µL of each dilution was then spotted on YPD plates and incubated for 48 h at 30 ˚C. A separate normalized culture not treated with H_2O_2 was used as control to ensure consistent normalization and dilution. Strains with *CTT1* deletion were treated with a lower concentration of hydrogen peroxide (100 mM) for 30 mins in shaking incubator at 30 ˚C. Three biological replicates per strain were used for analysis.

Quantitative RT-PCR

The total RNA of each strain was extracted using Quick-RNA™ Fungal/Bacterial RNA Miniprep kit (Zymo Research) from mid-log phase culture grown in YPD at 30 ºC and 200 rpm. The quantitative reverse-transcriptase real-time PCR (qRT-PCR) was carried out with 50 ng of total RNA and 0.5 µM of each primer using qScript™ One-Step SYBR® Green qRT-PCR kit (Quanta Biosciences™) on the LightCycler® 96 (Roche). *ACT1* gene was used as housekeeping control in all experiments. All primers used are listed in **Table 9**. The qRT-PCR reaction protocol is as follows: 50 ˚C for 10 min for cDNA synthesis and 95 ˚C for 5 min for Taq activation, followed by 45 cycles of 95 ˚C for 10 sec, 60 ˚C for 20 sec and 72 ˚C for 30 sec. The melt curves were continuously acquired from 65 ˚C to 97 ˚C at 5 readings/˚C. ΔCt is obtained by the difference of expression between gene-of-interest and housekeeping gene (*ACT1*). ΔΔCt was obtained by the difference between delta ΔCt value of mutant strain and wild type strain.

Bioreactor study

The seed culture for fermentation was cultured using a single colony inoculated in 5 mL YPD overnightand then subcultured into 50 mL YPD media. Seed culture was grown for 24 h at 30 ºC and used to inoculate 1 L YPD media in a 3 L bioreactor (Applikon®) at 30 ºC with continuous air feed at 2.6 L/min and online dissolved oxygen (DO) measurement at 800 rpm. 2 M HCl and 2 M NaOH were used for pH adjustment set at pH 4.0. Batch bioreactor experiments were run for ~64 hours.

Lycopene cassette

The heterologous genes *CrtE* from *Blakeslea trispora, CrtB* from *Pantoea agglomerans*, and *CrtI* from *Blakeslea trispora* for lycopene biosynthesis (Chen et al. 2016) were codon optimized for *S. cerevisiae* and synthesized as gBlocks (Integrated DNA Technologies). The multigene plasmids were constructed using the MoClo-YTK plasmid kit (Lee et al. 2015) (Addgene).

Each chemically synthesized gBlock® was introduced into the entry vector YTK001 using BsmBI (Thermo Scientific) via Golden Gate assembly protocol described by (Lee et al. 2015). Each gene (in an entry vector) is then assembled into transcriptional unit plasmid along with the appropriate promoter and terminator part vectors using BsaI (Thermo Scientific) via Golden Gate assembly. Finally, different transcriptional units with each gene expressed under different promoters were joined together and introduced into pre-assembled URA3 integration vector available in the kit YTK096 using BsmBI. *E. coli* DH5α was used for cloning. Chemically competent *E. coli* cells were prepared using Zymo Mix & Go! E. coli Transformation Kit and Buffer Set™ (Zymo Research) and plated on LB agar supplemented with appropriate antibiotics. The lycopene cassette was transformed into strain FY2 and plated on SC-uracil plates for integration into the chromosome at the URA3 locus. Red colonies were selected after 3 days of incubation at 30 °C.

Results and Discussion

Genome resequencing revealed mutations in genes not known to be related with isoprenoids biosynthesis in hyperproducers

To determine the underlying adaptive mutations responsible for improved production, isolates SM12, SM13, and SM14 from one population and the ancestral strain YLH2 were subjected to whole genome resequencing in this work. Four to six *de novo* mutations were identified from each isolated mutant (see **Table 1**). Several identical mutations (*e.g. HIS7 389, SRO9/GFD2* int) were found in more than one isolated mutant, which was not surprising since these mutants originate from the same evolving population.

Table 1 List of identified mutations (taken from (Godara et al. 2019))

Strain	Chr	Location	Type	Gene	Mutation	Amino acid change
SM12	2	716,077	G->A	HIS7	HIS7 389	S130L
	3	58,825	T->A	SRO9/GFD2	SRO9/GFD2 int	intergenic
	15	977,984	T->G	TYE7	TYE7 86	D29A
	mito	954	-24Δbp	tP(UGG)Q/ST15SrRNA	-24Δbp	deletion in noncoding sequence

Table 1 Continued

Strain	Chr	Location	Type	Gene	Mutation	Amino acid change
SM13	1	204,327	G->A	FLO1	FLO1 925	V309I
	2	716,077	G->A	HIS7	HIS7 389	S130L
	3	58,825	T->A	SRO9/GFD2	SRO9/GFD2 int	intergenic
	6	22,874	A->G	AQY3/DAK2	DAK2/AQY3 int	intergenic
	7	354,893	C->T	SCY1	SCY1 1836	I612I
	mito	954	-24Δbp	tP(UGG)Q/ST15SrRNA	-24Δbp	deletion in noncoding sequence
SM14	6	88592	T>G	EPL1	EPL1 1754	L585W
	7	124954^124955	insA	MDS3	MDS3 insA	nonsense mutation
	13	361550	G>A	YMR045C & YMR046C	YMRCTy1-3 1078	H360Y
	15	330827	A>C	ALG6	ALG6 1411	S471R
	2	716077	G>A	HIS7	HIS7 389	S130L

Among the *de novo* mutations that are identified in more than one isolate, a 24 bp mitochondrial deletion was found 152bp downstream of the mitochondrial proline tRNA

gene *tP(UGG)Q* in mutants SM12 and SM13. Since this is a mutation in the mitochondria and difficult to analyze, this mutation was omitted from further characterization in this study. A missense mutation in *HIS7 389* was identified in all sequenced mutants, indicating that this mutation arose early on during the evolution experiment. Using kinetic quantitative-PCR, the relative frequency of this mutation was estimated in population samples (revived from frozen stocks) from the evolution experiment. The results showed that the *HIS7 389* mutation was present in ~99% of the population very early on during the evolution, after the first cycle of H_2O_2 challenge (Reyes et al. 2014), strongly suggesting this to be a jackpot mutation (Luria and Delbruck 1943), which is a mutation that arose very early on in the experiment. *HIS7* encodes the imidazoleglycerol phosphate synthase gene, which is involved in histidine biosynthesis, and does not have an obvious role in β-carotene production. Indeed, subsequent verification experiments revealed that the mutation in *HIS7 389* confers no benefit to general fitness nor oxidative stress tolerance (later sections). A common intergenic mutation between *SRO9/GFD2* was identified in 2/3 sequenced mutants. Sro9 is a La-motif-containing protein and is known to be involved in translation (Sobel and Wolin 1999), and Gfd2 is currently a protein of unknown function. Mutations in the intergenic region of these two genes may potentially impact the expression of one or both genes; and in the case of Sro9, potentially lead to global changes in gene expression if impacted.

A common intergenic mutation between *SRO9/GFD2* was identified in 2/3 sequenced mutants. Sro9 is a La-motif-containing protein and is known to be involved

in translation (Sobel and Wolin 1999), and Gfd2 is currently a protein of unknown function. Mutations in the intergenic region of these two genes may potentially impact the expression of one or both genes; and in the case of Sro9, potentially lead to global changes in gene expression if impacted.

A total of 8 unique mutations were identified among the sequenced strains: 5 non-conservative missense mutations (*TYE7 86* in SM12, *FLO1 925* in SM13, and *EPL1 1754*, *YMRCTy1-3 1078*, and *ALG6 1411* in SM14), 1 intergenic mutation (between *DAK2/AQY3* in SM13), 1 frame shift mutation (in *MDS3* in SM14) and 1 synonymous point mutation (in *SCY1* in SM13). The intergenic mutation in *DAK2/AQY3* may play a role in stress adaptation as Dak2 is involved in stress adaptation and is required for detoxification of dihydroxyacetone (Molin et al. 2003), and Aqy3 is reported to play a role in passive glycerol diffusion (Oliveira et al. 2003). The mutation in *TYE7* can potentially lead to changes in gene expression as it is a regulator and acts as a transcriptional activator in Ty1-mediated gene expression (Lohning and Ciriacy 1994). Epl1 is involved in epigenetic regulation, and is known to control the expression of genes involved in DNA repair, ubiquitylation and other stress responses (Dohmen et al. 2007; Goossens et al. 2015; Stankunas et al. 1998), thus a mutation in *EPL1* can potentially impact general adaptation. Flo1 promotes adhesion of yeast cells and formation of cell clumps, which can protect cells from environmental stressors (Smukalla et al. 2008); however, based on microscopic inspection of SM13 (which contains the *FLO1* mutation), no obvious clumping of the cells was observed (data not shown). The mutation in *SCY1* is a synonymous point mutation, while we do not expect

the function of the gene to be altered, synonymous mutations have been shown to impact gene expression (Bailey et al. 2014; Kudla et al. 2009). Scy1 plays a role in sterol transport (Sullivan et al. 2009), thus a synonymous mutation in this gene can potentially alter lipid content in the cell. Alg6 encodes an endoplasmic glucosyltransferase (Reiss et al. 1996) and is involved in cell wall assembly (Arias et al. 2011). The mutation in *MDS3* identified in SM14 is a frameshift mutation near the N-terminus of the protein, strongly suggesting this to be an inactivating mutation. Mds3 regulates sporulation in *S. cerevisiae (Benni and Neigeborn 1997)* and deletion in *MDS3* has been shown to negatively impact resistance to oxidants (Higgins et al. 2002); thus the reason this mutation was selected for during oxidative stress challenge needs further investigation. The *YMRCTy1-3* encodes components of the Ty1 retrotransposon; Ty1 insertion mutations are known to impact fitness and are observed in laboratory evolution experiments (Gresham et al. 2008; Wilke and Adams 1992); thus, mutation in *YMRCTy1-3* can potentially impact fitness.

Assessing the impact of identified mutations on general fitness and β-carotene yield in reconstructed mutants in the carotenogenic ancestral strain YLH2

Mutations that may be selected for in our prior ALE experiment include those that lead to increased H_2O_2 resistance (*e.g.* via increased production of the antioxidant β-carotene) or confer benefits on general fitness. To assess impacts of each mutation, all mutations (except for the mitochondrial mutation) identified in SM12, SM13 and SM14

were introduced into YLH2 strain individually to generate strains YAG01 to YAG10. The growth kinetics results showed no significant differences in the specific growth rates (μ) of most of the reconstructed mutants compared with the parental strain YLH2 (See **Table 2**). However, all of the mutations led to significant reductions in the duration of the growth lag from ~4 h in parental strain YLH2 to between ~0.4 h – 2 h in the reconstructed mutants, which likely contributed to their selection in the ALE experiment.

Table 2 Growth kinetics of reconstructed mutants. Two-tailed student t-test was used to calculate p-values comparing with YLH2. Bold: p-value < 0.001 compared with YLH2 using two-tailed Student t-test (taken from (Godara et al. 2019))

Strain	Mutation	Specific growth rate (per hour)	Lag phase (hours)	OD after 24hr (Avg±Stdev)
YLH2		0.28±0.02	4.29±0.23	0.98±0.05
YAG01	HIS7 389 (a)	**0.23±0.01**	**0.76 ± 0.17**	0.82±0.06
YAG02	SRO9/GFD2 int (b)	0.28±0.00	**0.96 ± 0.12**	0.94±0.02
YAG03	TYE7 86 (c)	0.26±0.02	**0.66 ± 0.23**	0.91±0.03
YAG04	FLO1 925 (d)	**0.22±0.01**	**0.45 ± 0.17**	0.90±0.05
YAG05	DAK2/AQY3 int (e)	**0.23±0.01**	**0.91 ± 0.19**	1.01±0.09
YAG06	SCY1 1836 (f)	0.27±0.01	**0.45 ± 0.17**	0.97±0.03
YAG07	EPL1 1754 (g)	**0.20±0.00**	**1.87 ± 0.12**	0.94±0.04
YAG08	ALG6 1411 (h)	0.24±0.02	**1.16 ± 0.23**	1.00±0.05
YAG09	MDS3 ins (i)	0.27±0.01	**0.35 ± 0.12**	0.96±0.05
YAG10	YMRCTy1-3 1078 (j)	0.27±0.01	**0.40 ± 0.16**	1.11±0.12

The impacts of each mutation on carotenoids production was evaluated. Overall, the results showed that the majority of mutations alone conferred a positive impact on carotenoids production (see **Figure 4A**). The intergenic mutation between *SRO9/GFD2* (YAG02) led to a significant increase in β-carotene content, with yield of 17.9±0.1 mg/g

DCW in YAG02 compared with 9.8± 0.3 mg/g DCW in YLH2 and both evolved mutants that contain this mutation (SM12 [15.1± 0.2 mg/g DCW] SM13 [14.2± 0.1 mg/g DCW]). In SM14, the single mutation in *YMRCTy1-3* alone resulted in β-carotene yield of 20.5±0.2 mg/g DCW (YAG10), slightly exceeding the 19±0.1 mg/g DCW observed in strain SM14 (p-value ~ 0.003).

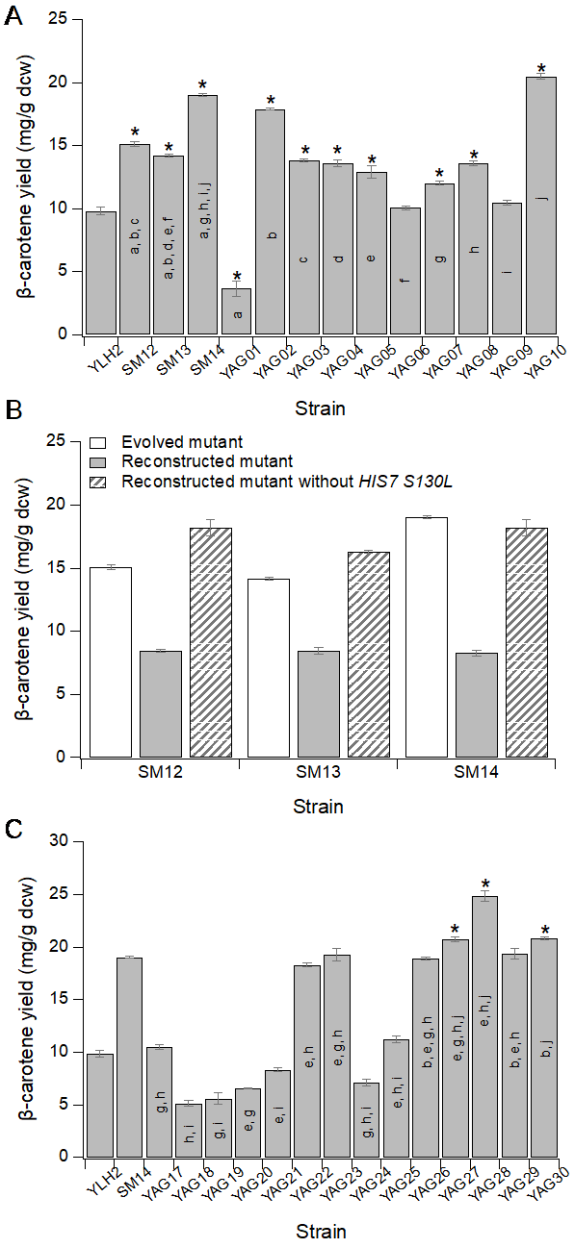


Figure 4 β-carotene production of evolved and reconstructed mutants. A) Single mutations; asterisks: p-value < 0.05 using 2 tailed Student t-test against YLH2. B) Reconstructed evolved mutants. C) Multiple mutations; asterisks: higher production compared with SM14 with p-value < 0.05 using 2 tailed Student t-test. Mutations: (a) *HIS7 389*, (b) *SRO9/GFD2 int*, (c) *TYE7 86*, (d) *FLO1 925*, (e) *DAK2/AQY3 int*, (f) *SCY1 1836*, (g) *EPL1 1754*, (h) *ALG6 1411*, (i) *MDS3 ins*, (j) *YMRCTy1-3 1078* (taken from (Godara et al. 2019))

These results suggest that one or more of the other mutations present in SM12, SM13 and SM14 were detrimental to product yield. Indeed, the jackpot mutation in *HIS7 389* that is present in all three sequenced mutants caused a significant decrease in β-carotene yield when introduced into YLH2 (3.7±0.6 mg/g DCW).

Assessing the effects of combinations of mutations on β-carotene production.

All mutations (except for the mitochondrial mutation) identified in each sequenced evolved mutant were reconstructed in the YLH2 parental background to determine whether the β-carotene yield in the isolated hyperproducers can be fully recapitulated. The results for the resulting strains YAG11 (SM12), YAG12 (SM13), and YAG13 (SM14) are shown in **Figure 4B**. Surprisingly, combining the selected mutations led to significant decrease in product yield (~8 mg/g DCW in all three reconstructed mutants) compared with SM12, SM13, and SM14. The results suggest that the mitochondrial mutation in SM12 and SM13 likely plays a role on β-carotene yield and/or that there are additional undetected beneficial mutations in the hyperproducers.

Since all three hyperproducers contain the detrimental mutation in *HIS7*, we reconstructed SM12, SM13, and SM14 strains without the *HIS7* mutation (resulting in strains YAG14, 15, and 16, respectively) to assess the extent of the negative impacts of this mutation in the multiple-mutations backgrounds. Results are shown in **Figure 4B**. YAG14-16 contain the same mutation combinations as that of YAG11-13, respectively,

with the exception of the *HIS7 389* mutation. Interestingly, strains YAG14 and YAG15 produced more β-carotene than SM12 and SM13 (p-value < 0.01), respectively, and YAG16 produced nearly as much as SM14. It is likely that due to the asexual nature of the ALE experiment, this detrimental *HIS7* mutation for β-carotene production was not able to be removed from the hyperproducers, leading to a reduction in the overall β-carotene yield in these strains.

Since multiple mutated genes were identified to have beneficial effects on β-carotene yield, we generated mutants containing combinations of mutations identified in different evolved strains (resulting strains YAG17 to YAG30) to determine whether superior combinations can be generated from these single mutations. All the superior recombinants identified (YAG22, YAG23, YAG26-YAG29) contain the *AQY3/DAK2* intergenic mutation, with the exception of YAG30 which contains the *SRO9*/*GFD2* intergenic mutation and the *YMRCTy1-3* mutation. The best combination comprises of the *YMRCTy1-3 1078*, *ALG6 1411*, and the *AQY3/DAK2 int* mutations (YAG28), which exhibited a β-carotene yield of 24.9 ± 0.5 mg/g DCW (**Figure 4C**). This superior strain exceeded the yield of the best evolved mutant SM14 by ~30%. The performance of YAG28 was further quantified using a benchtop bioreactor, and a β-carotene yield of 37.8±1.9 mg/g DCW, which is ~82% higher compared with SM14 (20.7±3.1 mg/g DCW), was observed. YAG28 also showed higher growth in the fermenter, reaching an OD_{600} around 20 as compared to SM14 which only reaches OD_{600} around 17 after 64 hours (p-value = 0.015 using two-tailed student t-test).

Decoupling the impacts of each mutation on hydrogen peroxide stress tolerance and β-carotene production

We had previously hypothesized that oxidative stress tolerance can be used as a selectable phenotype for increased β-carotene production in our prior ALE experiment, which led to the use of oxidative stress as the selective pressure (Reyes et al. 2014). It is possible that the mutations identified in the isolated hyperproducers conferred oxidative stress tolerance independent of β-carotene production, and that the increased productivity was a secondary effect. To determine whether this is the case, all single mutations (except for the mitochondrial mutation) were reconstructed in a wild-type S288c strain (FY2) and subjected to short-term H_2O_2 challenge. Since β-carotene itself is an antioxidant and protects against oxidative stress damage, we reconstructed these mutations in a strain that does not contain the carotenogenic cassette to decouple oxidative stress tolerance conferred by the mutation alone from that conferred from production of β-carotene.

As shown in **Table 3**, mutations in *MDS3* (YAG39), *EPL1* (YAG37), *FLO1* (YAG34), and *ALG6* (YAG38) conferred increased tolerance to H_2O_2 challenge in the non-carotenogenic background. On the other hand, mutations in *SCY1* (YAG36), *YMRCTy1-3* (YAG40), and the intergenic mutations in *DAK2/AQY3* (YAG35) and *SRO9/GFD2* (YAG32) resulted in decreased tolerance to hydrogen peroxide stress. This result runs counter to prior study showing that null mutation in *MDS3* led to decreased tolerance to H_2O_2 (Higgins et al. 2002). One possible explanation for this discrepancy is

the level of hydrogen peroxide used. In Higgins *et al*, null mutation in *MDS3* was shown to exhibit reduced growth in the presence of 4 mM of H_2O_2 in the culture media (Higgins et al. 2002) whereas in our study, the *MDS3 ins* mutation was found to exhibit increased viability after 1 h exposure to 1 M H_2O_2. Prior study found *MDS3* null mutant to exhibit reduced cell death in the presence of antifungal compounds (tunicamycin and dithiothreitol) that disrupt proteins/enzymes in the endoplasmic reticulum (ER) (Kim et al. 2012), suggesting that Mds3 is a death-promoting factor induced by ER stress. It may be possible that Mds3 has a death-promoting effect in the presence of high concentration of H_2O_2 similar to exposure to tunicamycin or dithiothreitol, and the inactivating *MDS3 ins* mutation led to increased viability. Interestingly, *ALG6* null mutant exhibits increased resistance to fluconazole, which disrupts a protein localized in the ER (Kapitzky et al. 2010). It is possible that the *ALG6 1411* mutation led to a similar tolerance mechanism as *MDS3 ins*; however, it is not known if the missense mutation is an inactivating mutation.

The intergenic mutation between *SRO9/GFD2* and the missense mutation in *YMRCTy1-3* that led to the highest increase in β-carotene yields in YLH2 background (YAG02 and YAG10, respectively), both caused significant reductions in hydrogen peroxide tolerance in the non-carotenogenic background (strains YAG32 and YAG40, respectively); suggesting that the two most beneficial mutations for β-carotene production were not selected due to beneficial effects in H_2O_2 tolerance that are independent of carotenoids production.

To test the validity of the original hypothesis that increased β-carotene production enhances H_2O_2 tolerance, the β-carotene biosynthesis cassette was introduced into each of the reconstructed mutants in FY2 to generate carotenogenic versions of each mutant. The resulting strains are YAG46-YAG55 with wild-type control YAG45. The β-carotene yields from the FY2 reconstructed mutants (**Table 3**) were, in general, lower compared with those in the YLH2 background, but the trend as to which mutations conferred the highest increase in production was consistent. The two mutations (*SRO9/GFD2 int* and *YMRCTy1-3 1078*) that caused significant decreases in H_2O_2 tolerance in the non-carotenogenic background (YAG32 and YAG40) showed the highest level of β-carotene yield in their carotenogenic counterparts (YAG47 and YAG55). YAG47 and YAG55 also exhibited significant increases in H_2O_2 tolerance, demonstrating that increased production of β-carotene is the cause of increased oxidative stress tolerance in these strains. Two of the mutations (*SCY1 1836* and the *DAK2/AQY3 int*) that exhibited decreased H_2O_2 tolerance compared with the wild-type control in the non-carotenogenic background both exhibited no difference in H_2O_2 tolerance relative to wild-type control in the carotenogenic versions along with modest increases in β-carotene yield; suggesting that increased β-carotene production helped to overcome the negative impacts of *SCY1 1836* and *DAK2/AQY3 int* mutations on oxidative stress tolerance. For the mutations that showed enhanced H_2O_2 tolerance in the non-carotenogenic background, the level of H_2O_2 tolerance appears to be independent of β-carotene production.

Table 3 Impacts of mutation on relative H_2O_2 tolerance in non-carotenogenic, carotenogenic, and carotenogenic strains lacking *CTT1* (taken from (Godara et al. 2019))

	FY2		FY2+carotenogenic			FY2+Δctt1+carotenogenic		
Mutation	Strain	relative H_2O_2 tolerance	Strain	relative H_2O_2 tolerance	β-carotene (mg/g DCW)	Strain	relative H_2O_2 tolerance	β-carotene (mg/g DCW)
wild-type	FY2		YAG45		7.2 ± 0.1	YAG56		6.8 ± 0.2
HIS7 389	YAG31	ND	YAG46	ND	**2.5 ± 0**	YAG57	ND	**2.4 ± 0.3**
SCY1 1836	YAG36	-	YAG51	ND	**9.4 ± 0.9**	YAG62	ND	**9.3 ± 0.1**
MDS3 ins	YAG39	+	YAG54	+	**9.7 ± 0.3**	YAG65	+	**9.3 ± 0.1**
EPL1 1754	YAG37	+	YAG52	+	**10.5 ± 0.1**	YAG63	+	**10.3 ± 0.2**
DAK2/AQY3 int	YAG35	-	YAG50	ND	**10.5 ± 0**	YAG61	ND	**10.3 ± 0.2**
FLO1 925	YAG34	+	YAG49	+	**11.2 ± 0.2**	YAG60	+	**10.8 ± 0.5**
ALG6 1411	YAG38	+	YAG53	ND	**11.4 ± 0.1**	YAG64	ND	**11.3 ± 0.1**
TYE7 86	YAG33	ND	YAG48	ND	**11.2 ± 0.1**	YAG59	ND	**11.2 ± 0.2**
SRO9/GFD2 int	YAG32	-	YAG47	+	**15.5 ± 0**	YAG58	+	**15.3 ± 0.2**
YMRCTy1-3 1078	YAG40	-	YAG55	+	**16.4 ± 0.1**	YAG66	+	**15.7 ± 0.3**

ND: not different from wild-type
+: ~10X higher survival compared to wild-type
-: ~10X lower survival compared to wild-type
Samples were treated with 1 M H_2O_2 for 1 hour. The Δ*ctt1* samples were treated with 100 mM H_2O_2 for 30 minutes.

Bold: p-value < 0.005 using two-tailed Student t-test compared with wild-type

We further generated a Δ*CTT1* strain to determine whether the presence of the cytosolic catalase gene significantly influences the relative H_2O_2 tolerance and β-carotene production in these mutants. Results showed very similar trends in both relative oxidative stress tolerance and β-carotene yields in the wild-type versus Δ*CTT1* strains (see **Table 3**), suggesting that the conclusions drawn regarding the impacts of each mutation on oxidative stress tolerance and β-carotene production are valid in both the *CTT1* wild-type and Δ*CTT1* backgrounds.

Overall, the data strongly suggest that the ALE strategy selected for few mutations that primarily increased H_2O_2 tolerance independent of β-carotene production (*e.g. MDS3 ins*), several mutations that conferred increased H_2O_2 tolerance both independent and dependent of increased β-carotene production (*EPL1 1754*, *FLO1 925*, and *ALG6 1411*), and mutations that conferred increased oxidative stress tolerance solely due to increased β-carotene production (*SCY1 1836*, *DAK2/AQY3 int*, *TYE7 86*, *SRO9/GFD2 int*, and *YMRCTy1-3 1078*). The mutations that conferred the highest increases in β-carotene yield were not selected for as a result of impacts other than increased product formation, suggesting that the oxidative stress challenge strategy used in the ALE experiments primarily selected for mutants with higher product yield as intended.

Total lipid content in reconstructed strains

Previous global gene expression profiling of a subset of the isolated adaptive mutants revealed upregulation of key genes involved in lipid biosynthesis (Reyes et al. 2014), suggesting lipid metabolism may be perturbed in the carotenoids hyperproducers. Indeed, when cultured under low nitrogen conditions, which has been shown to increase lipid content of oleaginous yeasts (EVANS and RATLEDGE 1984; Kolouchova et al. 2016; Sitepu et al. 2013) and algae (Fakhry and El Maghraby 2015; Richardson et al. 1969), strain YLH2 showed increased β-carotene yield (by ~298%) and total fatty acid content (Olson et al. 2016). To determine whether any of the mutations analyzed here impact lipid content, the relative lipid contents of these mutants were analyzed using Raman spectroscopy. Based on prior reports (Czamara et al. 2015; Weatherston et al. 2018) and comparison between the spectra of carotenogenic strain YLH2 and a non-carotenoid producing strain FY2, Raman peaks were identified for total fatty acid at 1449 cm^{-1}, unsaturated fatty acid at 1663 cm^{-1}, and carotenoid at 1156 cm^{-1}. Raman analysis was conducted for YLH2, SM14, all reconstructed single mutants (YAG01 to YAG10), and the best producers generated YAG22 and YAG28.

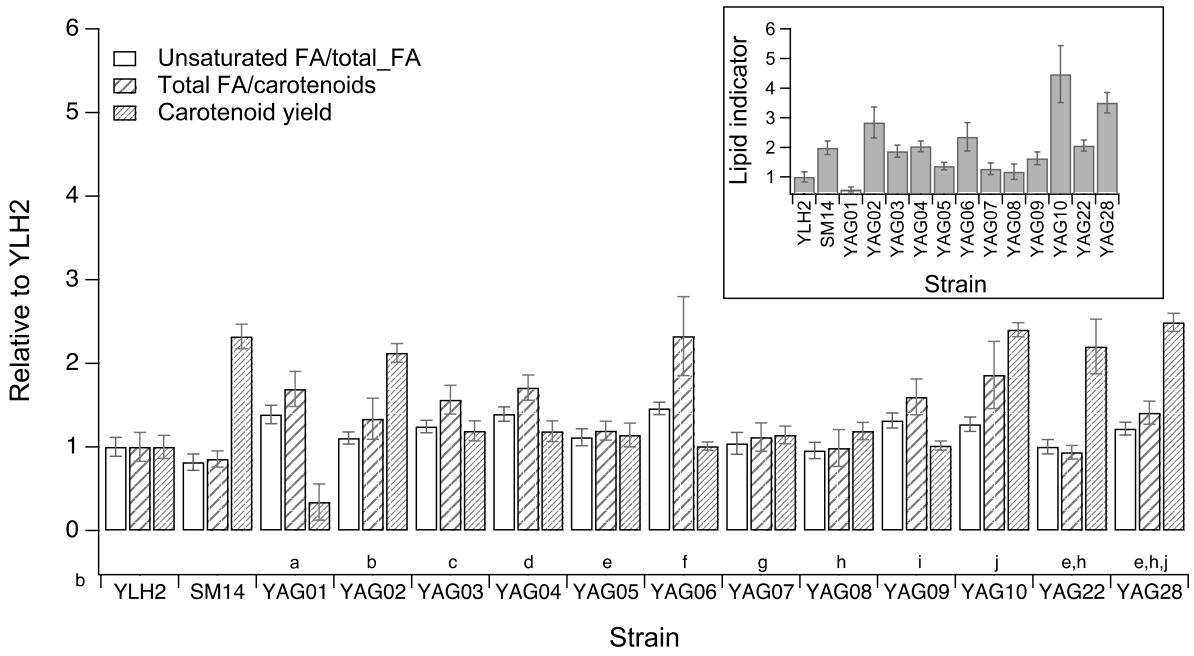


Figure 5 Relative Raman peak intensity ratios compared with YLH2. FA: fatty acid. Error bars represent a confidence interval about the sample mean, with p = 0.90. Mutations: (a) *HIS7 389*, (b) *SRO9/GFD2 int*, (c) *TYE7 86*, (d) *FLO1 925*, (e) *DAK2/AQY3 int*, (f) *SCY1 1836*, (g) *EPL1 1754*, (h) *ALG6 1411*, (i) *MDS3 ins*, (j) *YMRCTy1-3 1078* (taken from (Godara et al. 2019))

Results showed no significant differences in unsaturated fatty acid content in any of the reconstructed mutants compared with parental strain YLH2 (**Figure 5**). Thus, to compare between various mutants, the total fatty acid signal of each strain was normalized by its carotene signal and weighted by its carotene yield (equation shown in Materials and Methods). This yields a lipid indicator that is proportional to the quantity of total fatty acid per cell dry weight for each strain. All Raman data was further normalized against the carotenogenic parental strain, YLH2, in **Figure 5**. While the majority of strains tested showed relative lipid indicator values > 1, YAG10 and YAG28 exhibited the highest lipid content with relative lipid indicators ~4.5-fold and ~3.5-fold

greater than YLH2, respectively. Both YAG10 and YAG28 showed higher lipid indicators compared with SM14 (~2). These data suggest that increased β-carotene yield observed in strains YAG10 and YAG28 were likely partially due to increased lipid content of the strains.

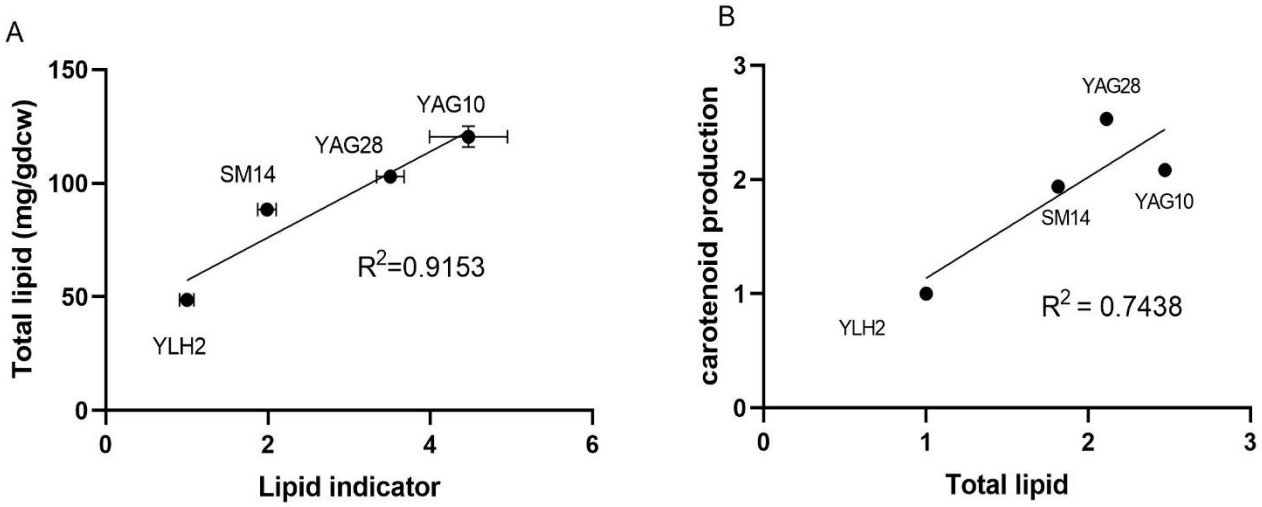


Figure 6 A) Correlation between lipid indicator obtained from Raman spectroscopy analysis and total lipid quantification results using the Bligh and Dyer method. Error bars are standard errors based on three biological replicates and three technical replicates each for lipid quantification results and three biological replicates and sixteen technical replicates for Raman spectroscopy analysis. YAG10 contains the *YMRCTy1-3 1078* mutation (j) and YAG28 contains *ALG6 1411* (h), *DAK2/AQY3 int* (e), and *YMRCTy1-3 1078* (j) mutations. B) Ratio of total lipid between hyperproducers vs ratio of beta-carotene production relative to YLH2 (taken from (Godara et al. 2019))

To validate the Raman spectral results, the total lipid content of strains YLH2, SM14, YAG10 and YAG28 was measured. Results are shown in **Figure 6**. The lowest amount of total liquid was observed in YLH2 with 48.7 ± 1.8 mg/g DCW. Results for SM14 showed 88.4 ± 2.8 mg/g DCW total lipid. Strains YAG10 and YAG28, which showed the highest lipid indicators, contained 120.5 ± 4.6 mg/g DCW and 103.0 ± 2.7

mg/g DCW of total lipid, respectively. The results from the gravimetric method and the lipid indicator values from Raman spectroscopy analysis showed linear correlation (R^2 value =0.9153) (**Figure 6A**). The level of increase in total lipid also positively correlates with increase in amount of β-carotene produced (**Figure 6B**). Overall, the increased lipid content in strains with higher β-carotene yields based on lipid indicators from Raman spectroscopy analysis is supported by total lipid quantification.

Mutations that confer increased β-carotene yield also increase lycopene production

To determine if the identified beneficial mutations for β-carotene production can also benefit the production of other carotenoids, we constructed the three best combinations of mutations ([*ALG6 1411* and *AQY3/DAK2 int*], [*ALG6 1411, EPL1 1754* and *AQY3/DAK2 int*] and [*ALG6, AQY3/DAK2 int* and *YMRCTy1-3 1078*]) in non-carotenogenic strain FY2, to generate strains YAG42, YAG43 and YAG44, respectively. The heterologous genes *crtE, crtB* and *crtI* for lycopene biosynthesis were integrated into these strains to generate lycopene-producing versions of these mutants. A control strain (YAG41) with just the lycopene biosynthetic genes integrated into FY2 was also constructed. As shown in **Figure 7**, the control strain produced 30.2 ± 0.3 mg/g DCW of lycopene. In strain YAG42, a 26% increase in lycopene production was observed, with total lycopene yield of 38.2 ± 0.3 mg/g DCW. The best combination of mutations for β-carotene production (*ALG6 1411, AQY3/DAK2 int* and *YMRCTy1-3 1078*) was also the best combination observed for lycopene production, with strain YAG44 producing

42.5 ± 0.3mg/g DCW. Based on knowledge gained from ALE experiments, we successfully generated a strain that achieved ~40% increase in lycopene yield compared with the wild-type YAG41.

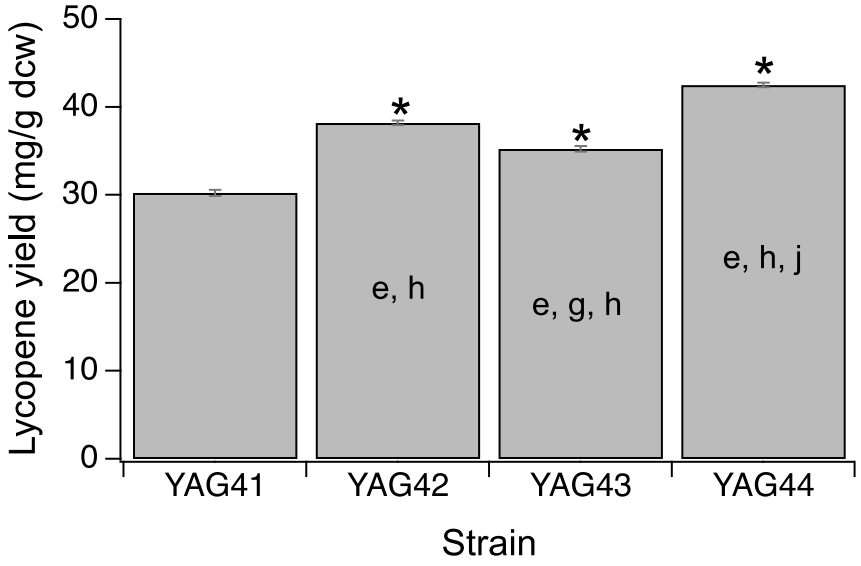


Figure 7 Effect of beneficial mutations on lycopene production. Asterisk: p-value < 0.05 using 2-tailed student t-test compared with YAG41. Mutations: (e) *DAK2/AQY3 int,* (g) *EPL1 1754,* (h) *ALG6 1411,* (j) *YMRCTy1-3 1078* (taken from (Godara et al. 2019))

Expression level of adjacent genes to intergenic mutations

Among the 9 reconstructed mutants in YLH2, 7 (in *TYE7*, *FLO1*, *EPL1*, *ALG6*, and *YMRCTy1-3*, intergenic region between *DAK2/AQY3* and intergenic region between *SRO9/GFD2*) showed statistically significantly increased β-carotene production, and 2 (in *SCY1* and *MDS3*) did not. Among the beneficial mutations for β-carotene yield, the missense mutation in *YMRCTy1-3* (strain YAG10) resulted in the largest increase in yield to 20.5±0.2 mg/g DCW (**Figure 2A**). Ty elements have been known to alter the

expression of neighboring genes; and a non-synonymous mutation can potentially result in changes in expression of downstream genes. Even though the two genes (*NUP116* and *IOC4*) flanking *YMRCTy1-3* are either upstream or divergently transcribed relative to *YMRCTy1-3*, we assessed the potential impact of this mutation on those genes using qRT-PCR (see supplementary info). As expected, the results showed no changes in gene expression of either of these two genes (**Table 11**). While the strain with the point mutation in *YMRCTy1-3* may have been adaptive due to possible changes in activity of the Ty1 element, it is difficult to envision how this may impact production and fitness in the reconstructed single mutant. The exact mechanism by which this point mutation led to increased β-carotene yield and general cellular fitness (reducing duration of lag phase) will be investigated in future work.

The remaining 5 beneficial mutations had moderate impacts on product yield; these include a mutation in the intergenic region between *DAK2* and *AYQ3*. Aqy3 is reported to play a role in passive glycerol diffusion (Oliveira et al. 2003). Dak2 is involved in stress adaptation and is required for detoxification of dihydroxyacetone (Molin et al. 2003), which may be formed from glycerol in the presence of hydrogen peroxide. The intergenic mutation may have led to changes in relative transcript abundance of *DAK2* and/or *AQY3*, potentially conferring the mutant a fitness advantage during the adaptive laboratory evolution by impacting cellular tolerance to H_2O_2; although the relationship between known functions of these genes and carotenoids production is not obvious. Surprisingly, qRT-PCR results revealed no significant changes in gene expression of either of these genes (see **Table 11**).

qRT-PCR data correlates with Raman and stress tolerance experiments

The conversion of HMG-CoA to mevalonate, catalyzed by Hmg1/Hmg2, is a rate-limiting step in sterol biosynthesis (Rodwell et al. 1976). Increasing expression of this gene has been successfully used as a strategy to improve flux towards terpene biosynthesis (Jackson et al. 2003; Ro et al. 2006). Quantitative reverse transcription PCR (qRT-PCR) was used to determine whether any of the mutations identified in the hyperproducers caused changes in relative transcript abundance of *HMG1*. Results showed no significant increase in expression of *HMG1* in all strains tested (**Table 12**). Hmg1 expression is also regulated at the translational and post-translational levels (Dimster-Denk et al. 1994; Hampton and Rine 1994), thus the qRT-PCR results do not preclude the possibility that one or more of the mutations being studied have an impact on Hmg1 activity.

In our earlier work, we observed a few genes involved in lipid biosynthetic process to be differentially upregulated in the carotenoids hyperproducers (Reyes et al. 2014). Since Raman spectroscopy analyses also showed increased lipid content in several of the reconstructed mutants, the expression of a subset of lipid biosynthesis related genes that were found to be upregulated in the evolved mutants (*CYB5*, *EEB1*, and *NSG1*) were measured using qRT-PCR in several of the reconstructed mutants. The *CYB5* gene encodes for cytochrome b5, which is involved in the sterol and lipid biosynthesis pathways by serving as an electron donor to cytochrome P450 (Truan et al. 1994). An increase in *CYB5* relative transcript abundance was observed in strains with the highest β-carotene yields (**see Table 12**); specifically, strains YAG02 (intergenic

mutation between *SRO9*/*GFD2*), YAG10 (mutation in *YMRCTy1-3*) and YAG28 showed at least a 3-fold increase in *CYB5* gene expression compared with YLH2. Interestingly, even though the intergenic mutation in *SRO9/GFD2* did not lead to any significant changes in the relative transcript abundances of *SRO9* or *GFD2*, the mutation caused a significant increase in *CYB5* expression in YAG02. The originally isolated hyperproducer SM14 also showed ~2.5 fold increase in *CYB2* gene expression. *EEB1* is involved in medium-chain fatty acid ethyl ester biosynthesis, with a preference for octanoyl-CoA as substrate (Saerens et al. 2006). Upregulation of this gene may lead to changes in cell membrane structure, although the exact relationship or mechanism of upregulation of this gene with increased carotenoids production is not clear. Besides being upregulated in SM14, *EEB1* also showed increased expression in reconstructed mutants with increased β-carotene yields (YAG10 and YAG28 strains). However, no clear trend correlating *EEB1* expression and β-carotene yield was observed. Interestingly, *EEB1* expression was significantly downregulated in strain YAG09, which contains the mutation in *MDS3*, compared with YLH2. Whether the differential expression of *EEB1* as a result of the mutation in *MDS3* plays any role in the negative epistasis between the *MDS3 ins* and other mutations on β-carotene yield remains to be explored. *NSG1* has been identified to play a key role in reducing Hmg2p degradation (Flury et al. 2005), and was found to be upregulated in β-carotene hyperproducers (Reyes et al. 2014); no correlation between expression of this gene and β-carotene yields was observed among all reconstructed strains tested. Overall, the data showed that

expression level of select genes in lipid biosynthesis appears to correlate with increase in lipid content but not necessarily in β-carotene yield.

Validation of spectroscopy data using high performance liquid chromatography

C18 column Agilent ZORBAX Eclipse Plus was used for all experiments with mobile phase 50:30:20 (*v/v/v*) of acetonitrile: methanol: isopropanol at 1 mL/min flow rate. Column temperature was kept constant at 40 ˚C and 450 nm wavelength was used for VWD detector for species detection. Standards of β-carotene and lycopene obtained from Sigma was dissolved in multiple amounts in dodecane to obtain a standard calibration curve HPLC. β-carotene in 3 yeast strains (YLH2, SM14 and YAG28) and their 3 biological replicates was tested, once extracted using cell disruption as described in section 2.6, using dodecane layer.

Amount of β-carotene observed from HPLC and amount observed from Tecan microplate reader at OD_{454} showed a strong correlation with R^2 values of 0.9916 as can be seen in **Figure 21**. Even though we also observed lycopene in β-carotene producing strains (**Figure 21 d**), the amount of lycopene produced is small (between 2.9%-5.3% of β-carotene). The relatively small amount of lycopene is not expected to significantly interfere with β-carotene quantification using absorbance measurements. Indeed, as shown in **Figure 21(e)**, β-carotene quantification results using absorbance measurements at OD_{454} correlated well with quantification using HPLC. Due to the ease of absorbance-based measurements, this was used for quantification of carotenoids in all strains tested.

Lycopene expression with different promoter

To optimize lycopene production in yeast, genes *CrtE*, *CrtB* and *CrtI* were expressed under different combinations of constitutive promoters with varying strengths: pTDH1, pCCW12 and pPGK1 in order of decreasing strength(Lee et al. 2015). Six integration plasmids at the URA3 locus were constructed and integrated into FY2 to generate 6 lycopene producing strains (see Table 7). Results show that promoter combinations have a significant impact on lycopene production (Table 13). The highest production was observed when *CrtE* was expressed under pPGK1, *CrtB* under pTDH3 and *CrtI* under pCCW12 promoter with 30.084 ± 0.4 mg/g DCW of lycopene production.

Potential synergistic effect between various mutations on β-carotene production

The existence of potential synergistic interactions between mutations on production was assessed using an additive model. Interestingly, all combinations tested exhibited negative epistasis except for 1 combination: AQY3/DAK2 int and ALG6 1411 mutations, which exhibited positive synergistic effect on carotenoids production (**Figure 22**). In addition, while the MDS3 ins mutation alone was neutral for β-carotene production, the results showed that the MDS3 ins mutation exhibited negative epistatic interactions in all combinations tested. As the insertional mutation in MDS3 likely led to inactivation of the gene, it is possible that a functional Mds3 is important for β-carotene production in yeast.

Conclusions

In our prior work, we hypothesized that increased productivity of β-carotene, a compound with antioxidant properties, is correlated with tolerance to oxidative stress. Using hydrogen peroxide as the selective pressure in adaptive laboratory evolution experiments, mutants with enhanced production of β-carotene were successfully isolated. In ALE experiments, the phenotypes of the mutants selected are dependent on the selection strategy. While we successfully isolated hyperproducers, it was not clear if β-carotene production was the primary phenotype being selected for, or if it was a secondary consequence of increased oxidative stress tolerance during the ALE experiment. In this work, we sequenced the β-carotene hyperproducers from the ALE experiments, identified beneficial mutations for β-carotene production, and demonstrated that the most impactful beneficial mutations for β-carotene production did not confer enhanced tolerance to hydrogen peroxide in the absence of carotenoids production. The results strongly suggest that the increased production of the desired antioxidant compound was indeed the phenotype selected for using this ALE strategy. By generating combinations of mutations identified from different β-carotene hyperproducers, we identified a superior combination of mutations that led to ~30% increase in β-carotene yield in test tubes and ~82% using a benchtop bioreactor compared with the best evolved hyperproducer SM14. Using Raman spectroscopy, total lipid quantification, and qRT-PCR, we found that the best β-carotene producers exhibit increased lipid biosynthesis. Finally, to demonstrate the broad utility of the findings from this work, we rationally constructed an improved lycopene producer using the best

combinations of mutations identified for β-carotene, resulting in a strain that produces ~40% more lycopene (42.5±0.3 mg/g DCW) compared with the baseline strain. In conclusion, we have successfully identified additional genes important for increasing carotenoids yields in yeast, which cannot currently be determined using rational approaches, and we have demonstrated that these mutations confer benefits to the production of carotenoids other than β-carotene.

CHAPTER III

ADAPTIVE LABORATORY EVOLUTION OF B-CARYOPHYLLENE PRODUCING *Saccharomyces Cerevisiae*

Background

Terpenoids are the largest and most diverse family of plant-derived compounds found in nature. Due to their wide industrial applications ranging from fuel alternatives, nutraceuticals, pharmaceutics etc., terpenoids are high-value compounds. Terpenes are secondary metabolites that are derived from C_5 precursor isopentenyl diphosphate (IPP) or dimethylallyl diphosphate (DMAPP) (Arigoni et al. 1997; Asadollahi et al. 2008). Extracting them from their natural source is usually not economical because of their low abundance. Microbial-based biosynthesis of terpenoids is a sustainable alternative for industrial production (Ignea et al. 2018; Wang et al. 2018; Wu et al. 2018).

Sesquiterpene is a family of terpenoids containing three isoprene units and they can be monocyclic, bicyclic or tricyclic in structure (Hüsnü et al. 2007; Russo et al. 2017). β-caryophyllene is a bicyclic sesquiterpene and has antioxidant and anti-inflammatory properties, with potential applications as an aircraft fuel alternative, antioxidant, anti-inflammatory product (Dahham et al. 2015; Tundis et al. 2017). Microbial production of β-caryophyllene has been demonstrated in *E. coli* and cyanobacteria reaching titers of ~1.5 g/L in a fed-batch bioreactor and 46.6 μg/L, respectively (Reinsvold et al. 2011; Yang and Nie 2016).

Apart from rationally engineering the known terpenoid biosynthesis pathway, due to complex and interlinked nature of metabolic network and cellular physiology,

genes and pathways not directly connected to the biosynthetic pathway may influence product formation (Alper et al. 2005; de Ruijter et al. 2017). A clear knowledge of the genotype-phenotype relation for product formation is not fully known. Thus, complementary techniques to rational engineering such as screening deletion or overexpression libraries (Arlt et al. 2011; Prelich 2012; Stevenson et al. 2001) and adaptive laboratory evolution (ALE)-based strategies (Portnoy et al. 2011; Reyes et al. 2014) can help to identify additional gene and/or pathway targets related with increasing product formation. In recent work, Promdonkoy *et al* utilized both rational engineering and ALE to improve the D-xylose utilization and isobutanol production in *Saccharomyces cerevisiae* (Promdonkoy et al. 2020). In another work, Rugbjerg *et al* showed evolved *Escherichia coli* MG1655 which utilizes glucose more efficiently due to a mutation in *rpoB* also exhibited improved mevalonate productivity (Rugbjerg et al. 2018).

Previously, we developed an environmental engineering-based strategy using ALE to improve product formation of an intracellular product with antioxidant properties by applying periodic oxidative stress challenge (Reyes et al. 2014). With extracellular products, the same strategy may fail due to potential "cheating" by non-producers in the population. In this work, we explore the potential application of ALE for the production of an extracellular product β-caryophyllene in *S. cerevisiae*. Initial strain optimization led to a strain producing ~3.8mg/g DCW β-caryophyllene. The optimized strain was used to design an optimum oxidative stress challenge strategy to improve product formation using ALE. Using the optimized strategy, evolved mutants

that exhibit a 4-fold increase in β-caryophyllene biosynthesis were isolated and characterized. Two mutations were identified to be responsible for the enhanced production phenotype observed.

Materials and methods

Strains, plasmids and growth conditions

All yeast strains used in this work are derivatives of S288c and listed in **Table 17**. *S. cerevisiae* strain BY4741 (MATa, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0)(Brachmann et al. 1998) are used as the base strains in this study. Yeast strains are cultured in Synthetic Complete (SC) media lacking amino acids for selection at 30 ºC and *E. coli* strains used for subcloning were cultured at 37 ºC in Luria Broth (LB) supplemented with appropriate antibiotics. Cytosolic catalase T (*CTT1*) gene was deleted in BY4741 strain as described in our earlier work (Reyes et al. 2014). For terpene quantification, yeast strains were cultured in YPD at 30 °C and 200 rpm for 72 h.

Plasmid construction for QHS1 and FPP overproduction

The heterologous genes *QHS1* from *Artemisia annua* were codon optimized for *Saccharomyces cerevisiae* and synthesized (Integrated DNA Technologies). Codon optimized *QHS1* gene was integrated into the yeast genome or expressed on plasmid under *URA3* selection marker. To increase flux towards FPP, *tHMG1, HMG2(K6R), UPC2-1* and *ERG20* were PCR amplified from yeast genome and added to chromosome

at *LEU2* locus under *LEU2* selection. The truncated *HMG1*, *tHMG1*, contains the1575-bp C-terminal part of *HMG1* and was amplified from yeast genome using PCR. The plasmids were constructed using the MoClo-YTK plasmid kit (Lee et al. 2015) (Addgene). In brief, each gene was introduced into the entry vector YTK001 using BsmBI (Thermo Scientific) via Golden Gate assembly protocol described by (Lee et al. 2015). Each gene (in an entry vector) is then assembled into transcriptional unit plasmid along with the appropriate promoter and terminator part vectors using BsaI (Thermo Scientific) via Golden Gate assembly. Finally, different transcriptional units are joined together to be used as plasmid or genomic integration with different selection markers for yeast. DH5α chemically competent *E. coli* cells were used for cloning and were prepared using Zymo Mix & Go! E. coli Transformation Kit and Buffer Set™ (Zymo Research). Yeast competent cells were prepared using Frozen-EZ Yeast Toolkit II Kit ™ (Zymo Research). To produce α-humulene *ZSS1* from *Zingiber zerumbet* was codon optimized for *Saccharomyces cerevisiae* and synthesized (Integrated DNA Technologies) and introduced into genome as described above. pTDH3 was used as promoter and tTDH1 was used as terminator for *ZSS1*.

Quantification of β-caryophyllene production

Organic layer of dodecane on top of media was used to capture volatile β-caryophyllene produced by strains. Unless specified a ratio of six-part media and one-part dodecane was used. Quantitative analysis of β-caryophyllene in dodecane layer was performed using Gas Chromatography(Sousa et al. 2011). Internal standard of alpha-

humulene was used for peak normalization. An Agilent J&W HP-5 (5%-phenyl)-methylpolysiloxane nonpolar column (30m x 0.32mm with 0.25μm film thickness) was used for this study. Gas chromatogram oven temperature was programmed from 100 °C initial temperature to 140 °C at 10 °C/min rate, followed by 2.5 °C/min to 180 °C, followed by 20 °C/min till finial temperature of 200 °C. FID detector was kept 280 °C whereas inlet was kept at 240 °C in a split less mode. Flow was kept at 2 ml/min with hydrogen flow at 30ml/min and ultra-pure air at 400 ml/min. Nitrogen was used for makeup flow at 25 ml/min. At least three biological replicates per strain were used for analysis.

Growth kinetics measurements

Microplate reader (TECAN Infinite ® M Nano) was used to measure growth curve for strains. Cells were grown for 24 h in test tube, then normalized to OD_{600} ~0.05 in 200 μL final volume in media in 96-well plates. Cells were cultured in the microplate reader for 72 h with orbital shaking at constant intervals. Shaking was performed using 3mm amplitude using kinetic cycles of 2 min incubation time, then orbital shaking for 3 min at 198 rpm followed by OD measurements. To obtain specific growth rate μ, duration of growth lag, and the maximum OD600 software grofit v1.1.1 was used(Kahm et al. 2010). Three biological replicates per strain each were used for analysis.

Oxidative stress tolerance using hydrogen peroxide

For spot assay overnight culture was normalized to OD 600 of 1.0 using phosphate-buffered saline (PBS). Strains were treated with hydrogen peroxide for 30 min intervals in a shaking incubator at 30 ºC at 200rpm. Cells were washed two times with PBS to eliminate the remaining hydrogen peroxide in solution. After the shock treatment, 10-fold serial dilutions were performed in (PBS) and spotted on SC plates lacking appropriate amino acid. Three biological replicates per strain were used for analysis. For stress optimization and evolution experiment the normalization step was omitted from the protocol.

Adaptive laboratory evolution

Single colonies of YAG115 were used to initiate the evolution experiment in 3 ml SC-ura-leu media on day 0. For populations that were evolved using the constant exposure strategy, 200 µl of cells were inoculated in fresh media supplemented with specified concentrations of H_2O_2. For populations that were subjected to periodic challenge, on day 1, 500 µl of culture was centrifuged, resuspended in phosphate buffer saline (PBS) and subjected to 50mM H_2O_2 challenge for 30 minutes under shaking (challenge period) and washed twice with PBS after exposure. Population samples were preserved in glycerol stocks each day. 200 µl of challenged cells were inoculated in 3 ml of SC-ura-leu media and 500 µl of dodecane was added for β-caryophyllene capture. On day 2, the population was allowed to recover by transferring 200 µl of overnight culture into fresh media (recovery period). The populations were challenged on odd days with

specified concentration of H_2O_2 and allowed to recover on even days. Production was quantified at the end of 24-hr by recovering the dodecane layer for populations evolved using either the continuous or periodic challenge strategies.

Isolating mutants from evolved populations

To select for hyperproducing mutants from the evolved populations, each population was subjected to exposure to a higher concentration of hydrogen peroxide than that used during the ALE experiments. 500 µl of each population sample was centrifuged and cells were resuspended in PBS with final concentration of 1 M of hydrogen peroxide. After 30 min of exposure, the cells were washed 2 time with PBS, and all cells were plated on SC -ura -leu plates. Plates were kept for 2 days at 30 °C. 8 colonies were randomly picked from each challenged population and streaked on SC-ura -leu plates to ensure we obtain individual clones. Single colonies were picked and cultured in 48 well plates containing 1 ml of media and ~166 µl dodecane overlay at 30ºC and shaking at 200 rpm. The dodecane layer was recovered to quantify β-caryophyllene production after 72 hr. To confirm the 48 well culture results, the high performing clones were grown in 3ml cultures in test tubes with 500 µl dodecane layer using three biological replicates.

Whole genome resequencing

Genomic DNA of isolated mutants, parental strain and population samples were extracted using YeaStar DNA extraction kit (Zymo Research). Library preparations and

NGS sequencing were performed by the Texas A&M Genomics Center for sequencing on the Illumina MiSeq platform using 300x300 paired end reads using Nexterra DNAFlex kit for library generation. An average coverage of >20-fold was obtained for each isolated mutant and for population samples a coverage of >150-fold was obtained. The sequencing data was aligned to *S. cerevisiae* S288c reference genome with breseq v0.29 (Deatherage and Barrick 2014). *De novo* mutations in isolated mutants were identified by comparing against the YAG113 parental strain, and verified via Sanger sequencing. The raw sequencing data were deposited in SRA database (https://www.ncbi.nlm.nih.gov/sra) with accession number PRJNA669136.

Reconstructing mutations into YAG114 and YAG118 strain

CRISPR-Cas9 using the one plasmid-system with pCRCT developed by (35) was used for site-directed mutagenesis to reconstruct identified mutations into YAG118. Briefly, 120 bp sequence including donor sequence and guide sequence (Table S3) was chemically synthesized by Twist Bioscience, USA, then introduced into the pCRCT plasmid using Golden Gate reaction using BsaI restriction sites, and transformed into E. coli cells, and plated on LB+Xgal for blue/white screening and incubated overnight at 37 °C. The bacterial colony with correct plasmid construction was verified by restriction digestion. The constructed CRISPR-Cas9 plasmid was then transformed into YAG114 and YAG118 strains and selected on SC -uracil plates. Colonies were picked and target mutations were first verified using PCR amplification refractory mutation system (ARMS) as described by Little et.al 2001 followed by confirmation via Sanger

sequencing. Cas9 plasmid was cured by serially passaging the strain in YPD 3 times and verified by PCR amplification.

Results and Discussion

Optimization of oxidative stress for use in adaptive laboratory evolution

Selecting an appropriate selective pressure (stressor) is key to successfully use of adaptive laboratory evolution. Since β-caryophyllene is a known antioxidant (Calleja et al. 2013; Dahham et al. 2015), an oxidative stress-based selective pressure can be used to aid the coupling of cellular growth or survival with production (Reyes et al. 2014). The strategy is based on the hypothesis that in an oxidative environment, the strain that produces more of an antioxidant product will have a growth advantage. To determine whether the production of β-caryophyllene increased cell survival, the catalase gene encoded by *CTT1* was first deleted in BY4741 to reduce the yeast native defense against hydrogen peroxide. Then strains that produce varying levels of β-caryophyllene were constructed in the Δ*CTT1* strain, resulting in YAG110 (BY4741 *ΔCTT1*), YAG111 (YAG110 with the β-caryophyllene synthase *QHS1* gene integrated in the genome), YAG114 (YAG110 with a genome-integrated *FPP* overproduction cassette) and YAG115 (YAG110 with *QHS1* and FPP overproduction cassette integrated in the genome). Production results showed YAG115 producing higher amount of β-caryophyllene than YAG111 (**Table 14**). All four strains were challenged with different concentrations of H_2O_2 (0mM, 50mM, 100mM, 150mM and 200mM) and their relative

viabilities were assessed (results are shown in **Figure 21**). The non-producers YAG110 and YAG114 showed the lowest levels of tolerance with ~10% survival with 50mM H_2O_2. Strain YAG111 showed an intermediate level of oxidative stress tolerance with ~10% survival after challenge with 150mM H_2O_2. The strain with the highest level of β-caryophyllene production YAG115 exhibited the highest tolerance with ~10% survival at 200mM H_2O_2, demonstrating the benefit of β-caryophyllene production on oxidative stress protection.

In addition to survival after exposure to 30 min challenge with higher concentrations of H_2O_2, cellular growth during continuous exposure at growth-permissible concentrations were also evaluated for each strain. H_2O_2 concentrations ranging from 0 mM, 25 mM, 50 mM, 75 mM and 100 mM were directly added to the media, and growth kinetics were measured. The data (in **Table 15**) showed no growth in the non-producer YAG110 for any H_2O_2 concentrations above 0 mM. Strain YAG111 and YAG114 showed some growth in 100 mM and 50 mM respectively. The higher tolerance of YAG114 over YAG110 may be attributed to increased accumulation of sterol because of the presence of the FPP overproduction cassette. The highest producer, YAG115, was the only strain able to grow in 100 mM hydrogen peroxide, although with a very long lag phase and with a low final cell density. This data showed the possibility of using continuous exposure as a possible alternative to periodic challenge as environmental stress for ALE experiment.

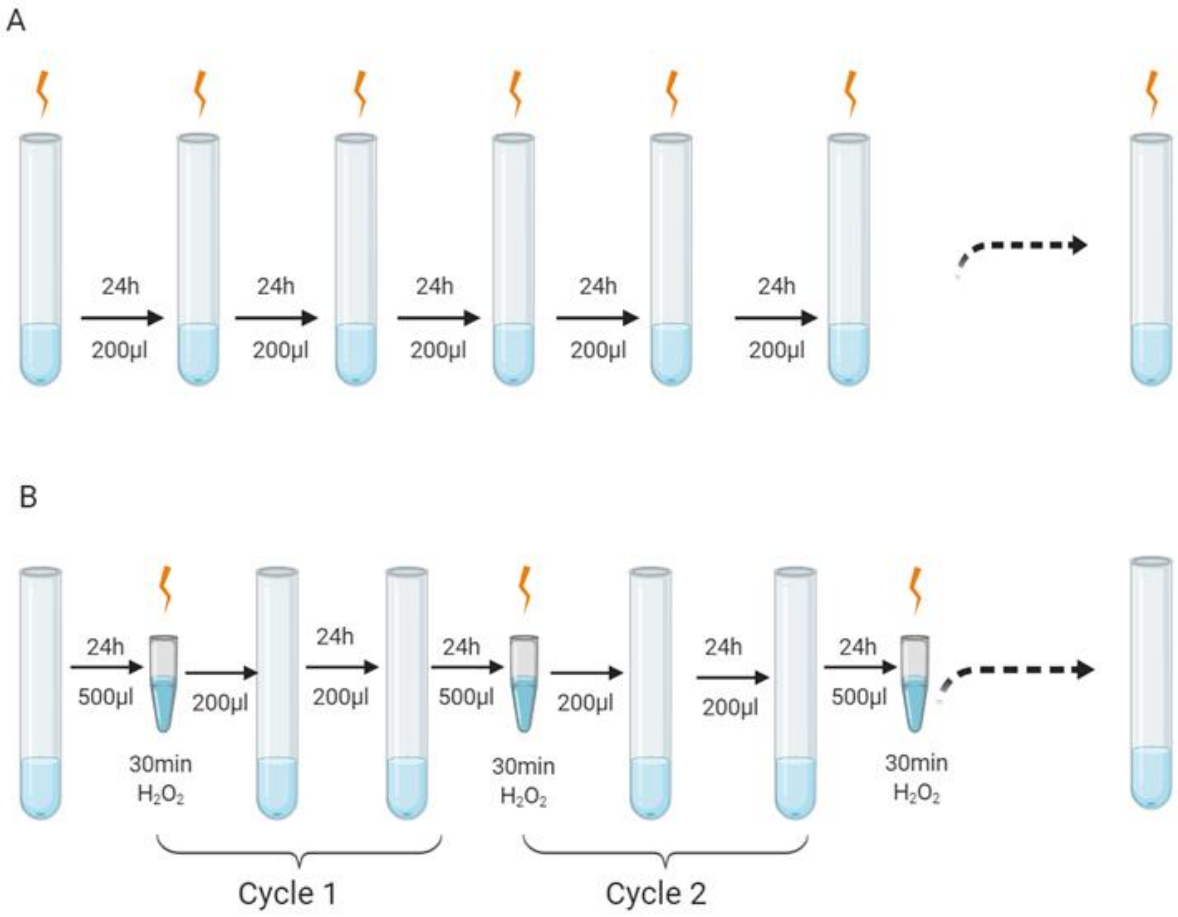


Figure 8 Schematic of different stress strategy tested. A. Continuous exposure with H_2O_2 directly added to media. B. Periodic exposure to H_2O_2 for 30 minutes, followed by a recovery period.

YAG115 was chosen as the parental strain for ALE to improve β-caryophyllene productivity since this strain exhibited the highest tolerance to oxidative stress. Prior to initiating an evolution experiment, two types of oxidative stress selection strategies were evaluated for their ability to improve β-caryophyllene productivity. A continuous

exposure strategy, where the cultures are continuously exposed to growth-permissible H_2O_2 in the media, and a periodic exposure strategy, where the cultures are subjected to a 30-min H_2O_2 challenge followed by a recovery period as described in **Figure 8**, were used. In all cases, a dodecane layer (500 µL) was added to each 3 mL of culture for β-caryophyllene capture. After every 24-hr of growth, ~7% (200 µL) of the culture was used to inoculate the subsequent culture (3mL of fresh media). Each strategy was evaluated for β-caryophyllene titer and yield after 8 days (continuous strategy) or 4 cycles (periodic challenge strategy). Results showed no significant increase in yield using continuous exposure at 25mM H_2O_2 compared to the no stressor control (**Figure 9**). A decrease in total yield was observed for all other H_2O_2 concentrations using continuous exposure. On the other hand, significant increases of up to 1.7-fold in production were observed in populations subjected to periodic challenge with 50 mM H_2O_2. Periodic challenge with 100 mM H_2O_2 resulted in an insignificant increase in production, with higher concentrations showing complete loss of production. The results revealed that periodic challenge at a H_2O_2 concentration that resulted in ~10% survival was an optimal selection strategy for increasing β-caryophyllene production using ALE.

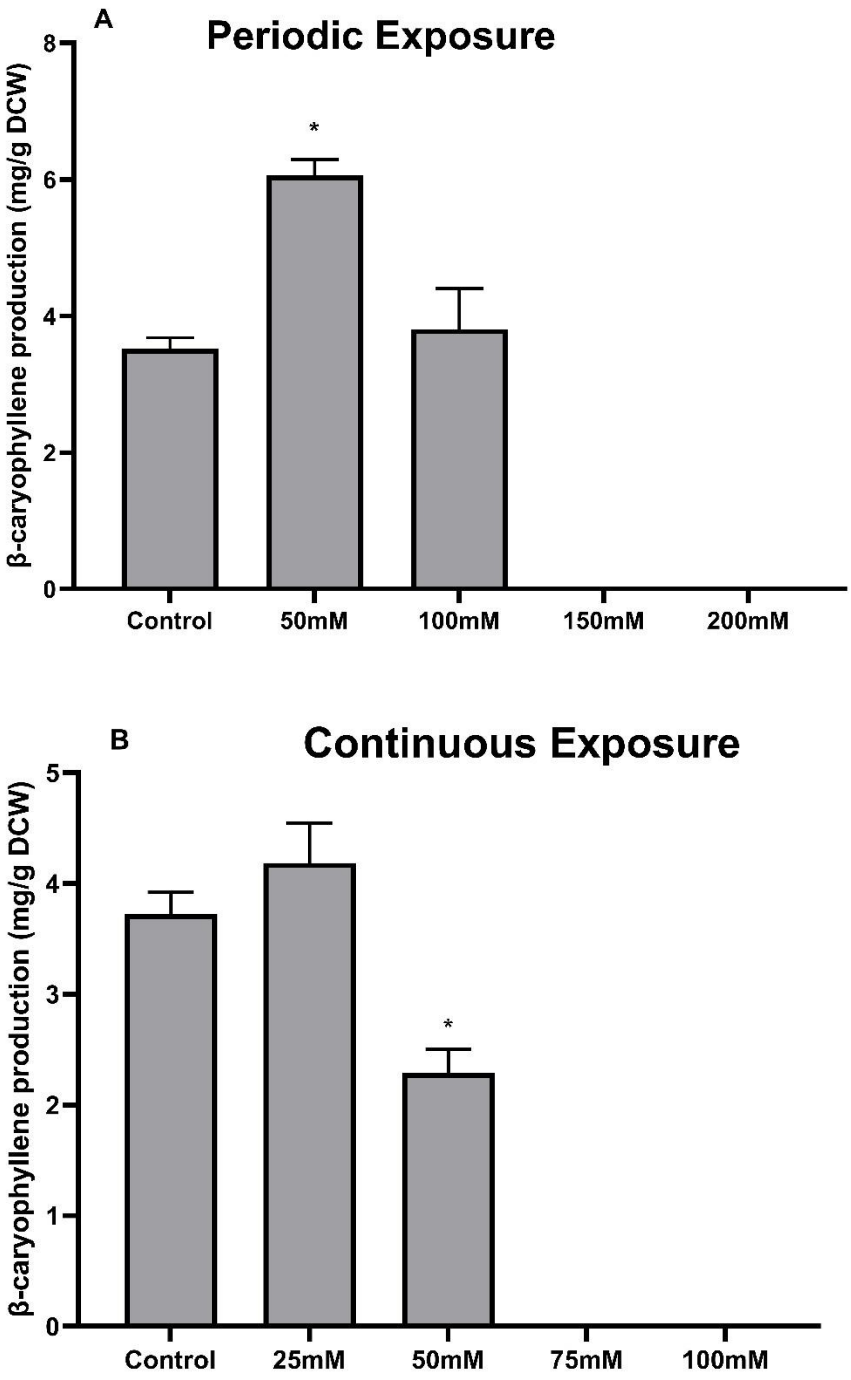


Figure 9 Average β-caryophyllene production observed after short-term selection using hydrogen peroxide for 8 days using A) periodic challenge and B) continuous exposure at various hydrogen peroxide concentrations. Asterisks: p value <0.05 using two-tailed Student's t-test against control.

Increased β-caryophyllene production via adaptive laboratory evolution using periodic H_2O_2 challenge

An ALE experiment was initiated with YAG115 using the periodic challenge strategy with an initial concentration of 50 mM H_2O_2. To further optimize the ramp-up in the ALE, at the end of cycle 2 (day 5), the culture was split into two populations, P1 and P2, which were exposed to 50 mM and 100 mM H_2O_2 exposure, respectively, in ramping up the selection pressure (**Figure 10**). The next split and ramp-up was done at the end of cycle 4 (day 9) from the P2 population (100mM exposure) into P3 and P4, which were subsequently exposed to 150mM and 200 mM H_2O_2 exposure, respectively. After the split, four populations (P1, P2, P3, and P4) were maintained. At cycle 5, P1 and P2 reached a peak in production, reaching ~3-fold increase in β-caryophyllene production before a stagnation or decrease in production were observed. On the other hand, populations P3 and P4 showed consistent decreases in production over time. Thus, population P2 was chosen for further analyses.

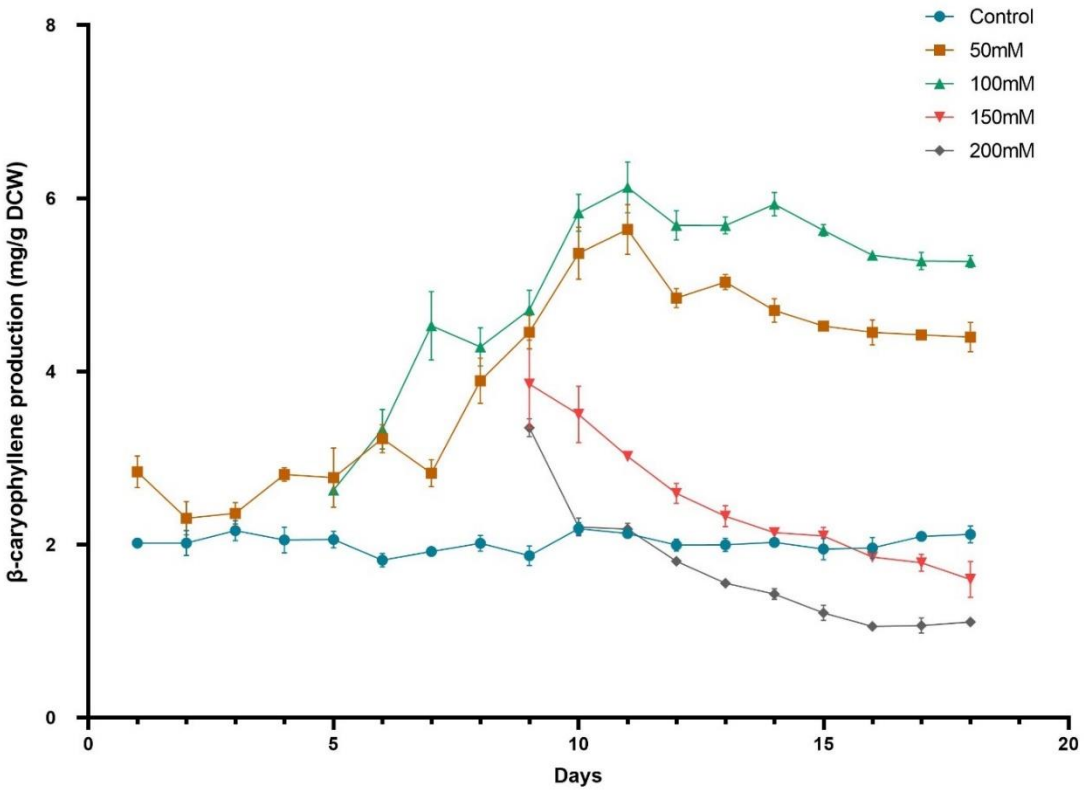


Figure 10 Production observed during evolution experiments.

Isolated evolved mutants exhibit significant increase in production

Under the assumption that evolved mutants with higher production of β-caryophyllene should have enhanced survival after oxidative stress challenge compared to the parental strain, we first identified a concentration of H_2O_2 in which the parental strain has negligible viability after a 30-minute exposure. As shown in **Figure 11**, 1 M H_2O_2 exposure resulted in no growth of the parental strain, whereas evolved population P2 (day11) had ~10% viability. In addition, we hypothesized that colony size after H_2O_2 challenge can be used to estimate relative productivity. Population P2 (day 11) was exposed to 1 M hydrogen peroxide stress for 30 min and plated on SC -ura -leu plate.

Three sizes of colonies were observed, small, medium and large. These colonies were cultured and their productivities were quantified. A positive correlation was observed in size and production as medium and large colonies (see **Figure 22**). The result suggests a growth advantage of mutants with higher production of β-caryophyllene, indicating the evolution strategy resulted in growth/survival and production coupling as hypothesized.

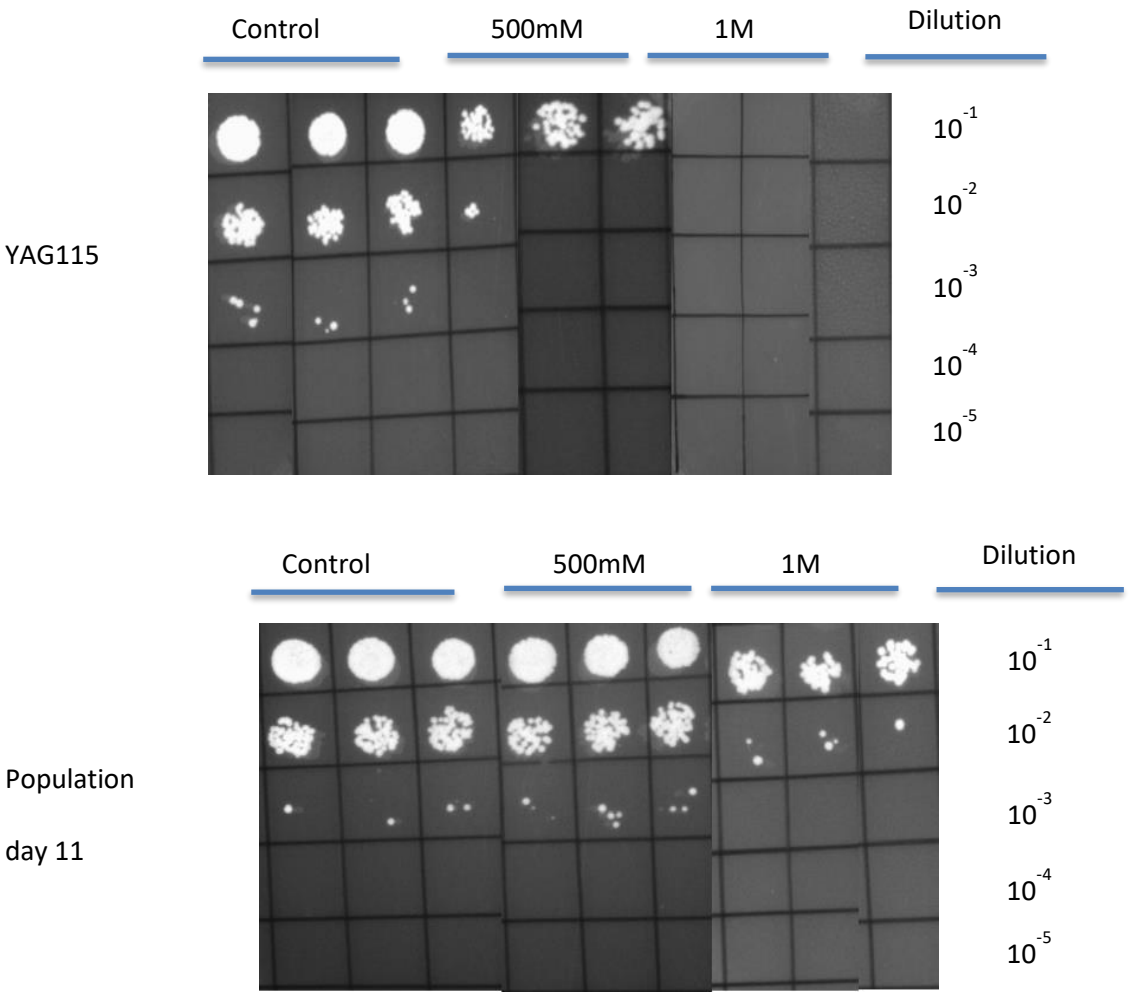


Figure 11 Spot assay for screening concentration of hydrogen peroxide

Using a 30-min exposure to 1 M hydrogen peroxide, we subjected four time point samples from the evolved population: P1 (day 4), P2 (day 8), P2 (day 11), P2 (day 14) and P2 (day 18) to screen for hyperproducers. We failed to identify any viable isolates from population P1 (day 4). For the remaining populations, eight individual colonies per population were randomly chosen. The isolated mutants were subjected to a second round of screening based on their β-caryophyllene production. **Figure 12A** shows the individual production obtained from each isolate. The top five overproducing evolved mutants were selected from the 32 total mutants, and their β-caryophyllene production were characterized in more detail. Results showed a 2-4 fold increase in production compared to the parental strain, with mutant P11M1 (population 2 day 11 mutant 1) being the best performer as shown in **Figure 12B**. The stability of P11M1 was assessed via four serial passages in SC -ura -leu media (3 days of growth each passage for a total of 12 days); β-caryophyllene production was monitored after each passage. The parental strain was included as control. P11M1 showed consistent production of ~16 mg/g DCW, whereas the control produced ~4 mg/g DCW.

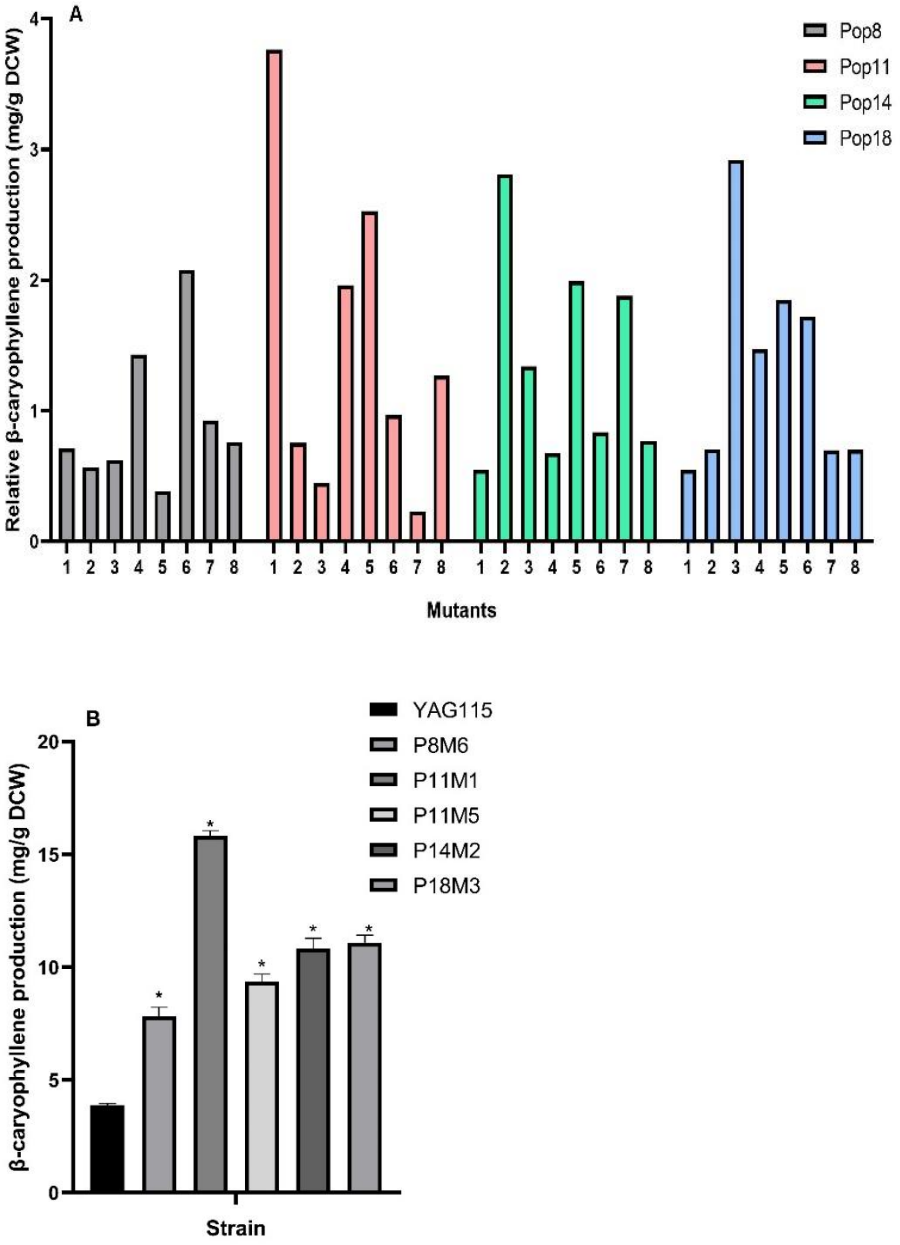


Figure 12 Isolated evolved mutants exhibited increased β-caryophyllene production over the parental strain. A) Initial screening of isolated mutants from each evolved population based on relative product yield in 48 well plates. Pop8 = P2 (day 8), Pop11 = P2 (day 11), Pop14 = P2 (day 14), and Pop18 = P2 (day 18). B) Confirmation of product formation in the five best performing mutants. The mutants are named first by the population they were isolated from followed by the isolate number. For example, P11M1 is the first mutant from Pop11. Asterisks: p value <0.05 using two-tailed Student's t-test against control.

To determine whether *QHS1* is the rate-limiting step in β-caryophyllene production in the evolved mutants, the *QHS1* gene was overexpressed using 2μ plasmid into both the parental and the best evolved isolate P11M1, resulting in strain YAG116 and YAG117, respectively. Overexpression of *QHS1* led to modest increases in production in the parental strain from ~4mg/g DCW to about ~5.3 mg/g DCW and P11M1 from ~16 to ~18 mg/g DCW (**Figure 23**), suggesting that β-caryophyllene synthase is not a major rate-limiting step in β-caryophyllene production in the evolved mutants.

Genome sequencing identified mutations outside β-caryophyllene pathway

To identify beneficial mutations for the observed increased productivity in the isolated mutants, we sequenced four time-course population samples from population P2 (day 8, day 11, day 14 and day 18), five of best isolated mutants, and the parental strain. A mutation frequency cutoff of greater of equal to 0.8 was used to narrow down the dominant mutations in the population. Large numbers (>200) of mutations above the cutoff threshold were observed in the population samples. The identified mutations in single hyperproducer isolates are shown in **Table 16**. The high frequency mutations (in populations) that were also identified in sequenced single isolates from the respective populations were selected for further analyses. In addition, since isolate P11M1 was the best producer identified, all mutations in P11M1 were also chosen, for a total of 13 unique mutations selected for further analyses (**Table 4**). To determine the impacts of each of these mutations on β-caryophyllene biosynthesis, single mutations were

introduced into parental strain using site-directed mutagenesis, resulting in strains YAG132-145. Only two mutations led to improved productivity, the mutation in *STE6* gene and a mutation in intergenic region of between *MST27/tR(UCU)G1* (**Figure 13**). *STE6 1025* mutation improved the productivity 3.7 fold to 12.6 mg/g DCW of β-caryophyllene. *STE6* encodes the plasma membrane ATP binding cassette (ABC) transporter known to export a-factor in MATa cells (Michaelis 1993). The *STE6 1025* mutation is a missense mutation resulting in a change in amino acid from T1025N. Thus, the potential benefit of *STE6 1025* mutation on productivity may be related to product transport. The *MST27/tR(UCU)G1* intergenic mutation improved productivity 3 fold to 10.3 mg/g DCW of β-caryophyllene compared to the parental strain. The *MST27/tR(UCU)G1 int* mutation is an insertion mutation between *MST27* and a Ty element (Kim et al. 1998). *MST27* is member of the *DUP240* multi-gene family and known to impact the vesicles formation (Sandmann et al. 2003). Ty1 elements is a transposon element and mutations in transposon elements are commonly found in evolution experiments and are known to impact fitness (Gresham et al. 2008; Wilke and Adams 1992). However, as this mutation is in the intergenic region and at the 3' end of *MST27*, it is difficult to ascertain the exact impact of the mutation on the cell. No further increase in production was observed when these two mutations were combined.

Table 4 Mutations chosen for detailed characterization. All mutations were verified by Sanger sequencing.

Chromosome	Position	Mutation	Amino acid change	Gene	Abbreviation
3	286312	C→T	nonsense mutation	CDC39	CDC39
10	715141	A→G	T200T	DAN4	DAN4 200
1	27105	A→G	T288T	FLO9	FLO9 288
3	151555	+A	Intergenic	tK(CUU)C/MAK32	tK(CUU)C/MAK32 int
7	128474	T→A	nonsense mutation	MDS3	MDS3
7	404475	+G	Intergenic	MST27/tR(UCU)G1	MST27/tR(UCU)G1 int
7	530034	A→C	S257S	MTL1	MTL1 257
2	754982	C→T	D709N	RIF1	RIF1 709
9	241053	$(A)_{21\rightarrow22}$	Intergenic	RNR3/FIS1	RNR3/FIS1 int
14	12986	$(T)_{11\rightarrow14}$	intergenic	SNO2/SNZ2	SNO2/SNZ2 int
11	43222	G→T	T1025N	STE6	STE6 1025
1	12690	A→T	Intergenic	YAL064W-B/TDA8	YAL064W-B/TDA8 int
8	2303	$(C)_{11\rightarrow12}$	Intergenic	YHL050C/YHL050C	YHL050C int

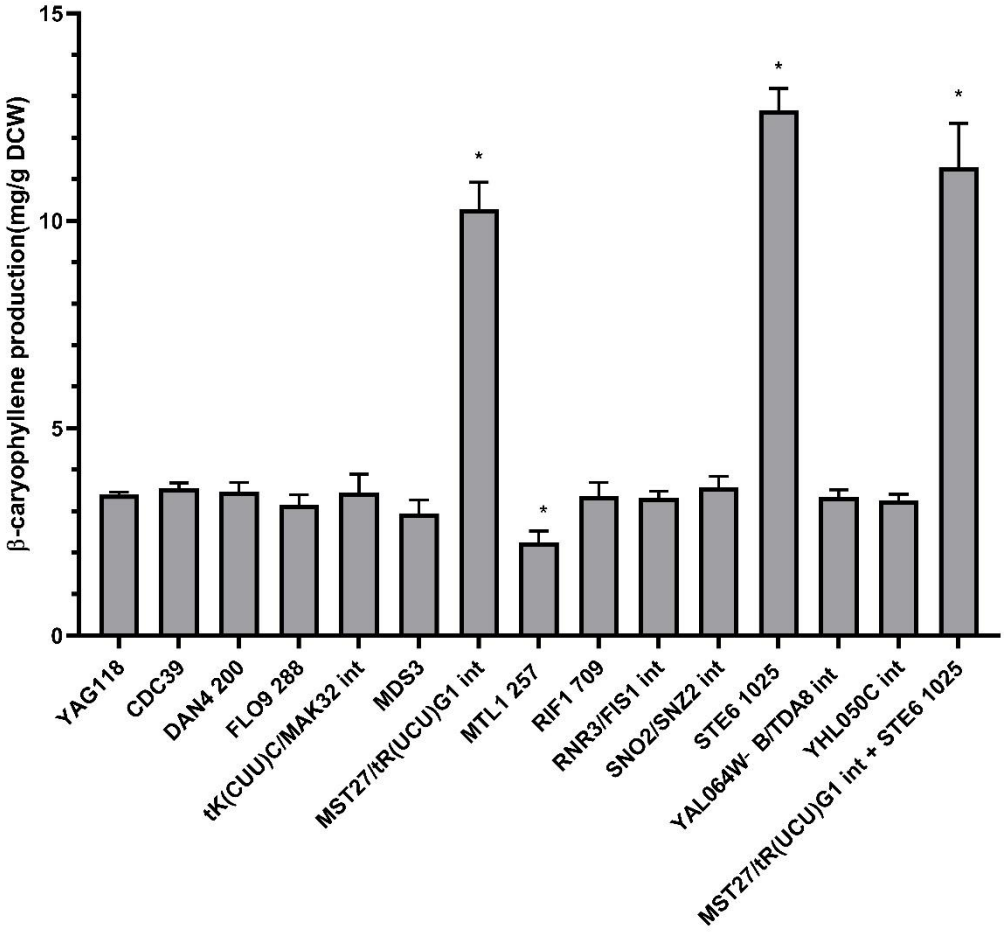


Figure 13 Production of β-caryophyllene in reconstructed single mutants. Asterisks: p-value <0.05 using two-tailed Student's t-test against control.

Interestingly, both the *STE6 1025* and *MST27/tR(UCU)G1* intergenic mutations were found in very low frequencies in the population (below detection limit). The *STE6 1025* mutation was found only in mutant P11M1 and *MST27/tR(UCU)G1* intergenic mutation was identified in all the isolated mutants sequenced. Note that the population samples sequenced were cultured in normal growth conditions before harvesting for next-gen sequencing, while the isolated mutants underwent a round of selection in 1 M

hydrogen peroxide. The findings that the *MST27/tR(UCU)G1* intergenic mutation was not detected in the population samples, but is present in all isolated mutants, and the positive correlation between cellular survival (based on colony size after H_2O_2 challenge; **Figure 22**) and product formation, suggest that the evolving population was highly heterogeneous with the most productive members consisting of a smaller fraction of the population.

Growth kinetics and oxidative tolerance of reconstructed mutations

In order to investigate the general fitness benefit conferred by the individual mutations identified, we did a growth kinetics study for all individual reconstructed strains. The data showed that five (*CDC39, tK(CUU)C/MAK32 int, RIF1 709, STE6 1025* and *YAL064W B/TDA8 int*) out of 13 mutations did not impact growth under normal culture conditions (see **Table 5**). The *FLO9 288, RNR3/FIS1 int*, and *SNO2/SNZ2 int* mutations exhibited increased growth rates. The *DAN4 200, MDS3, MTL1 257, MST27/tR(UCU)G1 int, and FLO9 288* mutations showed significant reduction in lag phase. Interestingly, mutation in *FLO9 288* positively impacted specific growth rate, lag phase, and max OD600, suggesting a significant fitness benefit, which likely allowed it to reach fixation in the evolving population. Mutations and changes in copy number in the *FLO* genes have been found to be involved in tolerance to environmental stressors (Fidalgo et al. 2006; Wallace-Salinas et al. 2015; Watanabe et al. 2013).

Table 5 Growth kinetics data for reconstructed strains. Bold : p value <0.05 using two-tailed Student's t test compared with control.

Strain	Mutations	Growth rate	stdev	lag phase	stdev	OD600Max	Stdev
YAG118	Control	0.169	0.006	3.347	0.139	1.052	0.016
YAG132	CDC39	0.184	0.012	3.340	0.273	1.094	0.059
YAG133	DAN4 200	0.178	0.017	**2.712**	**0.284**	1.037	0.045
YAG134	FLO9 288	**0.197**	**0.007**	**2.188**	**0.379**	**1.145**	**0.014**
YAG135	tK(CUU)C/MAK32 int	0.164	0.016	3.430	0.146	0.988	0.093
YAG136	MDS3	0.177	0.007	**0.716**	**0.135**	1.074	0.036
YAG137	MST27/tR(UCU)G1 int	0.174	0.007	**2.387**	**0.100**	1.064	0.039
YAG138	MTL1 257	0.175	0.011	**1.120**	**0.149**	1.050	0.048
YAG139	RIF1 709	0.186	0.012	3.370	0.105	1.088	0.019
YAG140	RNR3/FIS1 int	**0.196**	**0.008**	3.199	0.095	1.119	0.059
YAG141	SNO2/SNZ2 int	**0.204**	**0.005**	3.211	0.022	**1.178**	**0.022**
YAG142	STE6 1025	0.178	0.014	3.145	0.226	1.082	0.077
YAG143	YAL064W-B/TDA8 int	0.170	0.031	3.178	0.353	1.100	0.140
YAG144	YHL050C int	0.173	0.008	**3.679**	**0.027**	1.071	0.059
P11M1	Multiple	0.185	0.011	**2.336**	**0.175**	1.106	0.051

To assess the impacts of these mutations on oxidative stress tolerance, the individual reconstructed mutants (β-caryophyllene producers) were subjected to 200 mM hydrogen peroxide exposure for 30 minutes. The same set of mutations (listed in **Table 4**) were also reconstructed in a background strain lacking the β-caryophyllene synthase gene, resulting in strains YAG119-YAG131, which were used to study the impacts of these mutations on oxidative stress tolerance in the absence of production. In the absence of β-caryophyllene production, an increased 10X survival compared to the reference strain (YAG114) was observed in YAG121 (*FLO9 288*), YAG123 (MDS3), YAG124 (*MST27/tR(UCU)G1 int*) and YAG125 (*MTL1 257*) (**Table 6**). In the β-caryophyllene producing background, 10X increased tolerance was observed in YAG134 (*FLO9 288*), YAG136 (MDS3), YAG137 (*MST27/tR(UCU)G1 int*), YAG138 (*MTL1 257*) and

YAG142 (*STE6 1025*). Mutations in all *FLO9* and *MDS3* were found in detectable frequencies in all populations sequenced, and the *MTL1* mutation was present at ~49% frequency in the P11 population, suggesting that the increased oxidative stress tolerance conferred by these mutations likely contributed to their selection in the ALE experiment. Mutations in *FLO9* and *MTL1* were found in an industrial yeast strain that was evolved for growth on hydrolysates inhibitors (Wallace-Salinas et al. 2015), which have been shown to induced oxidative stress in *S. cerevisiae* (Almario et al. 2013). The increased hydrogen peroxide tolerance in YAG142 (*STE6 1025*) can be attributed to increased production, as without β-caryophyllene production there was no benefit to survival in oxidative stress challenge. The intergenic mutation in *MST27/tR(UCU)G1*, which is one of the two mutations found to be responsible for increased β-caryophyllene production, showed general fitness benefit with reduced lag phase in normal growth conditions and ~10X increase in survival in the presence of hydrogen peroxide stress. While this mutation is likely selected for due to the fitness it confers in the presence of strong oxidative stress, although it's frequency in the population is extremely low, it is unclear how it contributed to increased β-caryophyllene production.

Table 6 Hydrogen peroxide stress tolerance for strains with and without β-caryophyllene production after 200 mM H_2O_2 exposure for 30 minutes. ND: not different from wild-type. +: ~10X higher survival compared to wild-type. Bold: p value <0.05 compared with control using two-tailed Student's t test compared with control.

	No production		With Production		
Mutation	Strain	relative H_2O_2 tolerance	Strain	relative H_2O_2 tolerance	β-caroyophyllene (mg/g DCW)
Control	YAG114		YAG118		3.39
CDC39	YAG119	ND	YAG132	ND	3.55
DAN4 200	YAG120	ND	YAG133	ND	3.47
FLO9 288	YAG121	+	YAG134	+	3.15
tK(CUU)C/MAK32 int	YAG122	ND	YAG135	ND	3.44
MDS3	YAG123	+	YAG136	+	2.94
MST27/tR(UCU)G1 int	YAG124	+	YAG137	+	**10.27**
MTL1 257	YAG125	+	YAG138	+	2.24
RIF1 709	YAG126	ND	YAG139	ND	3.36
RNR3/FIS1 int	YAG127	ND	YAG140	ND	3.31
SNO2/SNZ2 int	YAG128	ND	YAG141	ND	3.57
STE6 1025	YAG129	ND	YAG142	+	**12.66**
YAL064W-B/TDA8 int	YAG130	ND	YAG143	ND	3.35
YHL050C int	YAG131	ND	YAG144	ND	3.25

Effect of *STE6 1025* on caryophyllene production

Since the missense mutation in *STE6* led to a large impact on β-caryophyllene productivity, we assessed whether the mutation is a loss-of-function or gain-of-function mutation using *STE6* knockout (YAG146) and overexpression (YAG147 with *STE6 1025* and YAG148 with wild-type *STE6*) strains, and quantifying their effects on β-caryophyllene productivity (data shown in **Figure 14**). The *STE6* knockout strain exhibited no change in production compared with the reference. Production was also not impacted when the wild type *STE6* was overexpressed. However, when we overexpressed the mutated *STE6* in the reference strain, β-caryophyllene production increased 4-fold to 13.8 mg/g DCW, suggesting the *STE6 1025* mutation is a gain-of-function mutation. Since the mutation is located between the transmembrane and ATP binding domains, it potentially impacted substrate recognition, allowing increased export of β-caryophyllene. However, further experiments are needed to identify the exact effect the mutation has on transport of β-caryophyllene.

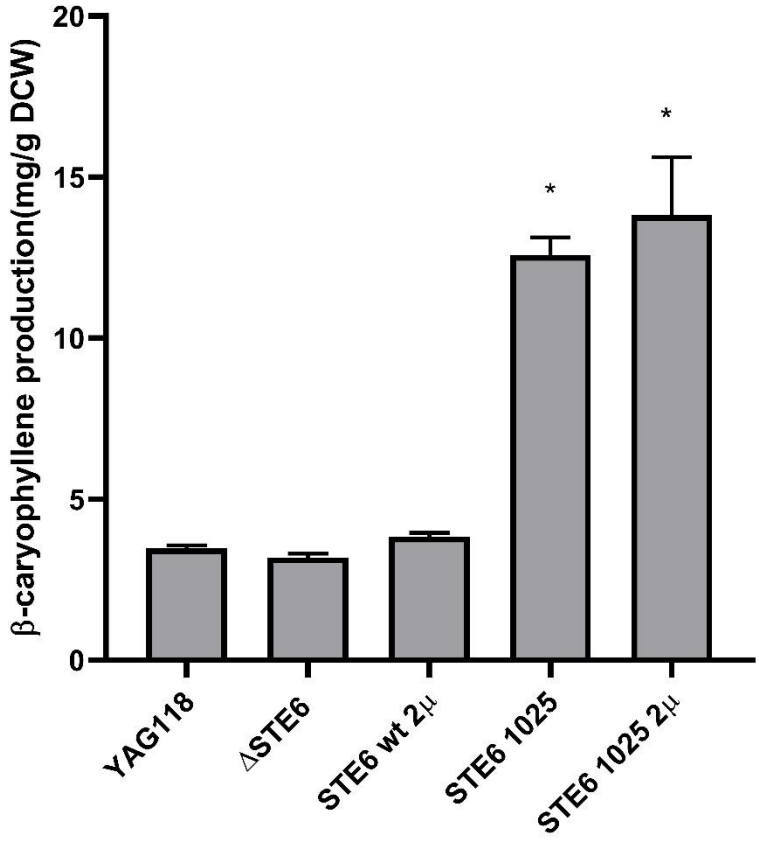


Figure 14 The effect of *STE6* expression on β-caryophyllene production. ΔSTE6: *ste6* deletion strain. *STE6* wt 2μ: wild-type *STE6* expressed on a 2μ plasmid. *STE6 1025*: reconstructed parental strain with Ste6 T1025N mutation. *STE6 1025* 2μ: parental strain expressing mutated *STE6* gene on a 2μ plasmid. A=Asterisks: p value <0.05 using two-tailed Student's t-test against parental strain.

We further explored the possibility that the *STE6 1025* mutation may influence the efflux and thus the productivity of other sesquiterpenes by overexpressing the mutated gene in an engineered α-humulene producer. No increase in production was observed in the strain that overexpresses *STE6 1025* compared to the reference (**Figure 24**), suggesting that the mutation specifically affects the transport of β-caryophyllene and does not significantly influence the efflux of sesquiterpenes in general.

Conclusions

In this work, we explored the use of adaptive laboratory evolution to improve the production of an extracellular product, β-caryophyllene, using oxidative stress challenge as selection. Initial metabolic engineering by overexpressing isoprenoid pathway genes resulted in a parental strain that produces ~4 mg/g DCW of β-caryophyllene. An ALE strategy was optimized by comparing periodic stress challenge versus continuous exposure. Using the optimized periodic H_2O_2 challenge method (from 50mM to 200mM H_2O_2), the ALE experiment resulted in evolving populations with up to 15.8 mg/g DCW (>3-fold increase) in β-caryophyllene production. Under the assumption that a positive correlation exists between oxidative stress tolerance and β-caryophyllene production, a final selection using 1 M H_2O_2 for 30 minutes was used to isolate mutants with the highest level of oxidative stress tolerance. This strategy yielded isolated mutants with up to 4-fold increase in β-caryophyllene production. Mutations in the intergenic region of *MST27/tR(UCU)G1* and a non-synonymous mutation in *STE6* were the only mutations found to benefit product formation. *STE6* is a known ABC exporter for a-factor in yeast. Deletion and overexpression studies demonstrated that the mutation is a gain-of-function mutation. It may be possible that the *STE6 1025* mutation resulted in binding and export of β-caryophyllene. However, the exact mechanism for how this point mutation is affecting β-caryophyllene productivity is still unclear. Overall, we demonstrated that by leveraging the antioxidant potential of terpenes that are exported extracellularly, product formation can be coupled with cellular growth/survival and be rapidly improved using short-term ALE experiments.

CHAPTER IV

CONCLUSION

For our carotenoid work we sequenced the β-carotene hyperproducers from the ALE experiments, identified beneficial mutations for β-carotene production, and demonstrated that the most impactful beneficial mutations for β-carotene production increased production of the desired antioxidant compound. Finally, to demonstrate the broad utility of the findings from this work, we rationally constructed an improved lycopene producer using the best combinations of mutations identified for β-carotene. We have successfully identified additional genes important for increasing carotenoids yields in yeast, which cannot currently be determined using rational approaches, and we have demonstrated that these mutations confer benefits to the production of carotenoids other than β-carotene.

For our caryophyllene work, the antioxidant property of β-caryophyllene was leveraged against hydrogen peroxide stress. Exposing cells to selective pressure with a recovery period turned out to be beneficial in order to evolve the strains compared to continuous exposure in media. ALE experiment resulted in population showing 3-fold increment in the production and mutant screening resulted in a 4-fold production. We demonstrated that by leveraging antioxidant potential of terpenes exported extracellularly the production can be growth coupled and be improved during very short period with appropriate and optimized stress.

Overall, we showed that we can leverage the product specific properties to create a selective pressure scheme which can be used to improve production in very short amount of period for different class of terpene both intracellular and extracellular. We gained deeper understanding on how the point mutations arising during the evolution experiment are able to impact the production even when the genes mutated are not directly related to the terpene biosynthetic pathway. We also showed how these mutations can or cannot impact the productivity of other terpenes from the same class. These novel mutations can be used to create a chassis strain which can be used to improve the production of the class of terpenes.

REFERENCES

Almario MP, Reyes LH, Kao KC (2013) Evolutionary engineering of Saccharomyces cerevisiae for enhanced tolerance to hydrolysates of lignocellulosic biomass Biotechnol Bioeng 110:2616-2623 doi:10.1002/bit.24938

Alper H, Miyaoku K, Stephanopoulos G (2005) Construction of lycopene-overproducing *E. coli* strains by combining systematic and combinatorial gene knockout targets Nat Biotechnol 23:612-616 doi:10.1038/nbt1083

Alter TB, Ebert BE (2019) Determination of growth-coupling strategies and their underlying principles BMC Bioinformatics 20:447 doi:10.1186/s12859-019-2946-7

An GH, Jang BG, Cho MH (2001) Cultivation of the carotenoid-hyperproducing mutant 2A2N of the red yeast *Xanthophyllomyces dendrorhous* (*Phaffia rhodozyma*) with molasses J Biosci Bioeng 92:121-125 doi:10.1016/S1389-1723(01)80211-5

Arias P, Diez-Muniz S, Garcia R, Nombela C, Rodriguez-Pena JM, Arroyo J (2011) Genome-wide survey of yeast mutations leading to activation of the yeast cell integrity MAPK pathway: novel insights into diverse MAPK outcomes BMC Genomics 12:390 doi:10.1186/1471-2164-12-390

Arigoni D, Sagner S, Latzel C, Eisenreich W, Bacher A, Zenk MH (1997) Terpenoid biosynthesis from 1-deoxy-D-xylulose in higher plants by intramolecular skeletal rearrangement Proc Natl Acad Sci U S A 94:10600-10605 doi:10.1073/pnas.94.20.10600

Arlt H, Perz A, Ungermann C (2011) An Overexpression Screen in *Saccharomyces cerevisiae* Identifies Novel Genes that Affect Endocytic Protein Trafficking Traffic 12:1592-1603 doi:10.1111/j.1600-0854.2011.01252.x

Asadollahi MA, Maury J, Moller K, Nielsen KF, Schalk M, Clark A, Nielsen J (2008) Production of plant sesquiterpenes in *Saccharomyces cerevisiae*: effect of *ERG9* repression on sesquiterpene biosynthesis Biotechnol Bioeng 99:666-677 doi:10.1002/bit.21581

Bailey SF, Hinz A, Kassen R (2014) Adaptive synonymous mutations in an experimentally evolved *Pseudomonas fluorescens* population Nature Communications 5:4076 doi:10.1038/ncomms5076

Bao Z et al. (2015) Homology-integrated CRISPR-Cas (HI-CRISPR) system for one-step multigene disruption in Saccharomyces cerevisiae ACS synthetic biology 4:585-594 doi:10.1021/sb500255k

Basso TO et al. (2011) Engineering topology and kinetics of sucrose metabolism in Saccharomyces cerevisiae for improved ethanol yield Metab Eng 13:694-703 doi:10.1016/j.ymben.2011.09.005

Benni ML, Neigeborn L (1997) Identification of a new class of negative regulators affecting sporulation-specific gene expression in yeast Genetics 147:1351-1366

Bhosale P, Bernstein PS (2004) Beta-carotene production by *Flavobacterium multivorum* in the presence of inorganic salts and urea J Ind Microbiol Biotechnol 31:565-571 doi:10.1007/s10295-004-0187-9

Bhosale P, Gadre RV (2001) Production of beta-carotene by a *Rhodotorula glutinis* mutant in sea water medium Bioresour Technol 76:53-55 doi:10.1016/S0960-8524(00)00075-4

Bligh EG, Dyer WJ (1959) A rapid method of total lipid extraction and purification Canadian journal of biochemistry and physiology 37:911-917 doi:10.1139/o59-099

Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, Boeke JD (1998) Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications Yeast 14:115-132 doi:10.1002/(sici)1097-0061(19980130)14:2<115::Aid-yea204>3.0.Co;2-2

Breil C, Abert Vian M, Zemb T, Kunz W, Chemat F (2017) "Bligh and Dyer" and Folch Methods for Solid-Liquid-Liquid Extraction of Lipids from Microorganisms. Comprehension of Solvatation Mechanisms and towards Substitution with Alternative Solvents International journal of molecular sciences 18 doi:10.3390/ijms18040708

Burgard AP, Pharkya P, Maranas CD (2003) Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization Biotechnol Bioeng 84:647-657 doi:10.1002/bit.10803

Cadiere A, Ortiz-Julien A, Camarasa C, Dequin S (2011) Evolutionary engineered Saccharomyces cerevisiae wine yeast strains with increased in vivo flux through

the pentose phosphate pathway Metab Eng 13:263-271 doi:10.1016/j.ymben.2011.01.008

Calleja MA, Vieites JM, Montero-Melendez T, Torres MI, Faus MJ, Gil A, Suarez A (2013) The antioxidant effect of beta-caryophyllene protects rat liver from carbon tetrachloride-induced fibrosis by inhibiting hepatic stellate cell activation The British journal of nutrition 109:394-401 doi:10.1017/s0007114512001298

Chen D, Han Y, Gu Z (2006) Application of statistical methodology to the optimization of fermentative medium for carotenoids production by *Rhodobacter sphaeroides* Process Biochem 41:1773-1778 doi:10.1016/j.procbio.2006.03.023

Chen Y, Xiao W, Wang Y, Liu H, Li X, Yuan Y (2016) Lycopene overproduction in *Saccharomyces cerevisiae* through combining pathway engineering with host engineering Microbial cell factories 15:113 doi:10.1186/s12934-016-0509-4

Czamara K, Majzner K, Pacia MZ, Kochan K, Kaczor A, Baranska M (2015) Raman spectroscopy of lipids: a review Journal of Raman Spectroscopy 46:4-20 doi:10.1002/jrs.4607

Dahham SS, Tabana YM, Iqbal MA, Ahamed MB, Ezzat MO, Majid AS, Majid AM (2015) The Anticancer, Antioxidant and Antimicrobial Properties of the Sesquiterpene beta-Caryophyllene from the Essential Oil of Aquilaria crassna Molecules (Basel, Switzerland) 20:11808-11829 doi:10.3390/molecules200711808

de Kok S et al. (2012) Laboratory evolution of new lactate transporter genes in a jen1Delta mutant of Saccharomyces cerevisiae and their identification as ADY2

alleles by whole-genome resequencing and transcriptome analysis FEMS Yeast Res 12:359-374 doi:10.1111/j.1567-1364.2012.00787.x

de Ruijter JC, Jurgens G, Frey AD (2017) Screening for novel genes of *Saccharomyces cerevisiae* involved in recombinant antibody production FEMS Yeast Res 17 doi:10.1093/femsyr/fow104

Deatherage DE, Barrick JE (2014) Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq Methods in molecular biology (Clifton, NJ) 1151:165-188 doi:10.1007/978-1-4939-0554-6_12

Dhar R, Sagesser R, Weikert C, Yuan J, Wagner A (2011) Adaptation of Saccharomyces cerevisiae to saline stress through laboratory evolution J Evol Biol 24:1135-1153 doi:10.1111/j.1420-9101.2011.02249.x

Dimster-Denk D, Thorsness MK, Rine J (1994) Feedback regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase in Saccharomyces cerevisiae Molecular biology of the cell 5:655-665

Dohmen RJ, Willers I, Marques AJ (2007) Biting the hand that feeds: Rpn4-dependent feedback regulation of proteasome function Biochim Biophys Acta 1773:1599-1604 doi:10.1016/j.bbamcr.2007.05.015

Dragosits M, Mattanovich D (2013) Adaptive laboratory evolution – principles and applications for biotechnology Microbial cell factories 12:64 doi:10.1186/1475-2859-12-64

Ekkers DM, Branco dos Santos F, Mallon CA, Bruggeman F, van Doorn GS (2020) The omnistat: A flexible continuous-culture system for prolonged experimental evolution Methods in Ecology and Evolution 11:932-942 doi:10.1111/2041-210x.13403

Elena SF, Lenski RE (2003) Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation Nature Reviews Genetics 4:457-469 doi:10.1038/nrg1088

Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high throughput capability PloS one 3:e3647 doi:10.1371/journal.pone.0003647

EVANS CT, RATLEDGE C (1984) Effect of Nitrogen Source on Lipid Accumulation in Oleaginous Yeasts Microbiology 130:1693-1704 doi:10.1099/00221287-130-7-1693

Fakhry EM, El Maghraby DM (2015) Lipid accumulation in response to nitrogen limitation and variation of temperature in Nannochloropsis salina Botanical studies 56:6-6 doi:10.1186/s40529-015-0085-7

Feist AM, Zielinski DC, Orth JD, Schellenberger J, Herrgard MJ, Palsson BO (2010) Model-driven evaluation of the production potential for growth-coupled products of Escherichia coli Metab Eng 12:173-186 doi:10.1016/j.ymben.2009.10.003

Fidalgo M, Barrales RR, Ibeas JI, Jimenez J (2006) Adaptive evolution by mutations in the FLO11 gene Proceedings of the National Academy of Sciences of the United States of America 103:11228-11233 doi:10.1073/pnas.0601713103

Fletcher E, Feizi A, Bisschops MMM, Hallstrom BM, Khoomrung S, Siewers V, Nielsen J (2017) Evolutionary engineering reveals divergent paths when yeast is adapted to different acidic environments Metab Eng 39:19-28 doi:10.1016/j.ymben.2016.10.010

Flury I, Garza R, Shearer A, Rosen J, Cronin S, Hampton RY (2005) INSIG: a broadly conserved transmembrane chaperone for sterol-sensing domain proteins EMBO J 24:3917-3926 doi:10.1038/sj.emboj.7600855

Fong SS, Burgard AP, Herring CD, Knight EM, Blattner FR, Maranas CD, Palsson BO (2005) In silico design and adaptive evolution of Escherichia coli for production of lactic acid Biotechnol Bioeng 91:643-648 doi:10.1002/bit.20542

Fred W, Catherine D, L. R-HS (1995) Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C Yeast 11:53-55 doi:10.1002/yea.320110107

Gagneur J, Klamt S (2004) Computation of elementary modes: a unifying framework and the new binary approach BMC Bioinformatics 5:175 doi:10.1186/1471-2105-5-175

Garcia Sanchez R et al. (2010) Improved xylose and arabinose utilization by an industrial recombinant Saccharomyces cerevisiae strain using evolutionary engineering Biotechnol Biofuels 3:13 doi:10.1186/1754-6834-3-13

Germer S, Holland MJ, Higuchi R (2000) High-throughput SNP allele-frequency determination in pooled DNA samples by kinetic PCR Genome Res 10:258-266

Godara A, Kao KC (2020) Adaptive laboratory evolution for growth coupled microbial production World Journal of Microbiology and Biotechnology 36:175 doi:10.1007/s11274-020-02946-8

Godara A, Rodriguez MAG, Weatherston JD, Peabody GL, Wu H-J, Kao KC (2019) Beneficial mutations for carotenoid production identified from laboratory-evolved Saccharomyces cerevisiae Journal of Industrial Microbiology and Biotechnology 46:1793-1804 doi:10.1007/s10295-019-02241-y

González-Ramos D et al. (2016) A new laboratory evolution approach to select for constitutive acetic acid tolerance in Saccharomyces cerevisiae and identification of causal mutations Biotechnology for Biofuels 9:173 doi:10.1186/s13068-016-0583-1

Goossens KVY et al. (2015) Molecular Mechanism of Flocculation Self-Recognition in Yeast and Its Role in Mating and Survival mBio 6:e00427-00415 doi:10.1128/mBio.00427-15

Gresham D et al. (2008) The Repertoire and Dynamics of Evolutionary Adaptations to Controlled Nutrient-Limited Environments in Yeast PLoS Genet 4:e1000303 doi:10.1371/journal.pgen.1000303

Hadicke O, Klamt S (2011) Computing complex metabolic intervention strategies using constrained minimal cut sets Metab Eng 13:204-213 doi:10.1016/j.ymben.2010.12.004

Hampton RY, Rine J (1994) Regulated degradation of HMG-CoA reductase, an integral membrane protein of the endoplasmic reticulum, in yeast J Cell Biol 125:299-312

Higgins VJ, Alic N, Thorpe GW, Breitenbach M, Larsson V, Dawes IW (2002) Phenotypic analysis of gene deletant strains for sensitivity to oxidative stress Yeast 19:203-214 doi:10.1002/yea.811

Holwerda EK et al. (2020) Metabolic and evolutionary responses of Clostridium thermocellum to genetic interventions aimed at improving ethanol production Biotechnology for Biofuels 13:40 doi:10.1186/s13068-020-01680-5

Hüsnü K, Başer C, Demirci F (2007) Chemistry of Essential Oils. In: Berger RG (ed) Flavours and Fragrances: Chemistry, Bioprocessing and Sustainability. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 43-86. doi:10.1007/978-3-540-49339-6_4

Ignea C, Pontini M, Motawia MS, Maffei ME, Makris AM, Kampranis SC (2018) Synthesis of 11-carbon terpenoids in yeast using protein and metabolic engineering Nat Chem Biol 14:1090-1098 doi:10.1038/s41589-018-0166-5

Ingy IA, Wim JQ (2017) A Glimpse into the Biosynthesis of Terpenoids KnE Life Sciences 3 doi:10.18502/kls.v3i5.981

Jackson BE, Hart-Wells EA, Matsuda SPT (2003) Metabolic Engineering to Produce Sesquiterpenes in Yeast Organic Letters 5:1629-1632 doi:10.1021/ol034231x

Jakobsen AN, Aasen IM, Josefsen KD, Strom AR (2008) Accumulation of docosahexaenoic acid-rich lipid in thraustochytrid Aurantiochytrium sp. strain

T66: effects of N and P starvation and O2 limitation Appl Microbiol Biotechnol 80:297-306 doi:10.1007/s00253-008-1537-8

Jensen K, Broeken V, Hansen ASL, Sonnenschein N, Herrgård MJ (2019) OptCouple: Joint simulation of gene knockouts, insertions and medium modifications for prediction of growth-coupled strain designs Metabolic Engineering Communications 8:e00087 doi:https://doi.org/10.1016/j.mec.2019.e00087

Jeon Y-C, Cho C-W, Yun Y-S (2006) Combined effects of light intensity and acetate concentration on the growth of unicellular microalga *Haematococcus pluvialis* Enzyme Microb Technol 39:490-495 doi:10.1016/j.enzmictec.2005.12.021

Kahm M, Hasenbrink G, Lichtenberg-Fraté H, Ludwig J, Kschischo M (2010) grofit: Fitting Biological Growth Curves with R Journal of Statistical Software; Vol 1, Issue 7 (2010)

Kapitzky L et al. (2010) Cross-species chemogenomic profiling reveals evolutionarily conserved drug mode of action Mol Syst Biol 6:451 doi:10.1038/msb.2010.107

Kim H, Kim A, Cunningham KW (2012) Vacuolar H+-ATPase (V-ATPase) promotes vacuolar membrane permeabilization and nonapoptotic death in stressed yeast J Biol Chem 287:19029-19039 doi:10.1074/jbc.M112.363390

Kim J, Reed JL, Maravelias CT (2011) Large-scale bi-level strain design approaches and mixed-integer programming solution techniques PloS one 6:e24162 doi:10.1371/journal.pone.0024162

Kim JM, Vanguri S, Boeke JD, Gabriel A, Voytas DF (1998) Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by

the complete Saccharomyces cerevisiae genome sequence Genome Res 8:464-478

Klamt S, Gilles ED (2004) Minimal cut sets in biochemical reaction networks Bioinformatics 20:226-234 doi:10.1093/bioinformatics/btg395

Klamt S, Mahadevan R (2015) On the feasibility of growth-coupled product synthesis in microbial strains Metab Eng 30:166-178 doi:https://doi.org/10.1016/j.ymben.2015.05.006

Kolouchova I, Matatkova O, Sigler K, Masak J, Rezanka T (2016) Lipid accumulation by oleaginous and non-oleaginous yeast strains in nitrogen and phosphate limitation Folia Microbiol (Praha) 61:431-438 doi:10.1007/s12223-016-0454-y

Kot AM, Błażejak S, Kurcz A, Gientka I, Kieliszek M (2016) *Rhodotorula glutinis*-potential source of lipids, carotenoids, and enzymes for use in industries Appl Microbiol Biotechnol 100:6103-6117 doi:10.1007/s00253-016-7611-8

Kudla G, Murray AW, Tollervey D, Plotkin JB (2009) Coding-sequence determinants of gene expression in *Escherichia coli* Science (New York, NY) 324:255-258 doi:10.1126/science.1170160

LaCroix RA et al. (2015) Use of Adaptive Laboratory Evolution To Discover Key Mutations Enabling Rapid Growth of Escherichia coli K-12 MG1655 on Glucose Minimal Medium Appl Environ Microbiol 81:17 doi:10.1128/AEM.02246-14

Lange C, Mustafi N, Frunzke J, Kennerknecht N, Wessel M, Bott M, Wendisch VF (2012) Lrp of Corynebacterium glutamicum controls expression of the brnFE operon encoding the export system for L-methionine and branched-chain amino acids J Biotechnol 158:231-241 doi:10.1016/j.jbiotec.2011.06.003

Leavitt JM, Wagner JM, Tu CC, Tong A, Liu Y, Alper HS (2017) Biosensor-Enabled Directed Evolution to Improve Muconic Acid Production in Saccharomyces cerevisiae Biotechnology journal 12 doi:10.1002/biot.201600687

Lee ME, DeLoache WC, Cervantes B, Dueber JE (2015) A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly ACS synthetic biology 4:975-986 doi:10.1021/sb500366v

Lee S, Kim P (2020) Current Status and Applications of Adaptive Laboratory Evolution in Industrial Microorganisms Journal of microbiology and biotechnology 30:793-803 doi:10.4014/jmb.2003.03072

Little S (2001) Amplification-refractory mutation system (ARMS) analysis of point mutations Current protocols in human genetics Chapter 9:Unit 9.8 doi:10.1002/0471142905.hg0908s07

Lohning C, Ciriacy M (1994) The *TYE7* gene of *Saccharomyces cerevisiae* encodes a putative bHLH-LZ transcription factor required for Ty1-mediated gene expression Yeast 10:1329-1339 doi:10.1002/yea.320101010

Luo H et al. (2019) Coupling S-adenosylmethionine–dependent methylation to growth: Design and uses PLoS Biol 17:e2007050 doi:10.1371/journal.pbio.2007050

Luria SE, Delbruck M (1943) Mutations of Bacteria from Virus Sensitivity to Virus Resistance Genetics 28:491-511

Mahr R, Gatgens C, Gatgens J, Polen T, Kalinowski J, Frunzke J (2015a) Biosensor-driven adaptive laboratory evolution of l-valine production in Corynebacterium glutamicum Metab Eng 32:184-194 doi:10.1016/j.ymben.2015.09.017

Mahr R, Gätgens C, Gätgens J, Polen T, Kalinowski J, Frunzke J (2015b) Biosensor-driven adaptive laboratory evolution of l-valine production in Corynebacterium glutamicum Metab Eng 32:184-194 doi:https://doi.org/10.1016/j.ymben.2015.09.017

Malisorn C, Suntornsuk W (2008) Optimization of β-carotene production by *Rhodotorula glutinis* DM28 in fermented radish brine Bioresour Technol 99:2281-2287 doi:10.1016/j.biortech.2007.05.019

Mata-Gómez LC, Montañez JC, Méndez-Zavala A, Aguilar CN (2014) Biotechnological production of carotenoids by yeasts: an overview Microbial cell factories 13:12-12 doi:10.1186/1475-2859-13-12

Michaelis S (1993) STE6, the yeast a-factor transporter Semin Cell Biol 4:17-27 doi:https://doi.org/10.1006/scel.1993.1003

Molin M, Norbeck J, Blomberg A (2003) Dihydroxyacetone kinases in *Saccharomyces cerevisiae* are involved in detoxification of dihydroxyacetone J Biol Chem 278:1415-1423 doi:10.1074/jbc.M203030200

Morais ARC, Bogel-Lukasik R (2013) Green chemistry and the biorefinery concept Sustainable Chemical Processes 1:18 doi:10.1186/2043-7129-1-18

Mundhada H et al. (2017) Increased production of L-serine in Escherichia coli through Adaptive Laboratory Evolution Metab Eng 39:141-150 doi:10.1016/j.ymben.2016.11.008

Mustafi N, Grunberger A, Kohlheyer D, Bott M, Frunzke J (2012) The development and application of a single-cell biosensor for the detection of l-methionine and branched-chain amino acids Metab Eng 14:449-457 doi:10.1016/j.ymben.2012.02.002

Nanou K, Roukas T, Kotzekidou P (2007) Role of hydrolytic enzymes and oxidative stress in autolysis and morphology of *Blakeslea trispora* during beta-carotene production in submerged fermentation Appl Microbiol Biotechnol 74:447-453 doi:10.1007/s00253-006-0666-1

Nishizaki T, Tsuge K, Itaya M, Doi N, Yanagawa H (2007) Metabolic Engineering of Carotenoid Biosynthesis in *Escherichia coli* by Ordered Gene Assembly in *Bacillus subtilis* Appl Environ Microbiol 73:1355 doi:10.1128/AEM.02268-06

O'Brien EJ, Monk JM, Palsson BO (2015) Using Genome-scale Models to Predict Biological Capabilities Cell 161:971-987 doi:10.1016/j.cell.2015.05.019

Ohno S, Shimizu H, Furusawa C (2014) FastPros: screening of reaction knockout strategies for metabolic engineering Bioinformatics 30:981-987 doi:10.1093/bioinformatics/btt672

Oliveira R, Lages F, Silva-Graca M, Lucas C (2003) Fps1p channel is the mediator of the major part of glycerol passive diffusion in *Saccharomyces cerevisiae*:

artefacts and re-definitions Biochim Biophys Acta 1613:57-71 doi:10.1016/S0005-2736(03)00138-X

Olson ML, Johnson J, Carswell WF, Reyes LH, Senger RS, Kao KC (2016) Characterization of an evolved carotenoids hyper-producer of *Saccharomyces cerevisiae* through bioreactor parameter optimization and Raman spectroscopy J Ind Microbiol Biotechnol 43:1355-1363 doi:10.1007/s10295-016-1808-9

P. B, R.V. G (2001) Optimization of carotenoid production from hyper-producing *Rhodotorula glutinis* mutant 32 by a factorial approach Lett Appl Microbiol 33:12-16 doi:doi:10.1046/j.1472-765X.2001.00940.x

Phaneuf PV et al. (2020) Causal mutations from adaptive laboratory evolution are outlined by multiple scales of genome annotations and condition-specificity BMC Genomics 21:514-514 doi:10.1186/s12864-020-06920-4

Pharkya P, Burgard AP, Maranas CD (2004) OptStrain: a computational framework for redesign of microbial production systems Genome Res 14:2367-2376 doi:10.1101/gr.2872004

Portnoy VA, Bezdan D, Zengler K (2011) Adaptive laboratory evolution--harnessing the power of biology for metabolic engineering Curr Opin Biotechnol 22:590-594 doi:10.1016/j.copbio.2011.03.007

Prelich G (2012) Gene overexpression: uses, mechanisms, and interpretation Genetics 190:841-854 doi:10.1534/genetics.111.136911

Promdonkoy P, Mhuantong W, Champreda V, Tanapongpipat S, Runguphan W (2020) Improvement in d-xylose utilization and isobutanol production in S. cerevisiae by

adaptive laboratory evolution and rational engineering J Ind Microbiol Biotechnol 47:497-510 doi:10.1007/s10295-020-02281-9

Raja R, Hemaiswarya S, Rengasamy R (2007) Exploitation of Dunaliella for beta-carotene production Appl Microbiol Biotechnol 74:517-523 doi:10.1007/s00253-006-0777-8

Reinsvold RE, Jinkerson RE, Radakovits R, Posewitz MC, Basu C (2011) The production of the sesquiterpene beta-caryophyllene in a transgenic strain of the cyanobacterium Synechocystis J Plant Physiol 168:848-852 doi:10.1016/j.jplph.2010.11.006

Reiss G, te Heesen S, Zimmerman J, Robbins PW, Aebi M (1996) Isolation of the *ALG6* locus of *Saccharomyces cerevisiae* required for glucosylation in the N-linked glycosylation pathway Glycobiology 6:493-498

Reyes LH, Gomez JM, Kao KC (2014) Improving carotenoids production in yeast via adaptive laboratory evolution Metab Eng 21:26-33 doi:10.1016/j.ymben.2013.11.002

Richardson B, Orcutt DM, Schwertner HA, Martinez CL, Wickline HE (1969) Effects of nitrogen limitation on the growth and composition of unicellular algae in continuous culture Applied microbiology 18:245-250

Ro D-K et al. (2006) Production of the antimalarial drug precursor artemisinic acid in engineered yeast Nature 440:940 doi:10.1038/nature04640

https://www.nature.com/articles/nature04640#supplementary-information

Rodwell VW, Nordstrom JL, Mitschelen JJ (1976) Regulation of HMG-CoA reductase Advances in lipid research 14:1-74

Rugbjerg P, Feist AM, Sommer MOA (2018) Enhanced Metabolite Productivity of Escherichia coli Adapted to Glucose M9 Minimal Medium Frontiers in bioengineering and biotechnology 6 doi:10.3389/fbioe.2018.00166

Russo A, Perri M, Cione E, Di Gioia ML, Nardi M, Cristina Caroleo M (2017) Biochemical and chemical characterization of Cynara cardunculus L. extract and its potential use as co-adjuvant therapy of chronic myeloid leukemia J Ethnopharmacol 202:184-191 doi:10.1016/j.jep.2017.03.026

Saerens SM, Verstrepen KJ, Van Laere SD, Voet AR, Van Dijck P, Delvaux FR, Thevelein JM (2006) The Saccharomyces cerevisiae EHT1 and EEB1 genes encode novel enzymes with medium-chain fatty acid ethyl ester synthesis and hydrolysis capacity J Biol Chem 281:4446-4456 doi:10.1074/jbc.M512028200

Sandberg TE et al. (2014) Evolution of Escherichia coli to 42 degrees C and subsequent genetic engineering reveals adaptive mechanisms and novel mutations Mol Biol Evol 31:2647-2662 doi:10.1093/molbev/msu209

Sandmann T, Herrmann JM, Dengjel J, Schwarz H, Spang A (2003) Suppression of coatomer mutants by a new protein family with COPI and COPII binding motifs in Saccharomyces cerevisiae Molecular biology of the cell 14:3097-3113 doi:10.1091/mbc.e02-11-0736

Savinell JM, Palsson BO (1992) Network analysis of intermediary metabolism using linear optimization. I. Development of mathematical formalism J Theor Biol 154:421-454 doi:10.1016/s0022-5193(05)80161-4

Schuster S, Dandekar T, Fell DA (1999) Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering Trends Biotechnol 17:53-60 doi:10.1016/s0167-7799(98)01290-6

Sitepu IR et al. (2013) Manipulation of culture conditions alters lipid content and fatty acid profiles of a wide variety of known and new oleaginous yeast species Bioresour Technol 144:360-369 doi:10.1016/j.biortech.2013.06.047

Smukalla S et al. (2008) *FLO1* is a variable green beard gene that drives biofilm-like cooperation in budding yeast Cell 135:726-737 doi:10.1016/j.cell.2008.09.037

Sobel SG, Wolin SL (1999) Two yeast La motif-containing proteins are RNA-binding proteins that associate with polyribosomes Molecular biology of the cell 10:3849-3862 doi:10.1091/mbc.10.11.3849

Sousa JP et al. (2011) Validation of a gas chromatographic method to quantify sesquiterpenes in copaiba oils J Pharm Biomed Anal 54:653-659 doi:10.1016/j.jpba.2010.10.006

Stankunas K, Berger J, Ruse C, Sinclair DA, Randazzo F, Brock HW (1998) The enhancer of polycomb gene of Drosophila encodes a chromatin protein conserved in yeast and mammals Development 125:4055-4066

Stevenson LF, Kennedy BK, Harlow E (2001) A large-scale overexpression screen in *Saccharomyces cerevisiae* identifies previously uncharacterized cell cycle genes

Proceedings of the National Academy of Sciences 98:3946-3951 doi:10.1073/pnas.051013498

Sullivan DP, Georgiev A, Menon AK (2009) Tritium suicide selection identifies proteins involved in the uptake and intracellular transport of sterols in *Saccharomyces cerevisiae* Eukaryot Cell 8:161-169 doi:10.1128/ec.00135-08

Sun XM, Ren LJ, Ji XJ, Chen SL, Guo DS, Huang H (2016) Adaptive evolution of Schizochytrium sp. by continuous high oxygen stimulations to enhance docosahexaenoic acid synthesis Bioresour Technol 211:374-381 doi:10.1016/j.biortech.2016.03.093

Tokuyama K, Toya Y, Horinouchi T, Furusawa C, Matsuda F, Shimizu H (2018) Application of adaptive laboratory evolution to overcome a flux limitation in an Escherichia coli production strain Biotechnol Bioeng 115:1542-1551 doi:10.1002/bit.26568

Truan G, Epinat JC, Rougeulle C, Cullin C, Pompon D (1994) Cloning and characterization of a yeast cytochrome b5-encoding gene which suppresses ketoconazole hypersensitivity in a NADPH-P-450 reductase-deficient strain Gene 142:123-127

Tundis R et al. (2017) Assessment of antioxidant, antitumor and pro-apoptotic effects of Salvia fruticosa Mill. subsp. thomasii (Lacaita) Brullo, Guglielmo, Pavone & Terrasi (Lamiaceae) Food Chem Toxicol 106:155-164 doi:10.1016/j.fct.2017.05.040

Verwaal R, Wang J, Meijnen JP, Visser H, Sandmann G, van den Berg JA, van Ooyen AJ (2007) High-level production of beta-carotene in *Saccharomyces cerevisiae* by successive transformation with carotenogenic genes from *Xanthophyllomyces dendrorhous* Appl Environ Microbiol 73:4342-4350 doi:10.1128/aem.02759-06

von Kamp A, Klamt S (2014) Enumeration of smallest intervention strategies in genome-scale metabolic networks PLoS Comput Biol 10:e1003378 doi:10.1371/journal.pcbi.1003378

von Kamp A, Klamt S (2017) Growth-coupled overproduction is feasible for almost all metabolites in five major production organisms Nat Commun 8:15956 doi:10.1038/ncomms15956

Wallace-Salinas V, Brink DP, Ahrén D, Gorwa-Grauslund MF (2015) Cell periphery-related proteins as major genomic targets behind the adaptive evolution of an industrial Saccharomyces cerevisiae strain to combined heat and hydrolysate stress BMC Genomics 16:514 doi:10.1186/s12864-015-1737-4

Wang C, Liwei M, Park J-B, Jeong S-H, Wei G, Wang Y, Kim S-W (2018) Microbial Platform for Terpenoid Production: Escherichia coli and Yeast Frontiers in Microbiology 9 doi:10.3389/fmicb.2018.02460

Watanabe J, Uehara K, Mogi Y (2013) Adaptation of the osmotolerant yeast Zygosaccharomyces rouxii to an osmotic environment through copy number amplification of FLO11D Genetics 195:393-405 doi:10.1534/genetics.113.154690

Weatherston JD, Seguban RKO, Hunt D, Wu H-J (2018) Low-Cost and Simple Fabrication of Nanoplasmonic Paper for Coupled Chromatography Separation and Surface Enhanced Raman Detection ACS Sensors 3:852-857 doi:10.1021/acssensors.8b00098

Wilke CM, Adams J (1992) Fitness effects of Ty transposition in *Saccharomyces cerevisiae* Genetics 131:31-42

Wright J, Bellissimi E, de Hulster E, Wagner A, Pronk JT, van Maris AJ (2011) Batch and continuous culture-based selection strategies for acetic acid tolerance in xylose-fermenting Saccharomyces cerevisiae FEMS Yeast Res 11:299-306 doi:10.1111/j.1567-1364.2011.00719.x

Wu T et al. (2017) Membrane engineering - A novel strategy to enhance the production and accumulation of beta-carotene in *Escherichia coli* Metab Eng 43:85-91 doi:10.1016/j.ymben.2017.07.001

Wu W, Liu F, Davis RW (2018) Engineering Escherichia coli for the production of terpene mixture enriched in caryophyllene and caryophyllene alcohol as potential aviation fuel compounds Metabolic Engineering Communications 6:13-21 doi:https://doi.org/10.1016/j.meteno.2018.01.001

Xie W, Lv X, Ye L, Zhou P, Yu H (2015) Construction of lycopene-overproducing Saccharomyces cerevisiae by combining directed evolution and metabolic engineering Metab Eng 30:69-78 doi:10.1016/j.ymben.2015.04.009

Yang J, Guo L (2014) Biosynthesis of beta-carotene in engineered *E. coli* using the MEP and MVA pathways Microbial cell factories 13:160 doi:10.1186/s12934-014-0160-x

Yang J, Nie Q (2016) Engineering Escherichia coli to convert acetic acid to β-caryophyllene Microbial cell factories 15:74-74 doi:10.1186/s12934-016-0475-x

APPENDIX

SUPPLEMANTORY FIGURES AND TABLES

Table 7 List of strains used.

Strain	Genotype	Source
SM12	GSY1136 YIplac211YB/I/E* Δctt1	Reyes et al. 2014
SM13	GSY1136 YIplac211YB/I/E* Δctt1	Reyes et al. 2014
SM14	GSY1136 YIplac211YB/I/E* Δctt1	Reyes et al. 2014
YLH2	GSY1136 YIplac211YB/I/E* Δctt1	Reyes et al. 2014
FY2	Matα ura3-52	Fred et. Al 1995
YAG01	GSY1136 YIplac211YB/I/E* Δctt1, HIS7 389	This work
YAG02	GSY1136 YIplac211YB/I/E* Δctt1, SRO9/GFD2 int	This work
YAG03	GSY1136 YIplac211YB/I/E* Δctt1, TYE7 86	This work
YAG04	GSY1136 YIplac211YB/I/E* Δctt1, FLO1 925	This work
YAG05	GSY1136 YIplac211YB/I/E* Δctt1, DAK2/AQY3 int	This work
YAG06	GSY1136 YIplac211YB/I/E* Δctt1, SCY1 1836	This work
YAG07	GSY1136 YIplac211YB/I/E* Δctt1, EPL1 1754	This work
YAG08	GSY1136 YIplac211YB/I/E* Δctt1, ALG6 1411	This work
YAG09	GSY1136 YIplac211YB/I/E* Δctt1, MDS3 ins	This work
YAG10	GSY1136 YIplac211YB/I/E* Δctt1, YMRCTy1-3 1078	This work
YAG11	GSY1136 YIplac211YB/I/E* Δctt1, HIS7 389, SRO9/GFD2 int, TYE7 86	This work
YAG12	GSY1136 YIplac211YB/I/E* Δctt1, HIS7 389,SRO9/GFD2 int, FLO1 925, DAK2/AQY3 int, SCY1 1836	This work
YAG13	GSY1136 YIplac211YB/I/E* Δctt1, HIS7 389, ALG6 1411, EPL1 1754, MDS3 ins, YMRCTy1-3 1078	This work
YAG14	GSY1136 YIplac211YB/I/E* Δctt1, SRO9/GFD2 int, TYE7 86	This work

YAG15	GSY1136 YIplac211YB/I/E* Δctt1, SRO9/GFD2 int, FLO1 925, DAK2/AQY3 int, SCY1 1836	This work
YAG16	GSY1136 YIplac211YB/I/E* Δctt1, ALG6 1411, EPL1 1754, MDS3 ins, YMRCTy1-3 1078	This work
YAG17	GSY1136 YIplac211YB/I/E* Δctt1, ALG6 1411, EPL1 1754	This work
YAG18	GSY1136 YIplac211YB/I/E* Δctt1, MDS3 ins, ALG6 1411	This work
YAG19	GSY1136 YIplac211YB/I/E* Δctt1, MDS3 ins, EPL1 1754	This work
YAG20	GSY1136 YIplac211YB/I/E* Δctt1, EPL1 1754, DAK2/AQY3 int	This work
YAG21	GSY1136 YIplac211YB/I/E* Δctt1, DAK2/AQY3 int, MDS3 ins	This work
YAG22	GSY1136 YIplac211YB/I/E* Δctt1, ALG6 1411, DAK2/AQY3 int	This work
YAG23	GSY1136 YIplac211YB/I/E* Δctt1, ALG6 1411, DAK2/AQY3 int, EPL1 1754	This work
YAG24	GSY1136 YIplac211YB/I/E* Δctt1, MDS3 ins, EPL1 1754, ALG6 1411	This work
YAG25	GSY1136 YIplac211YB/I/E* Δctt1, DAK2/AQY3 int, MDS3 ins, ALG6 1411	This work
YAG26	GSY1136 YIplac211YB/I/E* Δctt1, ALG6 1411, DAK2/AQY3 int, EPL1 1754, SRO9/GFD2 int	This work
YAG27	GSY1136 YIplac211YB/I/E* Δctt1, ALG6 1411, DAK2/AQY3 int, EPL1 1754, YMRCTy1-3 1078	This work
YAG28	GSY1136 YIplac211YB/I/E* Δctt1, ALG6 1411, DAK2/AQY3 int, YMRCTy1-3 1078	This work
YAG29	GSY1136 YIplac211YB/I/E* Δctt1, ALG6 1411, DAK2/AQY3 int, SRO9/GFD2 int	This work
YAG30	GSY1136 YIplac211YB/I/E* Δctt1, SRO9/GFD2 int, YMRCTy1-3 1078	This work

YAG31	Matα ura3-52, HIS7 389	This work
YAG32	Matα ura3-52, SRO9/GFD2 int	This work
YAG33	Matα ura3-52, TYE7 86	This work
YAG34	Matα ura3-52, FLO1 925	This work
YAG35	Matα ura3-52, DAK2/AQY3 int	This work
YAG36	Matα ura3-52, SCY1 1836	This work
YAG37	Matα ura3-52, EPL1 1754	This work
YAG38	Matα ura3-52, ALG6 1411	This work
YAG39	Matα ura3-52, MDS3 ins	This work
YAG40	Matα ura3-52, YMRCTy1-3 1078	This work
YAG41	FY2::URA3 P_{TDH3}-CrtB-T_{TDH1}-P_{CCW12}-CrtI-T_{ENO2}-P_{PGK1}-CrtE-T_{PGK1}	This work
YAG42	FY2::URA3 PTDH3-CrtB-TTDH1-PCCW12-CrtI-TENO2-PPGK1-CrtE-TPGK1, ALG6 1411, DAK2/AQY3 int	This work
YAG43	FY2::URA3 PTDH3-CrtB-TTDH1-PCCW12-CrtI-TENO2-PPGK1-CrtE-TPGK1, ALG6 1411, DAK2/AQY3 int, EPL1 1754	This work
YAG44	FY2::URA3 PTDH3-CrtB-TTDH1-PCCW12-CrtI-TENO2-PPGK1-CrtE-TPGK1, ALG6 1411, DAK2/AQY3 int, YMRCTy1-3 1078	This work
YAG45	Matα ura3-52 YIplac211YB/I/E*	This work
YAG46	Matα ura3-52 YIplac211YB/I/E*, HIS7 389	This work
YAG47	Matα ura3-52 YIplac211YB/I/E*, SRO9/GFD2 int	This work
YAG48	Matα ura3-52 YIplac211YB/I/E*, TYE7 86	This work
YAG49	Matα ura3-52 YIplac211YB/I/E*, FLO1 925	This work
YAG50	Matα ura3-52 YIplac211YB/I/E*, DAK2/AQY3 int	This work
YAG51	Matα ura3-52 YIplac211YB/I/E*, SCY1 1836	This work
YAG52	Matα ura3-52 YIplac211YB/I/E*, EPL1 1754	This work
YAG53	Matα ura3-52 YIplac211YB/I/E*, ALG6 1411	This work

YAG54	Matα ura3-52 YIplac211YB/I/E*, MDS3 ins	This work
YAG55	Matα ura3-52 YIplac211YB/I/E*, YMRCTy1-3 1078	This work
YAG56	Matα ura3-52 YIplac211YB/I/E* Δctt1	This work
YAG57	Matα ura3-52 YIplac211YB/I/E* Δctt1, HIS7 389	This work
YAG58	Matα ura3-52 YIplac211YB/I/E* Δctt1, SRO9/GFD2 int	This work
YAG59	Matα ura3-52 YIplac211YB/I/E* Δctt1, TYE7 86	This work
YAG60	Matα ura3-52 YIplac211YB/I/E* Δctt1, FLO1 925	This work
YAG61	Matα ura3-52 YIplac211YB/I/E* Δctt1, DAK2/AQY3 int	This work
YAG62	Matα ura3-52 YIplac211YB/I/E* Δctt1, SCY1 1836	This work
YAG63	Matα ura3-52 YIplac211YB/I/E* Δctt1, EPL1 1754	This work
YAG64	Matα ura3-52 YIplac211YB/I/E* Δctt1, ALG6 1411	This work
YAG65	Matα ura3-52 YIplac211YB/I/E* Δctt1, MDS3 ins	This work
YAG66	Matα ura3-52 YIplac211YB/I/E* Δctt1, YMRCTy1-3 1078	This work
YAG67	FY2::URA3 p_{TDH3}-CrtE-t_{TDH1}-p_{CCW12}-CrtB-t_{ENO2}-p_{PGK1}-CrtI-t_{PGK1}	This work
YAG68	FY2::URA3 p_{TDH3}-CrtE-t_{TDH1}-p_{CCW12}-CrtI-t_{ENO2}-p_{PGK1}-CrtB-t_{PGK1}	This work
YAG69	FY2::URA3 p_{TDH3}-CrtB-t_{TDH1}-p_{CCW12}-CrtE-t_{ENO2}-p_{PGK1}-CrtI-t_{PGK1}	This work
YAG70	FY2::URA3 p_{TDH3}-CrtI-t_{TDH1}-p_{CCW12}-CrtE-t_{ENO2}-p_{PGK1}-CrtB-t_{PGK1}	This work
YAG71	Same as YAG41; FY2::URA3 p_{TDH3}-CrtB-t_{TDH1}-p_{CCW12}-CrtI-t_{ENO2}-p_{PGK1}-CrtE-t_{PGK1}	This work
YAG72	FY2::URA3 p_{TDH3}-CrtI-t_{TDH1}-p_{CCW12}-CrtB-t_{ENO2}-p_{PGK1}-CrtE-t_{PGK1}	This work

Table 8 Guide RNA (gRNA) sequences for site-directed mutagenesis using CRISPR/Cas9.

Gene	Donor sequence	Guide sequence
HIS7 389	CAGTACCAGAAATAGGTTGGAATTCTTGCATTCCCTTGGAAAACCTATTCTTTGGATTGGATCCATACAAGAGGTACTATTTCGTCCATTCTTTTGCTGC	CCCTCGGAAAACCTATTCTT
SRO9/GFD2 int	AATGGTGCACGGATGTACCCACGCAATGAAAAATTTTTTTCACTTCTGGTAGGTCAAGAACTAAGAAAGAAAAGTACAAGTGTAAACATTTCTTAACATC	AAATTTTTTTCACTTCTGGT
TYE7 86	TTAATTAAACCTGAATCTGAATTTGATAATTGGTTGTCGGCTGAAAATGACGGAGCTAGTCATATCAACGTCAACAAGGACTCCTCGTCAGTTCTTTCTG	TGGTTGTCGGATGAAAATGA
FLO1 925	ATGGTGCTAGCAGTTGTTGGAGTTCTGATGACAATGATGGTTTCGTCAGTTGGAACGCCGTTGGTACCGGTGACGGTGGTCATTTCAGTAGATGTAGAAG	ACAATGACGGTTTCGTCAGT
DAK2/AQY3 int	AGCGTTATACTGTGCGATTATACGCTTCTTTTTATATGAATAAGGGGGAGACATGGTGGAAAGGTACCAGAACTTTTGATCGACCAAGACTAGGTAAAGC	ATATGAATAAGGGGGAGACA
SCY1 1836	GGCAATTAACAAAATGTCTTCAGACATTCAAAAGCACCATATCGCCAAATTAGATGATAAAGTCAATGATATTGGCGAAGATGCCTTTCACAAAGTCATT	AAAGCACCATATCGCCAAAT
MDS3 ins	AAAGGAGCTTATTCTCTTTTTTGCAAAGGAAAAAAAAACAACGGTAGTTCTTTTAGAAATTTAAACGAGTGGATAAGTAAGGAGACGTTTTTCTTAGACTT	AAAACAACGGTAGTTCTTTT
EPL1 1754	AGTTTCAGATAGATAGATCTTTTTATTCTTCACATTTACCAGAATACTGGAAGGGGATATCTGATGACATCAGAATATATGATTCAAATGGCCGTTCGAG	ATTTACCAGAATACTTGAAG

ALG6 1411	TATCAGCAGCATCAATAGCGACTATAGAAGAAGACGCTTACTGCCATATAATGTGGTTTGGAAAAGTTTTATCATAGGAACGTATATTGCTATGGGCTTT	AAGCTTACTGCCATATAATG
YMRCTy1-3 1078	ATTCTATGTATGAAGAACAACAGGAATCAAAACGTAATAAATCTACTTATAGGAGAAGTCCGAGTGATGAGAAGAAAGACTCTCGCACCTATACGAATAC	AACGTAATAAATCTACTCAT

Table 9 Primer sequences.

Primer Name	**Target**	**Sequence**	**Usage**
yfl-dak _f _wt	intergenic between *AQY3* and *DAK2*	5'-AAGGGGGAGACATGATGA-3' (wild type)	PCR ARMS verification
yfl-dak _f _mt	intergenic between *AQY3* and *DAK2*	5'-AAGGGGGAGACATGATGG-3' (mutant)	PCR ARMS verification
yfl-dak _r	intergenic between *AQY3* and *DAK2*	5'-AAGGGCTGATCTGGGTTC-3' (reverse)	PCR ARMS verification
epl1_f_wt	*EPL1*	5'- CTTCACATTTACCAGAATGCTT-3' (wild type)	PCR ARMS verification
epl1_f_mt	*EPL1*	5'-CTTCACATTTACCAGAATGCTG-3' (mutant)	PCR ARMS verification
epl1_r	*EPL1*	5'-TCGTAGGACTTCGTACCC -3' (reverse)	PCR ARMS verification
alg6_f_wt	*ALG6*	5'-GCGACTATAGAAGAAGAA-3' (wild type)	PCR ARMS verification
alg6_f_mt	*ALG6*	5'-GCGACTATAGAAGAAGAC-3' (mutant)	PCR ARMS verification
alg6_r	*ALG6*	5'- AAGTCCTTCATGGATTTG -3' (reverse)	PCR ARMS verification

mds3_f_wt	*MDS3*	5'-GAGCTTATTCTCTTTTTTGCAAAGGAAAAAAAC -3' (wild type)	PCR ARMS verification
mds3_f_mt	*MDS3*	5'-GAGCTTATTCTCTTTTTTGCAAAGGAAAAAAAA-3' (mutant)	PCR ARMS verification
mds3_r	*MDS3*	5'- AGCAATCAACTCATATCCGC -3' (reverse)	PCR ARMS verification
tye7_f_wt	*TYE7*	5'-TTTGATAATTGGTTGTCGGA-3' (wild type)	PCR ARMS verification
tye7_f_mt	*TYE7*	5'-TTTGATAATTGGTTGTCGGC-3' (mutant)	PCR ARMS verification
tye7_r	*TYE7*	5'-TGGAAAAAGAGCAGATTCCT-3' (reverse)	PCR ARMS verification
ymr_f_wt	*YMRCTy1-3*	5'-CAAAACGTAATAAATCTACTC-3' (wild type)	PCR ARMS verification
ymr_f_mt	*YMRCTy1-3*	5'-CAAAACGTAATAAATCTACTT-3' (mutant)	PCR ARMS verification
ymr_r	*YMRCTy1-3*	5'-AGCGTTAATTGGTATATTTC-3' (reverse)	PCR ARMS verification
scy1_f_wt	*SCY1*	5'-ACAAAATGTCTTCAGACATC-3' (wild type)	PCR ARMS verification
scy1_f_mt	*SCY1*	5'-ACAAAATGTCTTCAGACATT-3' (mutant)	PCR ARMS verification
scy1_r	*SCY1*	5'-TCAATATTTTTAGCTGCAAC-3' (reverse)	PCR ARMS verification

sro-gfd_f_wt	intergenic between *SRO9* and *GFD2*	5’-CACTTGTACTTTTCTTTCTTT-3’ (wild type)	PCR ARMS verification
sro-gfd_f_mt	intergenic between *SRO9* and *GFD2*	5’-CACTTGTACTTTTCTTTCTTA-3’ (mutant)	PCR ARMS verification
sro-gfd_r	intergenic between *SRO9* and *GFD2*	5’-TTATCTCTCACCATTTTTTG-3’ (reverse)	PCR ARMS verification
his7_f_wt	*HIS7*	5’-TGGAATTCTTGCATTCCCTC-3’ (wild type)	PCR ARMS verification
his7_f_mt	*HIS7*	5’-TGGAATTCTTGCATTCCCTT-3’ (mutant)	PCR ARMS verification
his7_r	*HIS7*	5’-ATTAAGAGTTCCTTCTCTTCCG-3’ (reverse)	PCR ARMS verification
flo1_f_wt	*FLO1*	5’-CGTTCCAACTGACGAAACCG-3’ (wild type)	PCR ARMS verification
flo1_f_mt	*FLO1*	5’-CGTTCCAACTGACGAAACCA-3’ (mutant)	PCR ARMS verification
flo1_r	*FLO1*	5’-AGAAGTAGAGGTAAAAGTGTCGTTCC-3’ (reverse)	PCR ARMS verification
curing_f	pCRCT plasmid	5’-AAATACAGACCGCCACAGTA-3’ (forward)	Cas9 curing
curing_r	pCRCT plasmid	5’-TGACGTTCATGCTTCTTGTC-3’ (reverse)	Cas9 curing

act1_f	*ACT1*	5’-TGTTCTAGCGCTTGCACCAT-3’(forward)	qRT-PCR
act1_r	*ACT1*	5’-AGAAATCTCTCGAGCAATTGGGA-3’ (reverse)	qRT-PCR
nsg1_f	*NSG1*	5’-GAACCTGACATGGTGCCTGA-3’(forward)	qRT-PCR
nsg1_r	*NSG1*	5’-TGACCTGGTTTTTGGCTGGT-3’(reverse)	qRT-PCR
cyb5_f	*CYB5*	5’-TCGGTCATTCTGACGAAGCA-3’(forward)	qRT-PCR
cyb5_r	*CYB5*	5’-TTTTCCACAGAAACGCGCTC-3’(reverse)	qRT-PCR
eeb1_f	*EEB1*	5’-AGCAGCCATGCTAACGAACT-3’(forward)	qRT-PCR
eeb1_r	*EEB1*	5’-GGACCACCAATCATGAGCCA-3’(reverse)	qRT-PCR
hmg1_f	*HMG1*	5’-ATGTAACCCAAGCAGACCCG-3’(forward)	qRT-PCR
hmg1_r	*HMG1*	5’-CTGTAGAGGCGCTCAACCAA-3’ (reverse)	qRT-PCR
aqy3_f	*AQY3*	5’-CATCATCGAGACGAGGGTCG-3’(forward)	qRT-PCR
aqy3_r	*AQY3*	5'- CCAAGCTTTTTACCCGGTGC-3’ (reverse)	qRT-PCR

dak2_f	*DAK2*	5'-TGCAACCTTTTGTCGAAGCG-3'(forward)	qRT-PCR
dak2_r	*DAK2*	5'-CCCAACAAGGGCATCCATCT-3' (reverse)	qRT-PCR
nup116_f	*NUP116*	5'-TGACCTCATTAGGGGGCGTA-3'(forward)	qRT-PCR
nup116_r	*NUP116*	5'-AAAGCAGGTAATCCTGGCCC-3' (reverse)	qRT-PCR
ioc4_f	*IOC4*	5'-GGTCTTACGGCAAGCCAGAT-3'(forward)	qRT-PCR
ico4_r	*IOC4*	5'-CGCTTCTGACGCTGTCTACT-3' (reverse)	qRT-PCR

Table 10 Characteristic wavelengths for Raman spectroscopy and their corresponding species.

Raman Shift (cm^{-1})	**Vibration**	**Species**
1157	C-C stretch	β-carotene, lycopene
1449	CH_2 bend	fatty acids
1660	C=C stretch	unsaturated fatty acids
2935	C-H stretch	proteins, fatty acids

Table 11 qRT-PCR results for genes flanking up and downstream of identified intergenic or Ty1 mutations. Relative to YLH2.

	DAK2		*AQY3*	
Strain	2^-(ΔΔct)	p value	2^-(ΔΔct)	p value
SM13	0.953 ± 0.500	0.254	1.019 ± 0.555	0.177
YAG05	1.242 ± 0.394	0.876	0.952 ± 0.337	0.254
YAG28	0.942 ± 0.436	0.108	1.138 ± 0.438	0.163
	NUP116		*IOC4*	
Strain	2^-(ΔΔct)	p value	2^-(ΔΔct)	p value
SM14	1.537 ± 0.425	0.136	1.729 ± 0.480	0.215
YAG10	1.544 ± 0.462	0.146	1.769 ± 0.606	0.225
YAG28	1.057 ± 0.412	0.239	1.243 ± 0.450	0.876
	SRO9		*GFD2*	
Strain	2^-(ΔΔct)	p value	2^-(ΔΔct)	p value
SM12	1.606 ± 0.755	0.734	0.885 ± 0.378	0.487
SM13	0.805 ±0.440	0.318	1.245 ± 0.475	0.623
YAG02	1.333 ± 0.678	0.546	1.032 ± .239	0.387

Table 12 qRT-PCR results for relative expression levels of select genes involved in lipid biosynthesis, sterol biosynthetic pathway and oxidative stress response. Bold: p-value < 0.05 using a student t-test compared with YLH2.

Strain	CYB5	NSG1	EEB1	HMG1	Avg β-carotene yield
	2^-(ΔΔct)	2^-(ΔΔct)	2^-(ΔΔct)	2^-(ΔΔct)	
YLH2	1.000	1.000	1.000	1.000	9.83
SM14	**2.491 ± 0.499**	1.757 ± 0.369	**3.017 ± 0.613**	0.961 ± 0.286	19.04
YAG01	1.170 ± 0.415	1.012 ± 0.318	0.881 ± 0.267	**1.803 ± 0.523**	3.67
YAG02	**4.238 ± 2.766**	1.509 ± 0.775	2.066 ± 0.988	1.353 ± 0.613	17.91
YAG03	1.050 ± 0.563	**1.562 ± 0.391**	1.178 ± 0.307	1.223 ± 0.279	13.84
YAG04	**1.316 ± 0.263**	1.840 ± 0.693	1.251 ± 0.325	0.897 ± 0.209	13.60
YAG05	**1.544 ± 0.376**	**1.977 ± 0.563**	**2.707 ± 0.868**	1.165 ± 0.356	12.89
YAG06	**1.519 ± 0.400**	1.234 ± 0.282	1.173 ± 0.262	1.272 ± 0.489	10.07
YAG07	0.861 ± 0.184	1.200 ± 0.329	1.853 ± 0.608	**0.419 ± 0.116**	12.03
YAG08	0.853 ± 0.227	0.861 ± 0.243	**1.741 ± 0.469**	**0.349 ± 0.119**	13.61
YAG09	0.781 ± 0.271	**3.053 ± 1.160**	0.360 ± 0.176	2.235 ± 0.765	10.47
YAG10	**3.2574 ± 1.434**	**1.544 ± 0.509**	**1.941 ± 0.600**	0.614 ± 0.181	20.49
YAG28	**3.356 ± 1.427**	1.421 ± 0.611	2.579 ± 1.107	0.640 ± 0.263	24.87
Strain	GND1	HAP3	MSN4	MXR2	Avg β-carotene yield
	2^-(ΔΔct)	2^-(ΔΔct)	2^-(ΔΔct)	2^-(ΔΔct)	
YLH2	1.000	1.000	1.000	1.000	9.83
SM14	1.701 ± 0.548	0.950 ± 0.414	0.861 ± 0.157	2.037 ± 1.476	19.04
YAG01	1.471 ± 0.527	1.052 ± 0.389	**0.367 ± 0.105**	0.502 ± 0.451	3.67

YAG02	0.885 ± 0.389	1.133 ± 0.544	**0.248 ± 0.107**	1.636 ± 1.472	17.91
YAG03	**2.297 ± 0.714**	1.206 ± 0.267	0.532 ± 0.271	1.069 ± 0.772	13.84
YAG04	**2.664 ± 0.656**	1.243 ± 0.457	0.803 ± 0.187	1.125 ± 0.800	13.60
YAG05	**1.790 ± 0.233**	1.000 ± 0.140	**0.290 ± 0.037**	2.324 ± 1.485	12.89
YAG06	**2.888 ± 0.439**	1.471 ± 0.420	**0.683 ± 0.184**	1.323 ± 0.874	10.07
YAG07	**3.490 ± 1.224**	**1.753 ± 0.356**	0.918 ± 0.112	1.417 ± 0.929	12.03
YAG08	**3.523 ± 0.886**	1.376 ± 0.406	**1.686 ± 0.302**	1.454 ± 0.964	13.61
YAG09	**4.056 ± 1.190**	1.602 ± 0.920	0.642 ± 0.347	1.404 ± 1.008	10.47
YAG10	0.859 ± 0.284	**0.715 ± 0.209**	**0.196 ± 0.051**	1.369 ± 0.990	20.49
YAG28	1.231 ± 0.497	1.231 ± 0.497	0.543 ± 0.209	0.918 ± 1.123	24.87

Table 13 Effect of promoters on lycopene yield. Three biological replicates were used for analysis.

Strain	**lycopene (mg/g DCW)**	**Std. dev**
YAG67	21.701	0.309
YAG68	22.175	0.180
YAG69	15.430	0.420
YAG70	7.433	0.443
YAG71	30.243	0.337
YAG72	17.378	0.160

Table 14 Production of strain with just QHS1 gene (YAG111) and strain with QHS1 and FPP overproduction genes (YAG115) integrated into genome.

Strain	β-caryophyllene (mg/g DCW)	Standard deviation
YAG111	0.790466	0.089037
YAG115	3.091245	0.520148

Table 15 Growth kinetics for strains with continuous exposure to hydrogen peroxide. Cells were grown in 96 well plate in a microplate reader for 72 hr and growth curves were calculated using grofit v1.1.1, ~NG: No Growth observed

Strain		Control	25mM	50mM	75mM	100mM
BY4741	**μ**	0.095 ± 0.007	0.041 ± 0.001	0.054 ± 0.013	0.023 ± 0.013	NG
	Lag phase	4.798 ± 0.203	27.242 ± 0.123	31.296 ± 0.0912	27.584 ± 0.312	NG
	Max OD	1.101 ± 0.014	0.721 ± 0.005	0.850 ± 0.005	0.483 ± 0.008	NG
YAG110	**μ**	0.096 ± 0.010	NG	NG	NG	NG
	Lag phase	4.613 ± 1.137	NG	NG	NG	NG
	Max OD	1.127 ± 0.094	NG	NG	NG	NG
YAG114	**μ**	0.051 ± 0.002	0.021 ± 0.004	0.007 ± 0.001	NG	NG
	Lag phase	3.635 ± 0.320	12.353 ± 2.163	26.497 ± 3.320	NG	NG
	Max OD	0.475 ± 0.018	0.331 ± 0.034	0.155 ± 0.012	NG	NG
YAG111	**μ**	0.055 ± 0.002	0.041 ± 0.004	0.018 ± 0.003	0.016 ± 0.005	NG
	Lag phase	4.195 ± 0.110	14.823 ± 0.318	18.608 ± 1.281	33.678 ± 2.401	NG
	Max OD	0.706 ± 0.027	0.522 ± 0.039	0.299 ± 0.031	0.217 ± 0.036	NG
YAG115	**μ**	0.039 ± 0.001	0.028 ± 0.002	0.017 ± 0.001	0.011 ± 0.002	0.007 ± 0.000
	Lag phase	3.888 ± 0.147	11.465 ± 0.247	12.348 ± 0.543	12.812 ± 2.000	15.092 ± 1.513

	Max OD	0.482 ± 0.018	0.404 ± 0.020	0.288 ± 0.007	0.237 ± 0.020	0.196 ± 0.003

Table 16 List of mutations found, frequency of each mutation in corresponding population sample is also shown. Each of these mutations were further verified using Sanger sequencing. Not available: mutation was not confirmed in Sanger sequencing. Dash - not considered for sequencing

Mutant	Frequency in population	Chromosome	position	mutation	annotation	gene	Sanger confirmation
P8 M6	0	14	12986	$(T)_{11\rightarrow14}$	intergenic (-110/-281)	SNO2 ← / → SNZ2	-
	0.944	1	26973	A→G	F332F (TTT→TTC)	FLO9 ←	Confirmed
	0.949	1	26981	T→C	S330G (AGC→GGC)	FLO9 ←	Confirmed
	0.94	1	26983	T→G	N329T (AAC→ACC)	FLO9 ←	Confirmed
	0.71	1	27090	T→C	E293E (GAA→GAG)	FLO9 ←	-
	0.789	1	27105	A→G	T288T (ACT→ACC)	FLO9 ←	-
	0	13	87053	$(TAT)_{36\rightarrow34}$	intergenic (-314/+65)	PRE8 ← / ← RPM2	-
	0.133	7	128474	T→A	Y1259* (TAT→TAA)	MDS3 →	-
	0	9	265719	G→C	T702S (ACT→AGT)	SYG1 ←	-
	0	4	273653	$(T)_{24\rightarrow23}$	coding (1224/1224 nt)	QRI7 ←	-
	0	7	400571	$(T)_{21\rightarrow22}$	intergenic (+411/+300)	OLE1 → / ← ERV14	-
	0	4	403520	$(T)_{9\rightarrow10}$	intergenic (-229/+175)	MPS1 ← / ← MRX9	-
	0	7	404475	+G	intergenic (+84/-995)	MST27 → / → tR(UCU)G1	-
	0	12	468701	T→C	intergenic (-1832/-112)	RDN18-2 ← / → RDN5-2	-
	0	7	531875	$(TA)_{15\rightarrow17}$	intergenic (+194/+8)	tD(GUC)G1 → / ← THG1	-

	0	4	548565	$(T)_{19\rightarrow20}$	intergenic (-255/+197)	RPC11 ← / ← BAP3	-
	0.64	13	908174	2 bp →TC	coding (811-812/3423 nt)	YMR317W →	-
	0.655	13	908177	A→G	I272V (ATT →GTG)	YMR317W →	-
	0.672	13	908179	T→G	I272V (ATT →GTG)	YMR317W →	-
	0.632	13	908185	A→G	S274S (TCA →TCG)	YMR317W →	-
	0.68	13	908196	G→C	W278S (TGG →TCG)	YMR317W →	-
	0.652	13	908198	G→T	A279S (GCA →TCA)	YMR317W →	-
	0.613	13	908203	G→A	T280T (ACG →ACA)	YMR317W →	-
	0.662	13	908218	C→T	S285S (AGC →AGT)	YMR317W →	-
P11M1	0	8	2303	$(C)_{11\rightarrow12}$	intergenic (-406/+368)	YHL050C ← / ← YHL050C	Confimed
	0.357	1	12690	A→T	intergenic (+264/+673)	YAL064W-B → / ← TDA8	Confirmed
	0	14	12986	$(T)_{11\rightarrow14}$	intergenic (-110/-281)	SNO2 ← / → SNZ2	Confimed
	0.834	1	27105	A→G	T288T (ACT →ACC)	FLO9 ←	Confirmed
	0	11	43222	G→T	T1025N (ACC→AAC)	STE6 ←	Confimed
	0	13	87053	$(TAT)_{36\rightarrow34}$	intergenic (-314/+65)	PRE8 ← / ← RPM2	Not available
	0.738	7	128474	T→A	Y1259* (TAT →TAA)	MDS3 →	Confimed
	0	3	151555	+A	intergenic (-199/-1282)	tK(CUU)C ← / → MAK32	Confirmed
	0	9	241053	$(A)_{21\rightarrow22}$	intergenic (-345/+255)	RNR3 ← / ← FIS1	Confimed
	0	3	286312	C→T	Q2066* (CAA→TAA)	CDC39 →	Confirmed
	0	7	404475	+G	intergenic (+84/-995)	MST27 → / → tR(UCU)G1	Confimed

	0.863	8	475932	Δ1 bp	intergenic (-154/+67)	tV(CAC)H ← / ← KOG1	Not available
	0.488	7	530034	A→C	S257S (TCA→TCC)	MTL1 →	Confimed
	0.646	10	715141	A→G	T200T (ACT→ACC)	DAN4 ←	Confirmed
	0	2	754982	C→T	D709N (GAC→AAC)	RIF1 ←	Confimed
	1	13	908218	C→T	S285S (AGC→AGT)	YMR317W →	Not available
P11M5	0.52	8	1846	A→G	T231T (ACT→ACC)	YHL050C ←	-
	0	14	12986	$(T)_{11→14}$	intergenic (-110/-281)	SNO2 ← / → SNZ2	-
	0.786	1	27090	T→C	E293E (GAA→GAG)	FLO9 ←	-
	0.834	1	27105	A→G	T288T (ACT→ACC)	FLO9 ←	Confirmed
	0	13	87053	$(TAT)_{36→34}$	intergenic (-314/+65)	PRE8 ← / ← RPM2	-
	0.199	7	128053	C→A	S1119* (TCG→TAG)	MDS3 →	-
	0	7	400571	$(T)_{21→22}$	intergenic (+411/+300)	OLE1 → / ← ERV14	-
	0	4	403520	$(T)_{9→10}$	intergenic (-229/+175)	MPS1 ← / ← MRX9	-
	0.488	7	530034	A→C	S257S (TCA→TCC)	MTL1 →	-
	1	13	908174	2 bp →TC	coding (811-812/3423 nt)	YMR317W →	Not available
	1	13	908177	A→G	I272V (ATT→GTG)	YMR317W →	Not available
	1	13	908179	T→G	I272V (ATT→GTG)	YMR317W →	Not available
	1	13	908185	A→G	S274S (TCA→TCG)	YMR317W →	Not available

	1	13	908196	G→C	W278S (TGG→TCG)	YMR317W →	Not available
	1	13	908198	G→T	A279S (GCA→TCA)	YMR317W →	Not available
	1	13	908203	G→A	T280T (ACG→ACA)	YMR317W →	Not available
	1	13	908218	C→T	S285S (AGC→AGT)	YMR317W →	Not available
	0.127	4	958000	C→A	E114* (GAA→TAA)	YDR248C ←	-
P14 M2	0.295	8	1846	A→G	T231T (ACT→ACC)	YHL050C ←	-
	0	1	6755	$(A)_{19→20}$	intergenic (+4048/+480)	YAL067W-A → / ← SEO1	-
	1	1	26981	T→C	S330G (AGC→GGC)	FLO9 ←	Confirmed
	1	1	26983	T→G	N329T (AAC→ACC)	FLO9 ←	Confirmed
	1	1	27105	A→G	T288T (ACT→ACC)	FLO9 ←	Confirmed
	1	2	34298	G→A	E1455K (GAG→AAG)	YBL100W-B →	Confirmed
	0	13	87044	$(TAT)_{33→31}$	intergenic (-305/+74)	PRE8 ← / ← RPM2	-
	0	13	87050	$(TAT)_{36→33}$	intergenic (-311/+65)	PRE8 ← / ← RPM2	-
	0	13	88387	A→C	L782W (TTG→TGG)	RPM2 ←	-
	1	7	128474	T→A	Y1259* (TAT→TAA)	MDS3 →	Confirmed
	0	7	397854	$(A)_{20→19}$	intergenic (-236/-774)	SDS23 ← / → OLE1	-
	0	4	403520	$(T)_{9→10}$	intergenic (-229/+175)	MPS1 ← / ← MRX9	-
	0	7	404475	+G	intergenic (+84/-995)	MST27 → / → tR(UCU)G1	-
	0	12	468701	T→C	intergenic (-1832/-112)	RDN18-2 ← / → RDN5-2	-

	0	10	539401	A→C	intergenic (+135/-2107)	HIT1 → / → tD(GUC)J4	-
	0	12	612381	T→G	S5R (AGT→AGG)	THI7 →	-
	0	4	930049	T→C	E103E (GAA→GAG)	RTN1 ←	-
	0	4	1489588	A→T	intergenic (+96/+10)	YDR524W-C → / ← YDR524C-B	-
P18M3	1	15	92	Δ3 bp	intergenic (–/-491)	– / → YOL166W-A	Not available
	0.712	8	1846	A→G	T231T (ACT→ACC)	YHL050C ←	-
	0	14	12986	$(T)_{11\rightarrow14}$	intergenic (-110/-281)	SNO2 ← / → SNZ2	-
	0.92	1	26973	A→G	F332F (TTT→TTC)	FLO9 ←	Confirmed
	0.917	1	26981	T→C	S330G (AGC→GGC)	FLO9 ←	Confirmed
	0.921	1	26983	T→G	N329T (AAC→ACC)	FLO9 ←	Confirmed
	0.714	1	27090	T→C	E293E (GAA→GAG)	FLO9 ←	-
	0.82	1	27105	A→G	T288T (ACT→ACC)	FLO9 ←	Confirmed
	0	4	37075	C→A	G93G (GGC→GGA)	MFG1 →	-
	0	13	87053	$(TAT)_{36\rightarrow34}$	intergenic (-314/+65)	PRE8 ← / ← RPM2	-
	0.804	7	128474	T→A	Y1259* (TAT→TAA)	MDS3 →	Confirmed
	0	3	231024	$(T)_{19\rightarrow20}$	intergenic (+20/-476)	HCM1 → / → RAD18	-
	0	4	403520	$(T)_{9\rightarrow10}$	intergenic (-229/+175)	MPS1 ← / ← MRX9	-
	0	7	404475	+G	intergenic (+84/-995)	MST27 → / → tR(UCU)G1	-
	0.796	12	468701	T→C	intergenic (-1832/-112)	RDN18-2 ← / → RDN5-2	-
	0	4	688076	$(T)_{14\rightarrow15}$	intergenic (+236/-151)	APC4 → / → VBA4	-

	0	2	723678	$(A)_{15\rightarrow16}$	intergenic (+43/+58)	SRB6 → / ← TRS20	-
	0.665	13	908196	G→C	W278S (TGG→TCG)	YMR317W →	-
	0.67	13	908218	C→T	S285S (AGC→AGT)	YMR317W →	-

Table 17 List of strain for chapter 3.

Strain	Genotype
BY4741	Mata his3Δ1 leu2Δ0 met15Δ0 ura3Δ0
pQHS1CEN	pTDH3-QHS1-tTDH1-URA3 (CEN plasmid)
pQHS12m	pTDH3-QHS1-tTDH1-URA3 (2μ plasmid)
pFPPCEN	pTHD3-tHMG1-tTDH1-pCCW12-HMG2(K6R)-tENO2-pPGK1-UPC2-1-tPGK1-pHHF2-ERG20-tADH3 -LEU2(CEN Plasmid)
pFPP2m	pTHD3-tHMG1-tTDH1-pCCW12-HMG2(K6R)-tENO2-pPGK1-UPC2-1-tPGK1-pHHF2-ERG20-tADH3 -LEU2(2μ Plasmid)
pQHS12mH	pTDH3-QHS1-tTDH1-HIS3 (2μ plasmid)
YAG101	BY4741 ura3::p_{TDH3}-QHS1-t_{TDH1}
YAG102	BY4741 /pQHS1CEN
YAG103	BY4741 /pQHS12m
YAG104	BY4741 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}
YAG105	BY4741ura3:: p_{TDH3}-QHS1-t_{TDH1} leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH2}
YAG106	BY4741 ura3:: p_{TDH3}-QHS1-t_{TDH1} /pFPPCEN
YAG107	BY4741 ura3:: p_{TDH3}-QHS1-t_{TDH1} /pFPP2m
YAG108	BY4741 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH3} /pQHS12m
YAG109	BY4741 /pQHS12m /pFPP2m
YAG110	BY4741 Δctt1
YAG111	BY4741 Δctt1 ura3:: p_{TDH3}-QHS1-t_{TDH1}
YAG112	BY4741 Δctt1 /pQHS1CEN

YAG113	BY4741 Δctt1 /pQHS12m
YAG114	BY4741 Δctt1 leu2::p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}
YAG115	BY4741 Δctt1 ura3::p_{TDH3}-QHS1-t_{TDH1} leu2::p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH2}
YAG116	YAG115 /pQHS12mH
YAG117	Population 11 mutant 1 /pQHS12mH
YAG118	BY4741 Δctt1 his3:: p_{TDH3}-QHS1-t_{TDH1} ::LEU2 p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH2}
YAG119	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, CDC39
YAG120	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, DAN4 200
YAG121	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, FLO9 288
YAG122	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, tK(CUU)C/MAK32 int
YAG123	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, MDS3
YAG124	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, MST27/tR(UCU)G1 int
YAG125	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, MTL1 257
YAG126	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, RIF1 709
YAG127	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, RNR3/FIS1 int
YAG128	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, SNO2/SNZ2 int
YAG129	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, STE6 1025
YAG130	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, YAL064W-B/TDA8 int
YAG131	BY4741 Δctt1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, YHL050C int
YAG132	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}, CDC39

YAG133	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ DAN4 200
YAG134	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ FLO9 288
YAG135	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ tK(CUU)C/MAK32 int
YAG136	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ MDS3
YAG137	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ MST27/tR(UCU)G1 int
YAG138	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ MTL1 257
YAG139	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ RIF1 709
YAG140	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ RNR3/FIS1 int
YAG141	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ SNO2/SNZ2 int
YAG142	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ STE6 1025
YAG143	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ YAL064W-B/TDA8 int
YAG144	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ YHL050C int
YAG145	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ STE6 1025, MST27/tR(UCU)G1 int
YAG146	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ Δste6
YAG147	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ (STE6 on 2 micron URA3)
YAG148	BY4741 Δctt1 his3:: pTDH3-QHS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ (STE6 1025 on 2 micron URA3)

YAG149	BY4741 Δctt1 his3:: pTDH3-ZSS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-t_{ADH1}
YAG150	BY4741 Δctt1 his3:: pTDH3-ZSS1-tTDH1 leu2:: p_{THD3}-tHMG1-t_{TDH1}-p_{CCW12}-HMG2(K6R)-t_{ENO2}-p_{PGK1}-UPC2-1-t_{PGK1}-p_{HHF2}-ERG20-$t_{ADH1,}$ (STE6 1025 on 2 micron URA3)

Table 18 Donor sequence for mutations

Mutation	Donor sequence
CDC39	CTTTGGTCTCACCAAAACCTAATTGTTGTAAAATTAGTTTAATTTCGGGAACACTTTATACAAAGGGCAGGTCTAAAAGATTAATATCGTTATTGTTTATCAACTGAGTAAAGAAAACGAACACTTTGTACAAAGGGCGTTTTAGAGAGAGACCTTTC
DAN4200	CTTTGGTCTCACCAAAACGTGGTAGAAGTAGTAGAGGTTGTAGGAGTAGTCGATGTGGTAGAGGTTGTAGGAGTGGTAGAAGTAGTAGAGGTTGTAGGAGTAGTCGATGTGGTAGAAGGTCGATGTGGTAGAAGTTGTGTTTTAGAGAGAGACCTTTC
FLO9288	CTTTGGTCTCACCAAAACGTTGGTACCGGTGACGGTGGTCATTTCAGTAGATGTAGAGGTGAAAGTACCGGTCCATGGTTCCGTTGTAGTTATGGTAGTACTGACAGTATAATTTGAAAGATGTAGAAGTGAAAGTACGTTTTAGAGAGAGACCTTTC
tK(CUU)C/MAK32int	CTTTGGTCTCACCAAAACCCGTATATGATAATATATTGATAATATAACTATTAGTTGATAGACGATAGTGGATTTTTATTCCAACAATTCTATATACGTAAAATTATAGCCTTTACCACTATAGTTGATAGACGATAGGTTTTAGAGAGAGACCTTTC
MDS3	CTTTGGTCTCACCAAAACTGTTAGTGCTACTGCTGCTGATGGAACCTTAATTCATATCAAACAAATGAAGATTAAAACTTCCGATGTTGGTTTGTTGATCATCGTTATCATCAAAACTAATTCATATCAAACAAATGAGTTTTAGAGAGAGACCTTTC

MST27/tR(UCU)G1int	CTTTGGTCTCACCAAAACAAATGGGTCCAACACGAATCGACTTTTCGAGGCTTCCTTCGGCCGTTTTCGGGCCAGTTAGTGCTGATTATATATCATACTCTAGTTTATGTTCGCTTTAAGGCTTCTTCGGCCGTTTTCGTTTTAGAGAGAGACCTTTC
MTL1257	CTTTGGTCTCACCAAAACTGTGGAGAGGGTGAAATATGATGATGATGATGATGATGAGGATGAGGATGAGGATGATGATGAGGATGATGAGGATGAGGATGATGAAAATGAGGAGGAGTGATGATGATGATGAGGATGGTTTTAGAGAGAGACCTTTC
RIF1709	CTTTGGTCTCACCAAAACGATGATCTGCGGGGGCAGGCTTGAAATGTTTTTTAACTTTACACCTTCTGAAGCTATAACACTCATTGGGTGAATATGCTTTTTATATTTTCTTTCCAATTTTTAACTTTACACCTTCTGGTTTTAGAGAGAGACCTTTC
RNR3/FIS1int	CTTTGGTCTCACCAAAACCAACATTGCGTGCCGTTGTTCTTTTGTTTTTTTTTTTTTTTTTTTTTCGTTGTTGTCGCAGCAACGACACCTAGGCGCTGCTCAAAGGGGCAAAAACCCTTTTTTTTTTTTTTTTTCGGTTTTAGAGAGAGACCTTTC
SNO2/SNZ2int	CTTTGGTCTCACCAAAACCAGCAACCGGGGTCATGGTAAGTGTGCTTGCTACTCAAACAGAAAAAAATGGTCTTTTATCAATGAATAACTTTTTTTTTTTGTAGCTGAATGATGACTACTCAAACAGAAAAAAATGTTTTAGAGAGAGACCTTTC
STE61025	CTTTGGTCTCACCAAAACTGCTGCCAGTTGGATCTATAGGATTCTTGATGAAAAGCATAATAACCTAGAGGTTGAAAACAATAATGCTAGAACAGTGGGAATAGCTGGTCACACCTACTGAAAAGCATAATACCCTAGGTTTTAGAGAGAGACCTTTC
YAL064W-B/	CTTTGGTCTCACCAAAACGGTGGTAATGATGAAGTAATTTCCTGACTTGTTGTTGTACTGGTAACAGGGGGTAATGATGAAGTAATTTCCTGACTTGTTGTTGCACTGGTAACAGGTGGTTGTTGTACTGGTAACAGGGTTTTAGAGAGAGACCTTTC

TDA8 int	
YHL050C int	CTTTGGTCTCACCAAAACAAGAGATGGGGGGGGGGGGCTCTGTTATCTATTATCTAGAAA AAACAGTCGGGCCGCAAGGAATCGTAAGGGTGAATTGCCACCAATTAAGGAAGGCTGTA TTATCTAGAAAAAACAGTTGTTTTAGAGAGAGACCTTTC

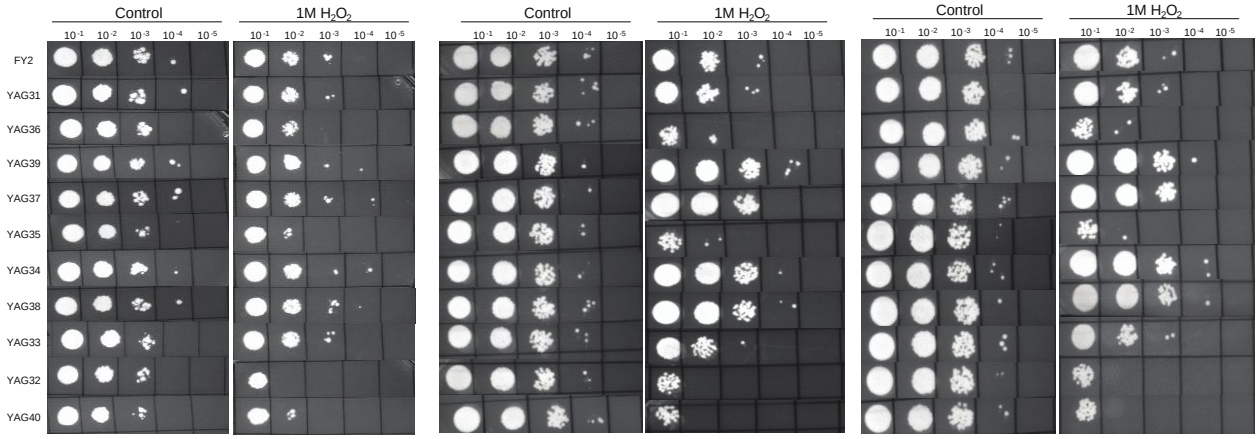


Figure 15 Oxidative stress tolerance of reconstructed single mutants in FY2 subjected to 1-hour exposure in 1 M H_2O_2. Strains are listed in increasing order of β-carotene observed in their YLH2 counterparts.

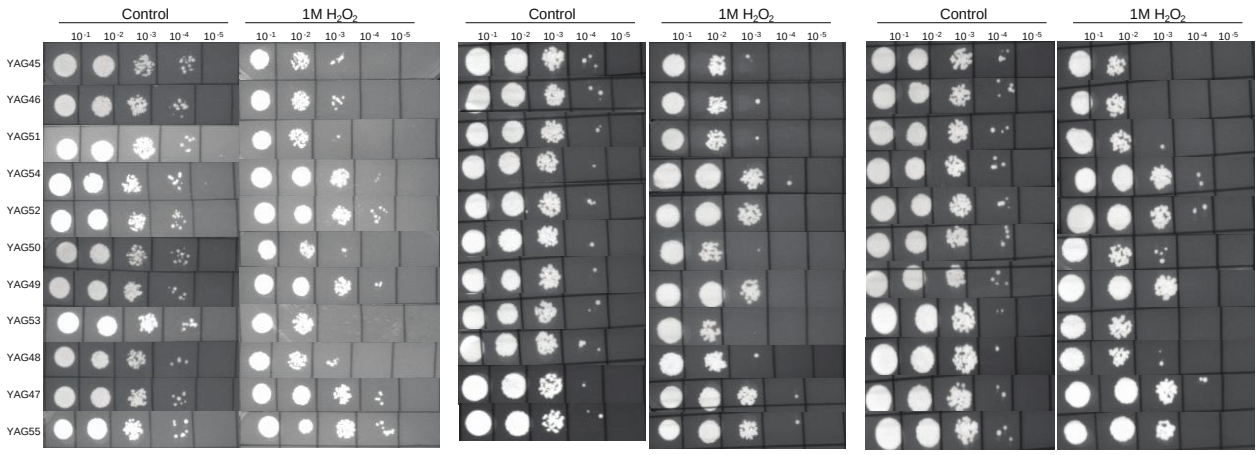


Figure 16 Oxidative stress tolerance of reconstructed single mutants in carotenogenic strain of FY2 subjected to 1-hour exposure in 1 M H2O2. Strains are listed in increasing order of β-carotene production.

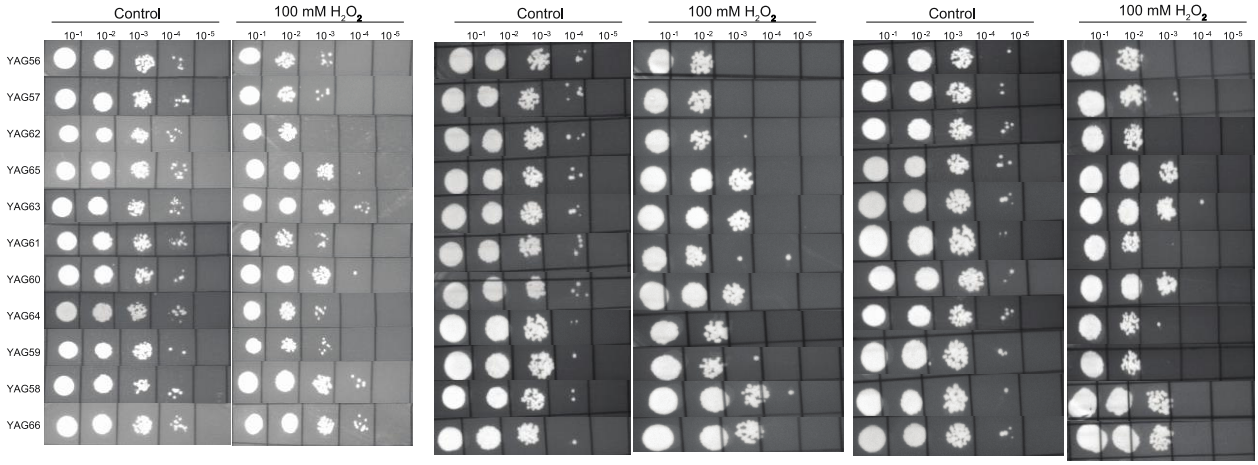


Figure 17 Oxidative stress tolerance of reconstructed single mutants in carotenogenic FY2 Δ*ctt1* strain subjected to 30 min exposure in 100 mM H_2O_2. Strains are listed in increasing order of β-carotene production.

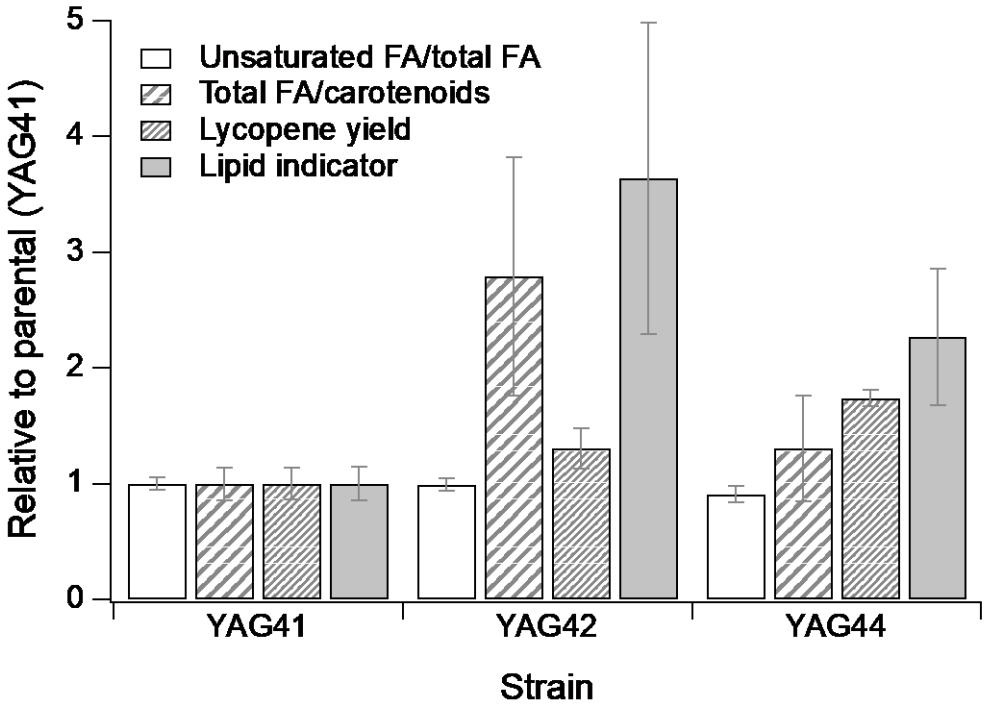
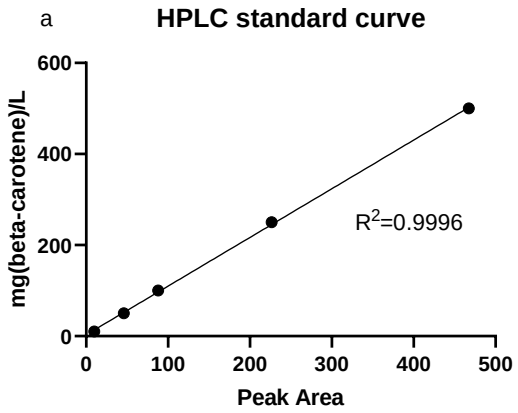
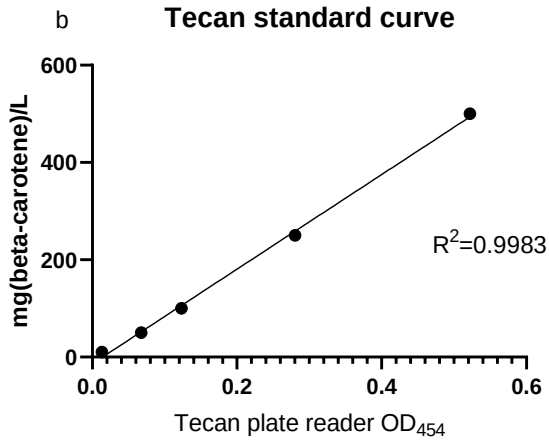


Figure 18 Raman results for lycopene producers. Relative peak intensity ratios compared with YAG41. FA: fatty acid. Error bars represent confidence interval about the sample mean with p = 0.90.

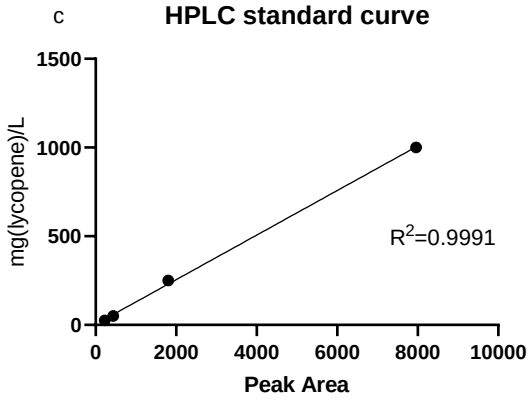
a


HPLC standard curve
mg(beta-carotene)/L
600
400
200
0
0
100
200
300
400
500
Peak Area
R2=0.9996

b


Tecan standard curve
mg(beta-carotene)/L
600
400
200
0
0.0
0.2
0.4
0.6
Tecan plate reader OD454
R2=0.9983

c


HPLC standard curve
mg(lycopene)/L
1500
1000
500
0
0
2000
4000
6000
8000
10000
Peak Area
R2=0.9991

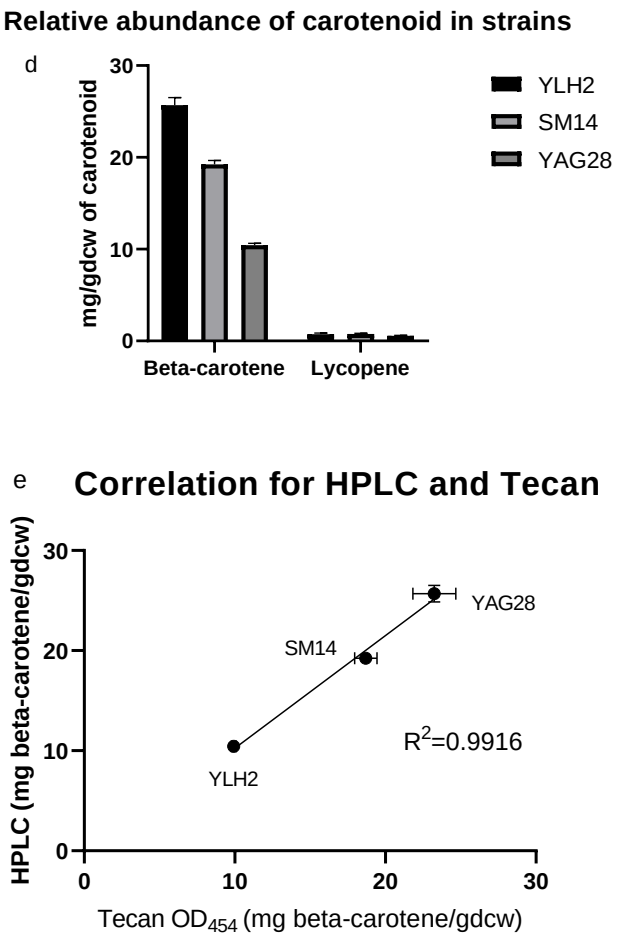


Figure 19 Correlation between β -carotene quantification using HPLC and using absorbance at OD454. a) and c) HPLC calibration curve using pure compounds for β -carotene and lycopene, respectively; peak area under the curve is shown. b) Standard curve for β -carotene using absorbance measurements at OD454. d) Amount of lycopene and β -carotene in strain YLH2, SM14 and YAG28 based on HPLC quantification. e) Correlation between HPLC and absorbance-based quantification of β-carotene for strains tested. Error bars are standard deviations based on 3 biological replicates per strain.

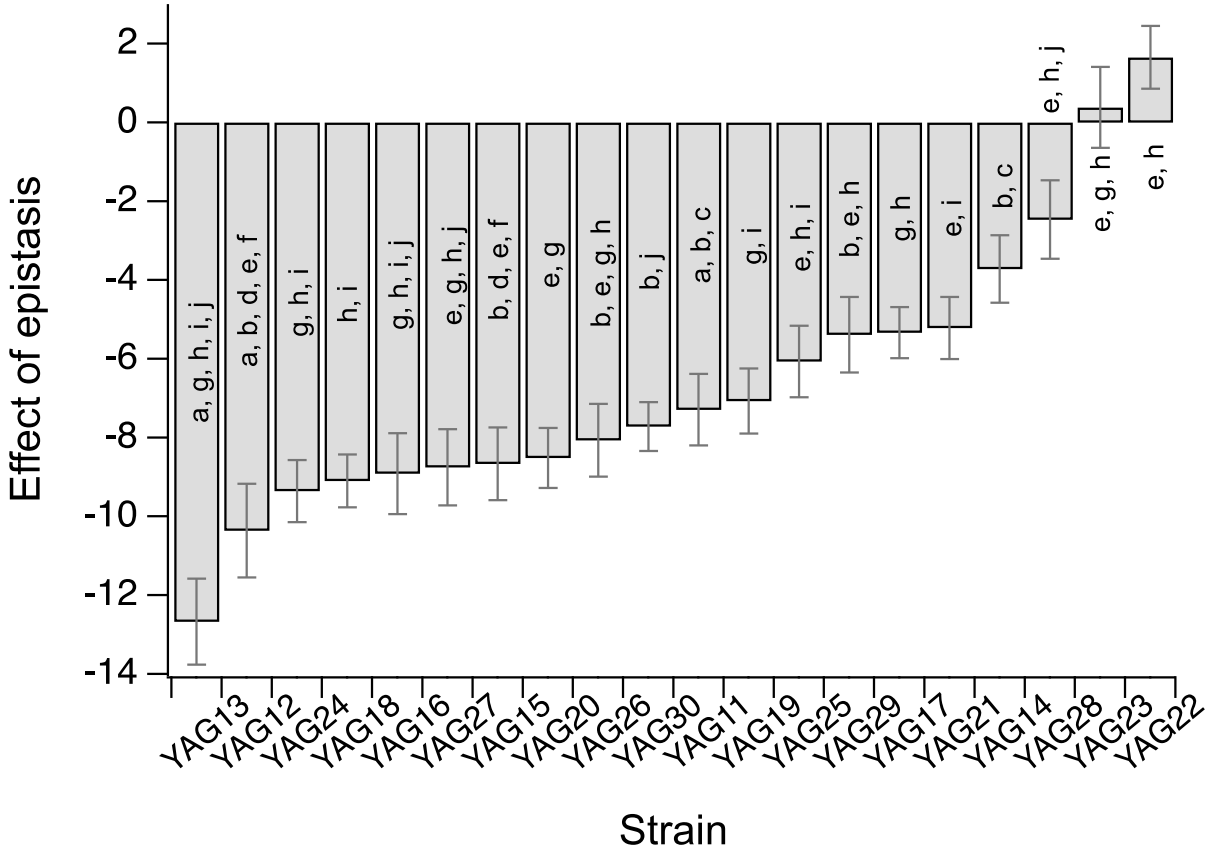
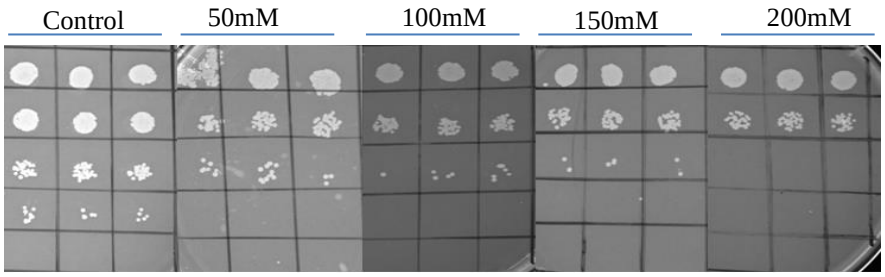
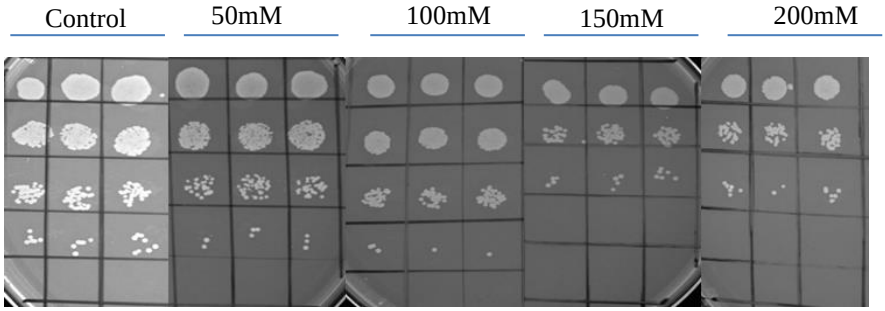


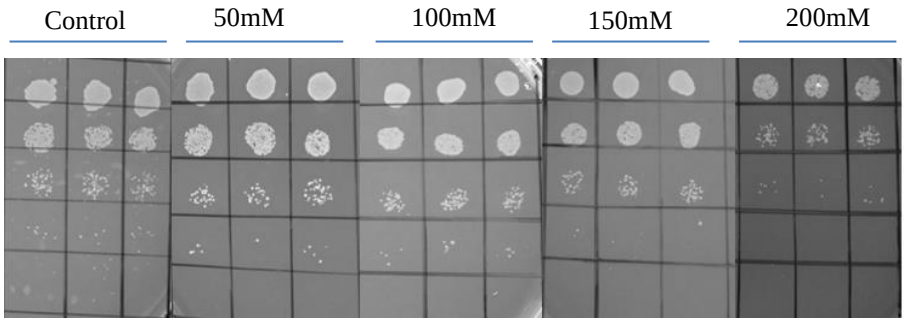
Figure 20 Epistatic interactions between different mutations on β-carotene production. Mutations: (a) HIS7 389, (b) SRO9/GFD2 int, (c) TYE7 86, (d) FLO1 925, (e) DAK2/AQY3 int, (f) SCY1 1836, (g) EPL1 1754, (h) ALG6 1411, (i) MDS3 ins, (j) YMRCTy1-3 1078.



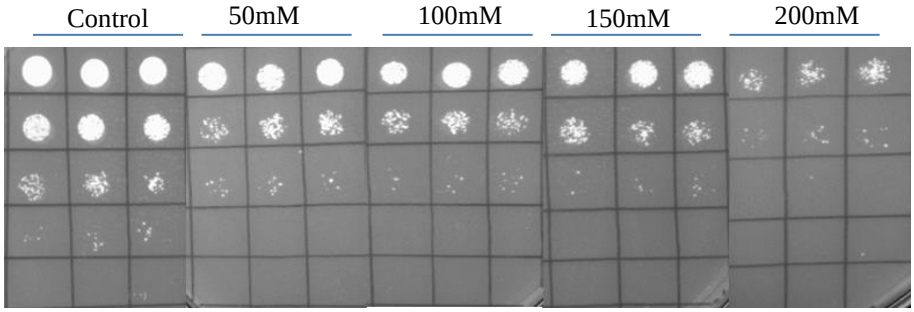
YAG110



YAG111



YAG115



YAG114

Figure 21 Antioxidant potential of beta-caryophyllene, 3 biologic replicates for each strain exposed to various hydrogen peroxide concentration and spot assay were done with different dilutions with increment as 10x dilutions.

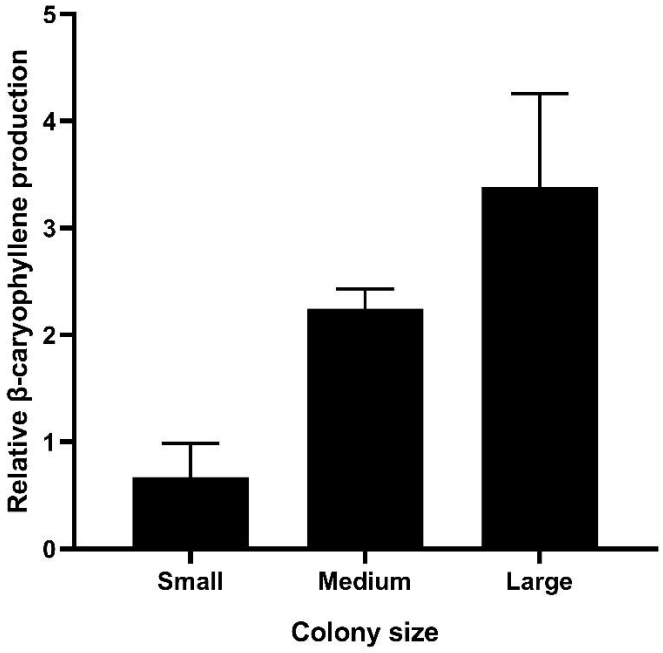


Figure 22 Relative caryophyllene production between average of 3 colonies picked of various sizes isolated from screening.

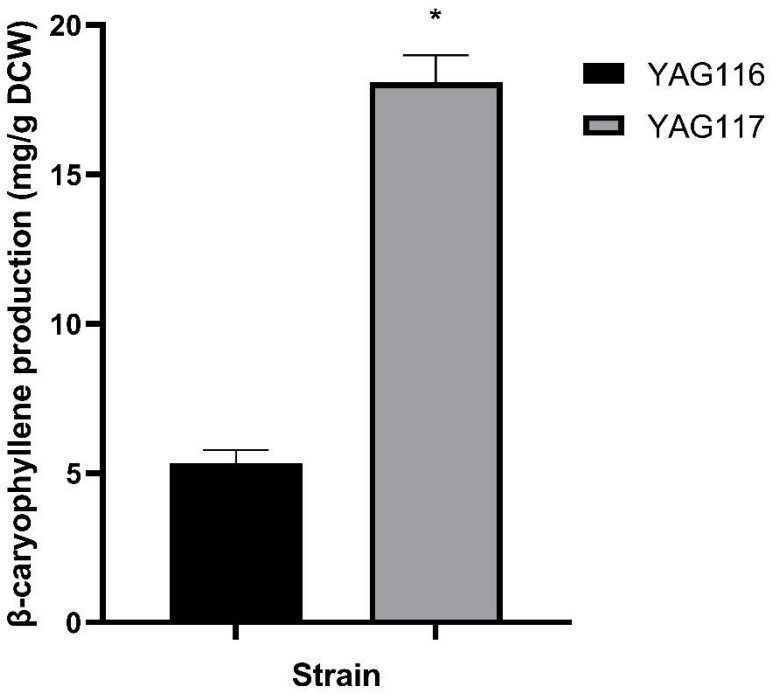


Figure 23 Production of β-caryophyllene in strains with additional copies of *QHS1* gene. Asterisks: p value <0.05 using two-tailed Student's t-test compared with YAG116.

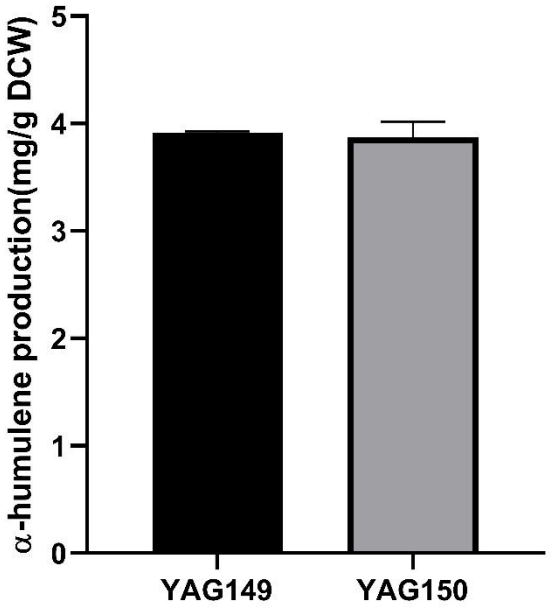


Figure 24 alpha humulene production under overexpressed *STE6 1025*