Title
Production of Simvastatin Acid in Saccharomyces cerevisiae

Author
Bond, Carly

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Production of Simvastatin Acid in *Saccharomyces cerevisiae*

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemical Engineering

by

Carly Bond

2018

ABSTRACT OF THE DISSERTATION

Production of Simvastatin Acid in *Saccharomyces cerevisiae*

by

Carly Bond

Doctor of Philosophy in Chemical Engineering

University of California, Los Angeles, 2018

Professor Yi Tang, Chair

Simvastatin is a semisynthetic cholesterol-lowering medication and one of the top-selling statins in the world. Currently, industrial production of simvastatin acid (SVA) is a multistep process starting from the natural product lovastatin and requires two organisms, *Aspergillus terreus* and *Escherichia coli*. For this reason, there is significant interest in direct production of simvastatin from a single microbial host. In this study, six heterologous genes were introduced into *Saccharomyces cerevisiae* and the acyl-donor dimethylbutyryl-S-methyl mercaptopropionate (DMB-SMMP) was added, resulting in initial production of 0.5 mg/L SVA. Switching the yeast strain from JHY686 to BJ5464-NpgA increased total polyketide production to over 60 mg/L. Conversion from dihydromonacolin L acid to monacolin J acid was increased from 60 to 90% by tuning the copy number of the P450 *lovA*. Increasing the media pH to 8.7 led to a further 10-fold increase in SVA production. Optimized chemical lysis of the cell walls *in situ* after maximum MJA

production led to 55 mg/L SVA, representing near complete from MJA and a 110-fold increase in titer from the initial SVA production strain. In addition, surface expression and secretion of the acyl-transferase LovD were explored as potential alternatives to cell lysis. MJA was added exogenously and these methods led to significant increases in percent conversion from MJA to SVA above cytosolic expression of LovD. The yeast strains developed in this work can be used as an alternative production method for SVA, and the strategies employed in this work can be broadly applied for heterologous production of other fungal polyketides and semisynthetic compounds in yeast.

The dissertation of Carly Bond is approved.

Harold G. Monbouquette

Yvonne Y. Chen

Sriram Kosuri

Yi Tang, Committee Chair

University of California, Los Angeles

2018

Table of Contents

List of figures

List of Tables

Acknowledgements

(*: Equal Contribution) Chapter 1 contains material from Bond, C.*, Tang, Y. & Li, L.* *Saccharomyces cerevisiae* as a tool for mining, studying and engineering fungal polyketide synthases. *Fungal genetics and biology : FG & B* **89**, 52-61, doi:10.1016/j.fgb.2016.01.005 (2016).

This work was supported by National Science Foundation CBET1605877. Carly M. Bond was supported by the NIH T32 Biotechnology Training Grant (2 T32 GM067555).

Thank you to my advisor Professor Yi Tang from the Department of Chemical and Biomolecular Engineering, UCLA for his advice, guidance, and support.

I extend my thanks and appreciation to my undergraduate and high school volunteers Youna Choi, Han Ngo, and Jessee Svoboda for helpful discussions and technical work.

Many thanks also to Anthony DeNicola for his advice, helpful discussions, and general support throughout my time in the Tang lab.

Finally, thank you to my fellow Tang lab members! We have a wonderful supportive, collaborative lab thanks to all of you.

Vita

2013	B.S., Chemical Engineering California Institute of Technology Los Angeles, California, USA
2013-2018	Graduate Research Assistant Department of Chemical and Biomolecular Engineering University of California, Los Angeles, USA
2014-2016	Biotechnology Training Grant Recipient (BTBSE) National Institute of General Medical Sciences University of California, Los Angeles, USA
2016, Summer	Biology Intern Amyris, Inc. Emeryville, California, USA
2017, 2018 Winter	Teaching Assistant Department of Chemical and Biomolecular Engineering University of California, Los Angeles, USA

PUBLICATIONS AND PRESENTATIONS (*: Equal Contribution)

Bond C*, Tang Y, Li L.* *Saccharomyces cerevisiae* as a tool for mining, studying and engineering fungal polyketide synthases. Fungal genetics and biology : FG & B. 2016 Apr;89:52-61. PubMed PMID: 26850128. Pubmed Central PMCID: 4789138.

Noey EL, Tibrewal N, Jimenez-Oses G, Osuna S, Park J, **Bond CM**, et al. Origins of stereoselectivity in evolved ketoreductases. Proc Natl Acad Sci U S A. 2015 Dec 22;112(51):E7065-72. PubMed PMID: 26644568. Pubmed Central PMCID: 4697376.

Bond, C.; Xu, W., Tang, Yi., "Engineering Production of Cholesterol Drugs Simvastatin and Lovastatin in *Saccharomyces cerevisiae*." Metabolic Engineering 11, 2016, The Western Awaji Island Conference Center, Japan, Poster

1. Introduction

1.1 *Saccharomyces cerevisiae* as a tool for mining, studying and engineering fungal polyketide synthases

Small molecule secondary metabolites produced by organisms such as plants, bacteria, and fungi form a fascinating and important group of natural products. Many of these natural products with diverse bioactivities have been important sources of medicines. Approximately 50% of all new chemical entity small molecules approved by the U.S. Food and Drug Administration (FDA) from 2000-2010 and nearly half of the drugs approved from 1994-2008 were derived from natural products[1,2].

Polyketides are one of the major classes of natural products with diverse chemical structures and biological activities. Fungi in particular have been important sources of polyketide pharmaceuticals or virulence factors, and several examples of well-known fungal polyketides are included in Fig. 1[3]. Many of these compounds display clinically relevant activities against human diseases. For example, simvastatin, lovastatin, and mevastatin are all structurally related, natural product-derived cholesterol-lowering medications currently on the market[4]. Squalestatin is an inhibitor of squalene synthases and also shows potential as a cholesterol-lowering medication[5]. In addition to the compounds displaying anti-hypercholesterolemia properties, several polyketides have been studied for their potential anticancer activities. Specifically, cytochalasin E has been shown to inhibit angiogenesis and tumor growth, and has been considered for use in cancer treatment and age-related macular degeneration[6]. Brefeldin A has also shown potential as an anticancer agent due to its induction of apoptosis in several cancer cell lines, shown to be via activation of the mitochondria-mediated cell death pathway in ovarian cancer cells[7,8]. Mycophenolic acid is a known immunosuppressant for organ-transplant patients but has shown potential as a method of inducing apoptosis in tumor

cells through inosine monophosphate dehydrogenase (IMPDH) inhibition[9]. Griseofulvin was launched as antifungal agent in the 1950s and has now attracted renewed attention for its anticancer and antiviral activities[10]. However, there are fungal polyketides hazardous to human health, such as aflatoxins, which are the primary mycotoxins produced by some *Aspergillus* sp. and considered to be the most potent naturally occurring carcinogens [11].

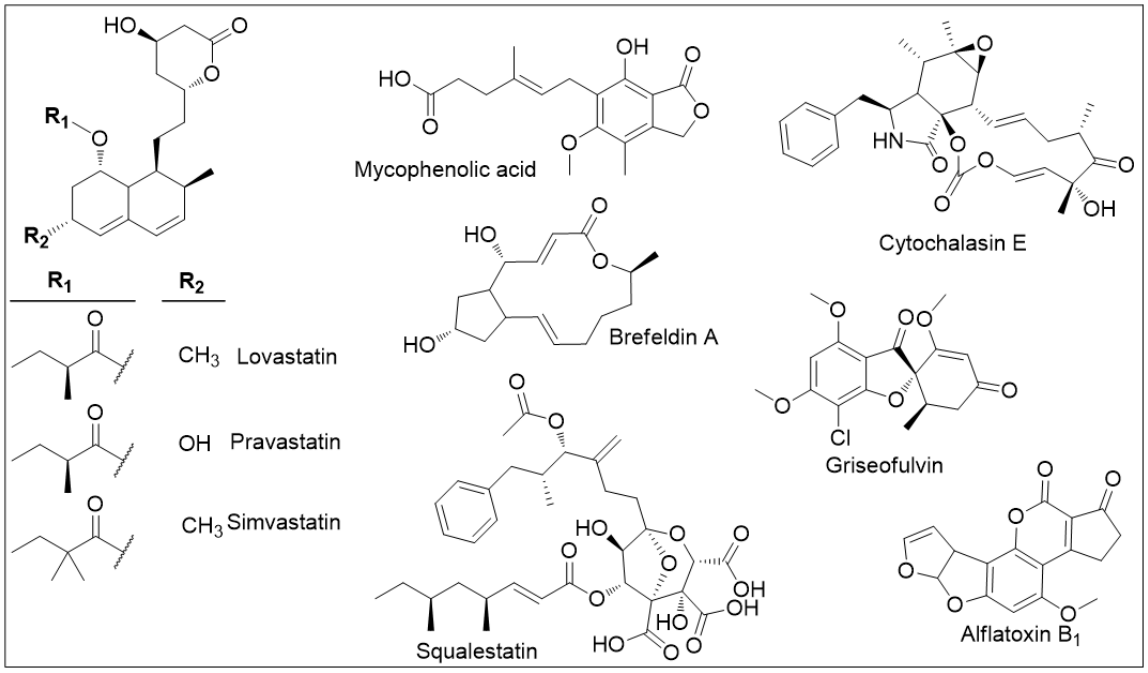

Figure 1: Examples of important fungal polyketides

While the structural complexity of these polyketides makes them interesting and useful bioactive compounds, these same features also make them difficult and expensive to prepare and scale-up using synthetic methods. Currently, nearly all commercial polyketides are prepared through fermentation or semi-synthesis[12,13]. Thus, elucidating the biosynthetic pathways of polyketides and characterizing the pathway enzymes is necessary for the production and diversification of natural products for pharmaceutical applications[14]. However, efforts to elucidate and engineer polyketide pathways in the

native filamentous fungi hosts are often hampered for the following reasons: 1) many fungal hosts lack established genetic tools; 2) the native hosts of desired natural products may have low biomass accumulation and produce low concentrations of the desired product; 3) fungal hosts often produce many secondary metabolites that can lead to significant background, allowing the possibility of cross-reactivity between pathways, and complicating the analysis of specific pathways; 4) transcriptional regulation of fungal natural product clusters is complex and not fully understood; and 5) there is no universal expression system with specified culture conditions that can be applied uniformly to fungal natural product pathways and their native hosts [15,16,17, 18]. For these reasons, a versatile heterologous host for the study and engineering of fungal polyketide synthases (PKSs) is desired.

As one of the most intensively studied single-celled eukaryotes in fundamental and applied molecular biology research, *Saccharomyces cerevisiae* has proven to be a useful and prominent industrial host for recombinant protein production[19]. *S. cerevisiae* is widely used not only in the food and beverage industry, but also in the production of bioethanol and fine chemicals[20]. Specifically, *S. cerevisiae* has many advantages beneficial to the elucidation and engineering of heterologous biosynthetic pathways from filamentous fungi: 1) A number of genetic tools for protein expression and pathway construction in yeast have been developed; 2) it is a unicellular organism well-suited for large scale fermentation; 3) *S. cerevisiae* has a limited native secondary metabolism, which minimizes the background and potential interference with heterologous pathways; 4) *S. cerevisiae* grows more rapidly than most filamentous fungi and it is considered a GRAS (generally regarded as safe) organism by the FDA[20]. In addition, yeast naturally produces common polyketide building blocks such as acetyl-CoA and malonyl-CoA; as well cofactors such as NADPH and *S*-adnesylmethioine, which facilitate the production of fungal

polyketides with minimal integration of heterologous genes[21]. Like filamentous fungi, yeast is also classified as a fungus and is expected to be a capable expression host for fungal proteins that are important for polyketide pathways. For example, *S. cerevisiae* can functionally express eukaryotic cytochrome P450s because these enzymes often anchor in the endoplasmic reticulum, which is absent in prokaryotes[22].

Using yeast as a heterologous host presents some challenges, such as required genes for PKS activation, an inability to splice most fungal introns, low production of necessary precursors, a lack of compartmentalization, and potential toxicity. However, the advantages of *S. cerevisiae* have allowed increased understanding of fungal polyketide pathways and biosynthesis[23]. For example, of the polyketides in Fig. 1, the lovastatin and brefeldin A biosynthetic pathways were identified using heterologous yeast expression[24,25,26] and the effects of statins, brefeldin A, mycophenolic acid, and griseofulvin have all been studied in yeast[27,28,29,30,31,32]. Aflatoxin pathway genes have been expressed heterologously in yeast in order to study aflatoxin biosynthesis, and its mode of action has also been studied using yeast[33,34,35]. This review will focus on the use of *S. cerevisiae* as a tool for the discovery, study, and production of fungal PKSs and their natural products.

1.2 Introduction to polyketide synthases

The diverse structures of polyketides are biosynthesized from short-chain carboxylic acid units by PKSs[3]. PKSs have been classified into type I, type II and type III based on their product profiles and catalytic domain architecture[36] Type I PKSs are large multidomain enzymes in which catalytic sites are juxtaposed in an assembly line fashion. The three essential domains for the elongation of the polyketide chain are β-ketoacyl synthase (KS), acyltransferase (AT) and acyl carrier protein (ACP). Other domains that control the degree

of reduction of β-keto groups may be present. These are ketoreductase (KR), dehydratase (DH) and enoyl reductase (ER) domains. Other frequently found tailoring domains include the methyltransferase (MT) domain, which introduces an α-methyl group immediately after a round of chain elongation; the thioesterase (TE) domain, which releases the polyketide product from the enzyme by hydrolysis, esterification or macrocyclization[37]; the reductase (R) domain which releases the polyketide product from the enzyme by two or four electron reduction or Dieckmann condensation[38]; a special type of TE domain called TE/CLC (Claisen-like cyclase) which catalyses Claisen-type condensations to release the products[39]. Type II PKSs are a set of multienzyme complexes that act iteratively and are frequently responsible for the biosynthesis of aromatic polyketides in bacteria. Type III PKSs are homodimers of KSs which catalyze the priming, extension, and cyclization of small polyketides, such as chalcone, in an ACP-independent fashion[36,40].

The majority of PKSs from filamentous fungi are type I PKSs. Unlike the multimodular bacterial type I PKSs that operate in a collinear fashion in which each set of domains (a module) are used once in the construction of the polyketide, fungal type I PKSs contain only one set of catalytic domains which are used iteratively in a well-programmed fashion to biosynthesize the final products. Hence, fungal PKSs are also known as iterative type I PKSs (iPKSs) (Fig. 2). Furthermore, based on the extent of β-keto reduction catalyzed by the iPKSs, the fungal iPKSs can be classified into three subgroups: nonreducing PKSs (NR-PKSs) that produce aromatic compounds such as norsolorinic acid, the precursor to aflatoxin[41], via the product template (PT) domain which is necessary for synthesizing these ring-shaped products[39]; partially-reducing PKSs (PR-PKSs) that produce compounds such as 6-methylsalicylic acid (6-MSA)[42]; and highly-reducing PKSs (HR-PKSs) that produce more reduced compounds such as lovastatin[39,43]. In addition, an HR-PKS can be fused to

a nonribosomal peptide synthetase (NRPS) module to form a hybrid PKS-NRPS, which can lead to the biosynthesis of tetramic acid-containing products[44]. Collectively, the different combinations and programming rules of the iPKSs, along with further tailoring of the initial scaffolds by other enzymes, have led to the tremendous structural and functional diversity of polyketides isolated from fungal species.

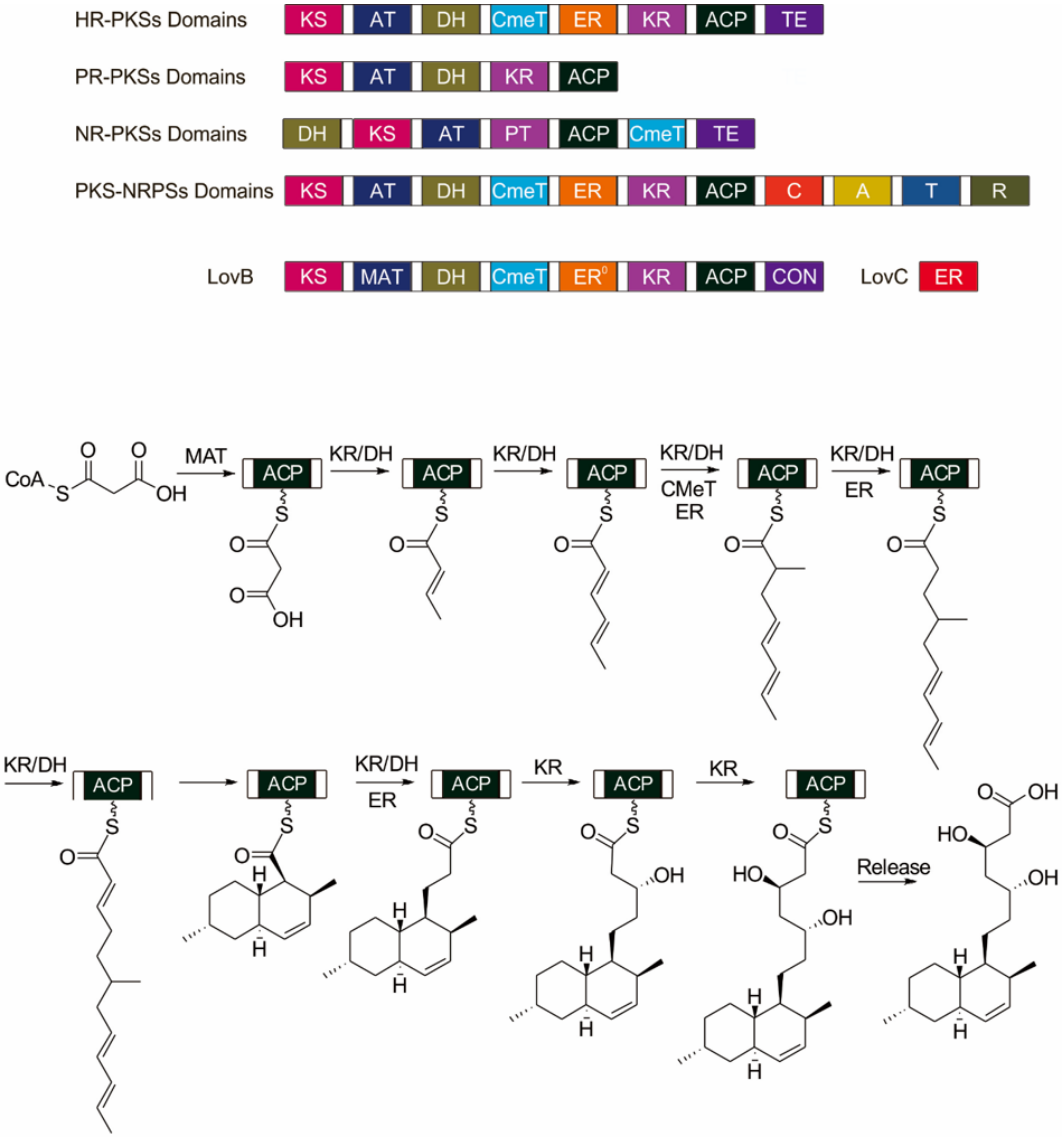

Figure 2: The domains of several types of iterative type I polyketide synthases (iPKSs) are shown here. The PKS domains abbreviated here are as follows: KS (Keto-synthase), AT (Acyltransferase), DH (Dehydratase), CmeT (C-methyltransferase), ER (Enoylreductase), KR (Ketoreductase), ACP (Acyl carrier protein), TE (Thioesterase), PT (Product template), MAT (malonyl-CoA:ACP acyltransferase), and CON (Condensation). The PKS-NRPS hybrid also contains Non Ribosomal Peptide Synthetase domains as follows: C (Condensation), A (Adenylation), T (Thiolation), and R (Reduction). The example below shows the action of a nonaketide synthase, LovB, and its partner enoylreductase, LovC, from the lovastatin pathway of *Aspergillus terreus*.

1.3 The *S. cerevisiae* toolbox for cloning and enzyme reconstitution

1.3.1 Molecular biology tools of *S. cerevisiae* for cloning and reconstitution of heterologous pathways

Polyketides and other secondary metabolites are biosynthesized by a series of enzymes encoded by genes that are typically clustered together in the genomes of the producing organisms[45]. The clustering of related genes has been instrumental in the discovery and engineering of natural product pathways in both bacteria and fungi[46]. Due to this clustering, the cloning and expression of polyketide biosynthetic enzymes can be accomplished with only one continuous genomic DNA fragment containing the entire cluster. However, the cloning or assembly of the pathway in a suitable vector is still a challenge because the large size of most gene clusters, or even a single PKS-encoding gene (iPKSs genes are typically ~10 kB), is usually too large to be amplified by PCR efficiently and correctly.

One strategy to capture entire gene clusters is to construct a genomic DNA library with a suitable vector and chemically screen for clones that may carry a functional cluster. For example, in the first heterologous expression of penicillin, a cosmid containing the penicillin biosynthetic gene cluster from *Penicillium chrysogenum* was screened from a cosmid library using a DNA probe of the homologous isopenicillin N synthetase from *Flavobacterium* sp. SC 12,154. This cosmid was then transformed into *Neurospora crassa* and *Aspergillus niger*, which led to the production of authentic penicillin V in both transformed hosts[47]. However, this method can fail when the size of the gene cluster is larger than the capacity of a cosmid, or the library does not include a cosmid clone that contains all of the genes involved in the biosynthesis of the natural product. Recently, Bok *et al*. constructed an unbiased shuttle BAC library of *Aspergillus terreus* ATCC2054 with the vector containing both the *Escherichia coli* replicon and *Aspergillus* autonomously

replication sequence (AMA1). The average insert size was about 100 kb, which can cover all genes and regulatory elements of the biosynthetic pathway and be used successfully in the heterologous expression of secondary metabolites[48]. However, the construction of high quality unbiased BAC library is time-consuming and labor-intensive, so alternative solutions have been explored using *S. cerevisiae*.

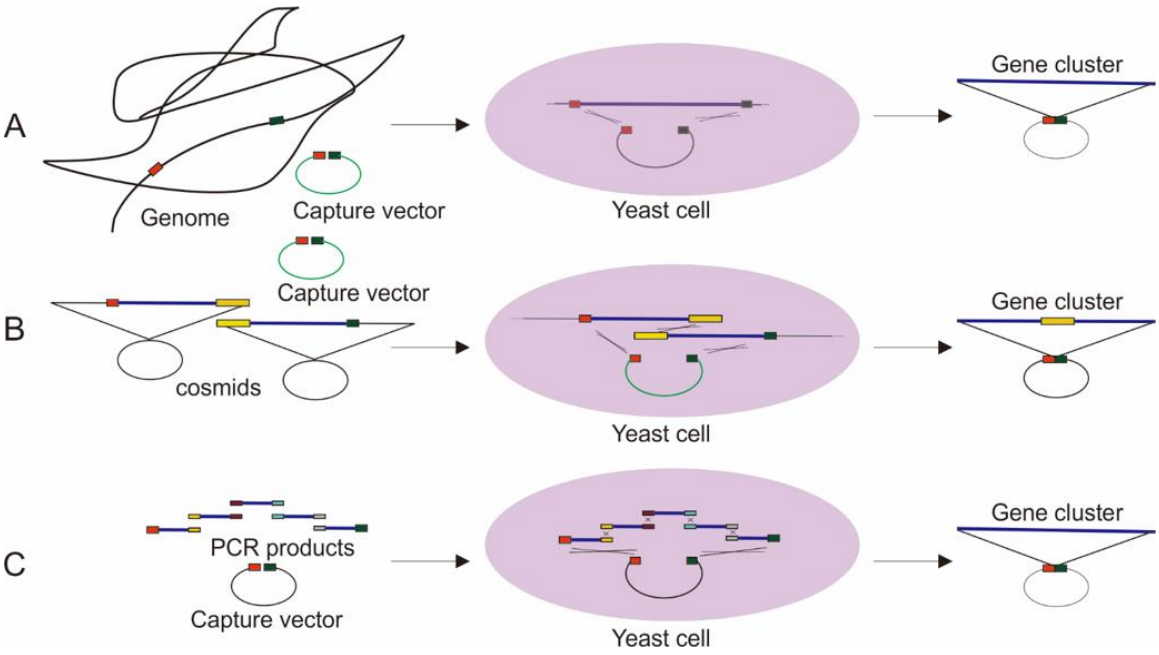

Figure 3: Cloning of large DNA fragments based on the recombination of yeast
A Capture of large DNA fragments from genomic DNA
B Assembly of interested gene cluster from overlapping cosmids
C Assembly of interested pathway from overlapping PCR products

Yeast has been developed extensively as a synthetic biology tool for the cloning and capture of entire biosynthetic gene clusters (Fig. 3). One significant advantage of yeast is that homologous recombination takes place far more frequently than ligation or non-homologous end joining during *S. cerevisiae* DNA repair[49]. This feature has been exploited to construct vectors containing large gene clusters. Overlapping DNA pieces and a linearized vector can be co-transformed into yeast spheroplasts, and the DNA fragments can then be joined by homologous recombination to form intact, selectable vectors. This

method is generally known as Transformation-associated Recombination (TAR) as well as other names such as DNA assembler[50,51]. Oldenburg *et al.* investigated the efficiency of TAR cloning with different lengths of homologous overlaps, and confirmed that while 40 bp of overlap produced optimal results, as few as 20 bp of overlap could generate the desired product[52]. TAR cloning is exceptionally simple and efficient, especially in the rapid cloning of large DNA genes or gene clusters, without using restriction enzymes or being limited by PCR product sizes[53].

TAR has been used to assemble two or more overlapping cosmids in one step. Feng *et al* used TAR in *S. cerevisiae* to assemble the entire fluostatins biosynthetic gene cluster from a bacterial environmental DNA (eDNA) library. Initially, when the cosmid library was screened, they found that the fluostatins cluster was located across two different cosmids. In order to assemble the entire cluster, a *S. cerevisiae*/*Escherichia coli*/*Streptomyces* capture vector was designed with 1 kb homology regions matched to the boundaries of the gene cluster. The linearized vector and the two cosmids were then co-transformed into yeast and homologously recombined[54]. Yeast based TAR has also been used to capture large DNA fragments from the genome directly[53]. Analysis of the genome of a marine bacterium, *Saccharomonospora* sp. CNQ-490, by Yamanaka *et al* revealed a putative NRPS gene cluster similar to the gene cluster responsible for synthesis of the antibiotic daptomycin in *Streptomyces roseosporus*. To mine this gene cluster, 1 kb DNA fragments matching the boundaries of the targeted gene cluster were cloned by PCR and inserted into the capture vector. The linearized capture plasmid and genomic DNA were co-transformed into yeast to capture the cluster. The resultant plasmid carrying the cluster was directly used for heterologous expression, leading to the heterologous production of taromycin A[55].

For accurate and efficient cloning of large fungal iPKSs genes, PCR combined with yeast homologous recombination has been widely used in recent years. A common strategy to clone large intron-free iPKSs genes from cDNA for yeast expression is to first PCR amplify several overlapping fragments from the cDNA of the iPKSs gene, followed by one-step recombination into the desired expression vector in yeast. The use of cDNA is to guarantee the correct mRNA processing in the yeast. One recent example is the cloning and reconstitution of iPKSs AurA from the aurovertin biosynthetic gene cluster in *Calcarisporium arbuscula*. Mao *et al* amplified three overlapping fragments covering the intron-free *aurA* cDNA. The two 5' and 3' pieces also contained overlapping regions with the 2μ yeast vector pXW55, which led to placement of the entire *aurA* gene in the vector under control of the *ADH2* promoter. The resulting yeast strain containing the desired plasmid was then used directly to elucidate the product of the iPKSs through expression of the encoded enzymes and analysis of the resulting products[56]. Other examples of TAR-based assembly of iPKSs include *cazF* and *cazM* from the chaetoviridin biosynthetic gene cluster in *Chaetomium globosum*[57,58]; *bref-PKS* from the biosynthetic gene cluster of brefeldin A in *Eupenicillium brefeldianum*[26]; and *fma-PKS* from the biosynthetic gene cluster of fumagillin in *Aspergillus fumigatus*[59]. A variation of this strategy is to first assemble the entire iPKSs gene using overlap extension PCR (OE-PCR), followed by transformation into yeast along with the vector to yield the intact expression plasmid via recombination. This ExRec (overlap Extension PCR-yeast homologous recombination) method described by Ishiuchi *et al* was used to successfully reconstitute iPKSs from *Chaetomium globosum* and *Coprinopsis cinerea*[60].

Yeast homologous recombination has also been used in the refactoring of multi-gene fungal biosynthetic pathways for heterologous reconstitution in model hosts. Pahirulzaman *et al* reconstructed the four gene tenellin biosynthetic pathway from

Beauveria bassiana for expression in *Aspergillus oryzae*. The three tailoring genes were first recombined into a single expression vector under the control of three different promoters. The PKS-NRPS gene was first cloned in a separate plasmid, followed by insertion into the three-gene plasmid via Gateway-mediated recombination *in vitro*[61]. Kakule *et al* also employed yeast homologous recombination to construct cryptic heterologous fungal PKS pathways in *Fusarium heterosporum*. The genes of interest were combined into a *S. cerevisiae*/*E. coli* /*F. heterosporum* shuttle vector via yeast homologous recombination, amplified in *E. coli*, and expressed in *F. heterosporum*. The PKS CpaS from *Aspergillus flavus*, the PKS LovB and enoylreductase LovC from *A. terreus*, and two putative PKS-NRPSs, PrlS and PrlC, from endophytic fungus NRRL 50135 were all successfully reconstituted in *F. heterosporum* via this strategy, leading to production of up to 1 g/L of the encoded polyketide[62].

1.3.2 *S. cerevisiae* as a platform for PKS protein expression and purification

Purified iPKSs enzymes from fungal PKS biosynthetic pathways are needed for the functional characterization of their unique programming rules. Complete understanding of PKSs will also enable their abilities to be exploited to benefit the structural diversification, activity optimization, and generation of "unnatural" natural products[14]. As an expression host for fungal proteins, *S. cerevisiae* has many advantages compared to prokaryotic or more complex eukaryotic hosts. Unlike bacterial hosts, such as *E.coli*, *S. cerevisiae* has the machinery for secretory pathways and post-translational modifications[19]. Compared to mammalian and plant cells, *S. cerevisiae* has fast growth and allows easy genetic manipulation.

Nevertheless, there are some genetic modifications that were made to *S. cerevisiae* to ensure functional expression of iPKSs. Deletion of the two vacuolar protease genes (*PEP4*

and *PRB1*), encoding the aspartyl protease and proteinase B respectively, significantly increases the expression level of recombinant fungal iPKSs. The strain BJ5464 (*MATα ura3-52 his3-Δ200 leu2-Δ1 trp1 pep4*::*HIS3 prb1 Δ1.6R can1 GAL*) was therefore chosen for protein expression[63,64]. In addition, the yeast host must ensure correct post-translational modification of the ACP domain of iPKSs. The active site serine of ACPs must be modified with a phophopantetheinyl moeity to afford the *holo* version. While yeast has its endogenous 4'-phosphopantetheine (pPant) transferase that transfers pPant from coenzyme A to ACP domain of fatty acid synthases, the ACP domain of most fungal iPKSs cannot be modified and hence the iPKSs is not active. Kealey *et al* confirmed that there were almost no functional 6-methylsalicylic acid synthases (6-MSAS) expressed in the *S. cerevisiae* without a heterologous pPant transferase[21,65]. To this end, *npgA*, a pPant transferase gene from *Aspergillus nidulans*[65,66] was integrated into the genome of *S. cerevisiae* BJ5464 to yield BJ5464-NpgA. Using a sensitive fluorescent assay, the authors showed that expressed fungal PKSs were efficiently phosphopantetheinylated in the engineered strain[65]. This strain has subsequently been used widely in the expression and purification of functional iPKSs, as first demonstrated with the 335 kDa lovastatin nonaketide synthase LovB[25].

1.4 Yeast as a platform for studying the function of fungal iPKSs

S. cerevisiae has been used extensively as a host for the expression and characterization of iPKS pathways, and to link fungal polyketide natural products to the gene clusters that produce them. For example, while investigating the biosynthesis of the protein transport-inhibitor Brefeldin A (BFA), Zabala *et al* sequenced the producing organism *Eupenicillium brefeldianum* ATCC 58665. From the numerous gene clusters that contained iPKSs, one putative gene cluster encoding an HR-PKS and numerous P450 genes was proposed to be most likely involved in BFA biosynthesis. Genetic manipulation of the producing organism

proved to be extraordinarily difficult. Therefore, the authors performed direct expression in yeast to investigate the role of the HR-PKS. The HR-PKS gene and the partnering thiohydrolase gene were cloned into two vectors from cDNA and heterologously expressed in *S. cerevisiae* BJ5464-NpgA. While the authors did not observe the completed core structure of BFA, an acyclic octaketide consistent with the length of BFA was recovered. The oxidation patterns of the octaketide product were consistent with those observed in BFA, thereby providing strong evidence linking this gene cluster to BFA[26].

Yeast has also been used to connect an orphan NR-PKS gene cluster to the biosynthesis of fungal aromatic polyketide TAN-1612. Li *et al* identified a candidate cluster in *Aspergillus niger* that was presumed to be involved in the production of known compound TAN-1612. To verify the function of this cluster, the putative NR-PKS encoded by the *adaA* gene was cloned from cDNA into a yeast 2μ expression plasmid and the three tailoring genes *adaB-D* from the cluster were cloned into a separate plasmid. After two days of culturing, *S. cerevisiae* BJ5464-NpgA expressing these two plasmids produced TAN-1612[67]. Subsequent investigations using the yeast host revealed the product of the NR-PKS alone, as well as the individual functions of the tailoring enzymes. Similarly, Zhou *et al* utilized yeast to confirm that two iPKSs, an HR-PKS Rdc5 and an NR-PKS Rdc1, were involved in the biosynthesis of the radicicol precursor in *Pochonia chlamydosporia*. Heterologous expression of these iPKSs, cloned from genomic DNA, was performed in *S. cerevisiae* BJ5464-NpgAand led to the production of (*R*)-monocillin II, an intermediate in radicicol biosynthesis. This result confirmed that the two iPKSs function collaboratively in the biosynthesis of the resorcylic acid lactone and allowed a closer study of the functions of these enzymes[68]. This led to the yeast-based reconstitution of other dual iPKSs clusters from fungi, including those of chaetoviridine, resorcylides, lasiodiplodins, and cuvularins[69].

Chooi *et al*. sought to identify virulence factors from *Parastagonospora nodorum*, a wheat pathogen affecting wheat yields globally. In the course of their investigation, they found that *SN477*, a PR-PKS gene, was highly upregulated during the pathogen infection. When *SN477*, was cloned from cDNA under the *ADH2* promoter and transformed into *S. cerevisiae* BJ5464-NpgA, (*R*)-mellein was produced by the yeast host, revealing that SN477 is the only enzyme required for (*R*)-mellein synthesis. Though (*R*)-mellein showed no relevance to the virulence against wheat, it was able inhibit the germination of wheat seeds[70].

In addition to serving as a host for linking polyketide metabolites to their corresponding iPKSs, the yeast iPKS expression platform has also proven to be useful in the mechanistic studies of these highly programmed machineries. One example is the characterization of LovB, an HR-PKS from *Aspergillus terreus* that is involved in the biosynthesis of the cholesterol-lowering compound lovastatin[71]. The low yield of functional LovB (335 kDa) recovered from *Aspergillus*-based expression hosts significantly hindered the biochemical study of this model HR-PKS. When expressed from *S. cerevisiae* BJ5464-NpgA, a purified LovB yield of ~4.5 mg/L was achieved, thereby providing sufficient amounts for *in vitro* biochemical investigations[25]. Reconstitution of LovB with its enoylreductase LovC demonstrated that the enzyme is precisely programmed to synthesize the expected product dihydromonacolin L when the needed cofactors (NADPH and SAM) are supplied. After perturbing the system through removal of one or more of the required cofactors or LovC, LovB produced a series of conjugated α-pyrones that are drastically different in structure from dihydromonacolin L. However, detailed structural characterization of these shunt products revealed that releasing pyrones is one mechanism by which LovB can clear its ACP of incorrectly tailored intermediates, thereby providing insight into how iPKSs can maintain their product fidelity. Subsequently, using the yeast expression

platform and *in vitro* characterization, Xu *et al* discovered a previously hidden thioesterase, LovG, encoded in the pathway to be the enzyme responsible for both the release of dihydromonacolin L from LovB and significant increases in its turnover rate[72].

Using the same yeast expression host, Xie *et al.* also reconstituted the activities of the lovastatin diketide synthase (LovF) using purified enzymes. While architecturally similar to LovB, the authors showed that this enzyme produces an enzyme-bound α-methylbutyrate diketide using Fourier Transform Mass Spectrometry (FTMS) of proteolyzed LovF fragments. Offloading of the product was demonstrated to be carried out by the acyltransferase LovD, which transfers the α-methylbutyrate to the C8-hydroxyl group of monacolin J to yield the final product, lovastatin. Kinetic analysis using LovF and smaller acyl mimics such as α -methylbutyryl-CoA or α -methylbutyryl-SNAC demonstrated protein-protein interactions between the LovF ACP domain and LovD are highly specific, as significant penalties to the acyltransfer reaction were observed when small molecules thioester carriers were used in place of the ACP domain. This was the first demonstrated acyltransferase mediated product release from an iPKSs and since then this mechanism has been found to be widely adopted by other fungal PKS pathways[39,73]. Together with the functional reconstitution of the lovastatin P450 LovA in yeast[24], the entire six-gene biosynthetic pathway of lovastatin has been functionally elucidated in *S. cerevisiae*, paving the way for heterologous pathway construction and engineering.

In another example, Wang *et al.* demonstrated the aryl-aldehyde formation in the biosynthesis of an NR-PKS through heterologous expression of a cryptic NR-PKS and an NRPS-like gene from *Aspergillus terreus* in yeast. When the cryptic NR-PKS ATEG03629 was expressed in *S. cerevisiae* BJ5464-NpgA, 5-methyl orsellinic acid (5-MOA) was produced. However, when both ATEG03629 and ATEG03630, an NRPS-like gene with a terminal reductase domain, were co-transformed into the yeast host, 2,4-dihydroxy-5,6-

dimethyl benzaldehyde was produced in addition to 5-MOA. Both ATEG03629 and ATEG03630 were cloned from genomic DNA using exons predicted from bioinformatics analysis. The *in vivo* results indicated that the NRPS-like protein catalyzes the aryl-acid to aryl-aldehyde conversion. To further confirm the activity of ATEG03630, the enzyme was purified from BJ5464-NpgA. *In vitro* experiments of ATEG03630 with the substrate 5-MOA and cofactor NADPH confirmed the catalytic activity of ATEG03630. Though these compounds had been reported previously, their biosynthetic origins had never been established[74].

As the above examples demonstrate, yeast is a useful and versatile tool for production of diverse polyketides and their associated mega-enzymes in sufficient quantities for *in vitro* mechanistic studies of iPKSs. These applications of *S. cerevisiae* improve the knowledge of biosynthesis of natural products and accumulate enzyme tools for drug discovery and the synthesis of new natural products.

1.5 Engineering production of fungal PKSs and polyketides in yeast

The deletion of native proteases and the integration of a heterologous pPant transferase have been discussed earlier as essential for the functional expression of fungal iPKSs in *S. cerevisiae*. Other methods of increasing the titer of the final polyketide product include increasing the supply of the common polyketide precursors acetyl-CoA and malonyl-CoA, and introducing self-resistance genes to mitigate the toxicity of final product. These metabolic engineering strategies can be employed in the production of both natural and engineered, unnatural polyketides in yeast. The route from the discovery of novel polyketides to high production in yeast is outlined in Fig. 4.

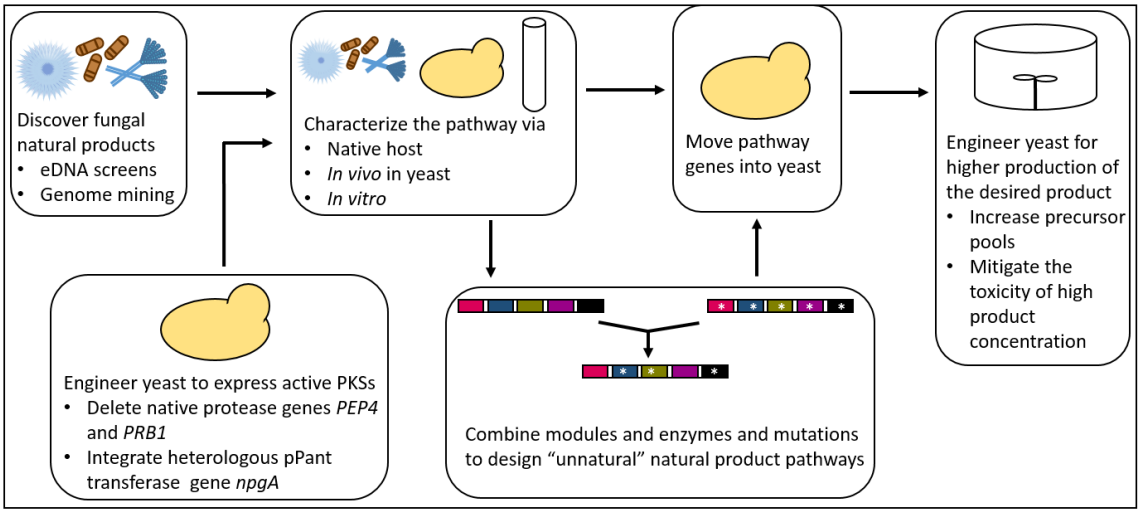

Figure 4: The route from the discovery of novel polyketides to high production in yeast, using the tools and methods discussed.

1.5.1 Strain engineering for high titers via increased precursor production

Increasing the supply of the precursors is one common metabolic engineering strategy to increase the titer of the final product[75]. Common polyketide precursors acetyl-CoA and malonyl-CoA are produced endogenously by yeast, but metabolic engineering is required to produce larger quantities of these substrates. Recent efforts to increase precursor production have focused both on enzyme overexpression and engineering, as well as on directing carbon towards the desired precursors via deletion of competing pathways.

The *ACC1* gene encodes acetyl-CoA carboxylase, which catalyzes the conversion of acetyl-CoA to malonyl-CoA. Wattanachaisaereekul *et al* overexpressed the native yeast *ACC1* through replacing the native promoter of *ACC1* with a strong constitutive TEF1 promoter in a 6-MSA polyketide production *S. cerevisiae* strain, increasing the titer of 6-MSA by 60%[76]. Choi *et al*. improved the activity of *ACC1* by site-directed mutagenesis, leading to 3-fold increases in 6-MSA levels[77].

Metabolic pathway analysis in the well-studied *S. cerevisiae* has been used to determine how to direct carbon towards the desired product. Cardenas *et al* specifically increased the titer of precursor acetyl-CoA through the deletion of 15 genes identified during pathway analysis of glucose-6-phosphate, an acetyl-CoA and malonyl-CoA precursor[63]. Lian *et al* redirected glycolytic flux to the production of acetyl-CoA via gene inactivation of *ADH1* and *ADH4* from the competing ethanol pathway and *GPD1* and *GPD2* from the competing glycerol pathway[78]. Chen *et al* over-expressed the endogenous *ADH2* and *ALD6* genes as well as a mutant heterologous acetyl-CoA synthetase from *Salmonella enterica* to redirect carbon flux from acetaldehyde to cytosolic acetyl-CoA. They also reduced carbon loss from the pool of acetyl-CoA by inhibiting *CIT2*, a peroxisomal citrate synthase, and *MLS1*, a cytosolic malate synthase. Using this platform to produce α-santalene led to a four-fold increase in titer compared to the reference strain[79].

1.5.2 Strain engineering for high titers via toxicity mitigation

One challenge of using yeast as a cell factory to produce high titers of fungal polyketides is that many bioactive secondary metabolites are toxic in high concentrations[80]. Some strategies for producing toxic compounds in cell factories include using inducible promoters to decouple cell growth from compound production[20]; overlaying an organic solvent like dodecane on the top of the culture to remove the toxic compound as it is produced[81]; and overexpressing native, broadly-specific transport proteins to remove toxic compounds from the cell[82]. Native producers often evolve their own solutions to protect themselves from the presence of toxic compounds, such as specific transporter proteins or other enzymes that confer self-resistance to the native host[80]. These self-resistance mechanisms can take the form of product-specific transporters like the putative transporter *cazK* from *Chaetomium globosum* and the putative efflux pump *lovI* from *A. terreus*[71,58]. Recently, Ley *et al* showed that integration of the putative efflux pump *mlcE*,

from the compactin PKS cluster of *P. citrinum,* into the *S. cerevisiae* genome conferred increased resistance to mevastatin, lovastatin, and simvastatin, as compared to the wild type strain. MlcE was shown to be a specific transporter and restored the growth rates of yeast in the presence of up to ~800 mg/L of exogenously added lovastatin[83]. Past analyses of fungal PKS clusters have primarily focused on reconstitution of the enzymes necessary for the biosynthesis of the encoded natural product, but, as Ley, *et al* showed, future work incorporating more of the putative self-resistance genes into heterologous polyketide production hosts has the potential to significantly increase final product titers[83].

1.6 Combinatorial biosynthesis of fungal PKSs in *S. cerevisiae*

Due to the contributions of natural products to the development of pharmaceuticals, there has been significant research toward discovering new natural products and engineering the biosynthesis of novel, unnatural bioactive compounds[84,2,1]. One future goal is to set up an algorithm which takes a molecule of interest as input and outputs a sequence of natural PKS modules to produce the desired molecule[85]. Although the synthesis of new "unnatural" natural products via combinatorial biosynthesis has been pursued in bacterial polyketide antibiotics for more than 15 years[86], there had been little progress in the field of fungal PKSs until recently. In 2013, Xu *et al.* reported the reprogramming of the first-ring cyclization of two benzenediol lactones (BDLs)[87] . BDLs are a family of fungal polyketides with diverse structural features and wide-ranging bioactivities. The BDL family is composed of resorcylic acid lactones (RALs), connected at C2–C7, and dihydroxyphenylacetic acid lactones (DALs), which feature a C3–C8 bond[88]. The biosynthesis of these fungal polyketides involves pairs of collaborating iPKSs: an HR-PKS which passes its product to an NR-PKS for further modifications[89]. Xu *et al* reconstituted heterologous HR-PKS and NR-PKS pairs from *A. terreus* and *Chaetomium chiversii* in BJ5464-NpgA and confirmed that the product template (PT) domains of the fungal NR-

PKSs regiospecifically catalyzed the first-ring aldol cyclization, leading to the characteristically different polyketide folding modes of RALs and DALs. Next, rational reprogramming of the regiospecific first-ring cyclization was realized by domain replacement and site-directed mutagenesis[87]. Using the same BJ5464-NpgA system, Xu *et al.* expressed chimeric iPKSs enzyme pairs, resulting in the biosynthesis of unnatural BDLs, and found that the thioesterase (TE) domain acts as the decision gate for releasing the final product from a fungal NR-PKS. This indicated that in combinatorial biosynthesis, the TE domain must be able to accept altered polyketide intermediates and release unnatural natural products with the desired structure[90].

After the characterization of different domains of HR-PKSs and NR-PKSs in the biosynthetic pathways of BDLs, Xu *et al.* deployed yeast as a tool to investigate whether noncognate HR-PKS and NR-PKS pairs could interact with each other efficiently[69]. Through the combinatorial expression of random pairs of iPKSs subunits from four BDL biosynthetic pathways from *A. terreus, C. chiversii, Lasiodiplodia theobromae,* and *Acremonium zeae* in BJ5464-NpgA, a diverse library of BDL congeners was created. One of these unnatural polyketides had heat shock response-inducing activity that had previously been shown to block multiple cancer-causing pathways[69]. This combinatorial work by Xu *et al.* provided more insight into PKS design rules, which will assist in future engineering of diverse natural products.

Another method for the biosynthesis of unnatural natural products is to feed artificial precursors to organisms heterologously expressing PKSs. Zhou *et al* and Gao *et al* both fed large acyl SNAC (N-acetylcysteamine thioester) substrates to fungal iPKSs expressed or purified from BJ5464-NpgA[68,91]. Both groups were able to produce unnatural analogs of the relevant natural product. Gao *et al* found that the efficiency of incorporation of the

unnatural precursor analogs depended on the nature of the structural changes between these analogs and the natural precursors[91].

Finally, novel scaffolds can be found by combining existing heterologous natural product pathways in yeast and analyzing the resulting compounds. Klein *et al* took genes from known natural product pathways for alkaloids, benzoxazinoids, flavonoids, flavonols, lignans, polyunsaturated fatty acids, tetra- and diterpenoids, and type III polyketides and 14 libraries of cDNA from 17 different organisms including plants, animals, fungi, and amaurochaetes and expressed them in different combinations on yeast artificial chromosomes in *S. cerevisiae*. The resulting compounds were screened for useful pharmaceutical activities and for the novelty of their scaffolds[92].

The successful combinatorial biosynthesis of fungal PKSs in *S. cerevisiae* has not only shown the utility of yeast for engineering improved production of natural products, but also extended this paradigm from bacterial polyketides to fungal polyketides[93].

1.7 Discovering new natural products through genome mining and expression in *S. cerevisiae*

Another method of discovering new natural product scaffolds is to find novel natural product gene clusters. Top-down approaches, in which newly-discovered fungi are cultivated and their products analyzed, can lead to the discovery of new polyketides but are limited to compounds that are naturally synthesized in relative abundance in the native environment or under laboratory conditions. However, recent advances in genome sequencing and increased availability of fungal genomes can facilitate the discovery and analysis of new polyketides from putative clusters in previously studied organisms.

For over two decades, it has been known that fungal secondary metabolites are often synthesized by genes physically clustered in the genome[45]. As more sequenced fungal genomes become available, extensive genome mining efforts have been launched, leading to improved algorithms to annotate putative polyketide clusters in various fungi [94,95,96]. However, the metabolites encoded by clusters located this way are often unknown, and subsequently termed cryptic or orphan. Some silent clusters found via genome mining have not lead to any product in the native host under laboratory conditions[15,97].

The products of fungal iPKSs cannot be predicted solely from their sequences, so the cluster must be induced in the native host, or expressed and analyzed in a heterologous host[18,98]. Despite the many advantages *S. cerevisiae* has as a heterologous host, introns in the fungal gene cluster coding regions must be completely removed before the heterologous pathway is expressed in yeast due to the significant differences between the introns of yeast and filamentous fungi[18,99]. However, the accuracy of fungal intron prediction has been improved by programs like FGENESH and homology alignment of known related genes to a putative cluster[100,101]. Using intron prediction programs or direct cloning from RNA, if possible, should allow entire clusters discovered through genome mining to be cloned and expressed in yeast. As discussed previously, Ishiuchi *et al* used their ExRec method of cloning from a pool of total RNA isolated from various fungi to express several heterologous PKSs in yeast[60]. Despite this progress, and its success as a platform for the exploration of silent bacterial PKS clusters, *S. cerevisiae* has been underutilized for exploring new silent fungal PKS clusters[54,102,103]. The technology and tools are available to discover new fungal natural products with yeast.

1.8 Simvastatin production; past, present, and future

Lovastatin is a natural polyketide pharmaceutical that was first isolated from the filamentous fungus *Aspergillus terreus* in 1978 and has since become a major drug used to treat hypercholesterolemia[104]. Simvastatin is a semisynthetic compound derived from lovastatin with a single methyl group addition on the side chain, and is one of the most highly prescribed cholesterol lowering medications, with annual sales topping $5 billion before it became generic in 2007, and it is listed as an essential drug by the World Health Organization[4,105,106]. The lovastatin pathway has been well studied through genetic and biochemical characterizations (Fig. 5). LovB, an HR-PKS, and LovC, an enoyl reductase, together synthesize dihydromonacolin L acid (DMLA) from nine malonyl-CoA acyl units, along with a methyl donor, *S*-adenosyl methionine (SAM), and the cofactor, nicotinamide adenine dinucleotide phosphate (NADPH)[71,107]. DMLA is then released from LovB by the thioesterase LovG and converted to monacolin L acid (MLA) and then monacolin J acid (MJA) by the cytochrome P450 monooxygenase (P450) LovA[72,24]. During the multistep oxidation reactions, LovA receives reducing equivalents from its redox partner, the cytochrome P450 reductase (CPR)[24]. The diketide synthase LovF synthesizes the α-S-methylbutyryl-ACP product, and the acyl unit is transferred to MJA by the acyltransferase LovD, completing the synthesis of lovastatin acid (LVA)[71,73].

The semisynthetic analog simvastatin, with a α,α-dimethylbutyryl side chain, is a more effective cholesterol lowering drug and is the active pharmaceutical ingredient of the blockbuster drug Zocor ®. The semisynthesis of simvastatin from lovastatin is a multistep chemical process that inculdes protection and deprotection steps, requiring multiple purification steps and is of poor atom economy[108,109]. These laborious processes contributed to simvastatin being nearly five times the cost of lovastatin, which in turn encouraged the search for a more economical method of producing simvastatin[110]. As

simvastatin differs from the natural product lovastatin by only one methyl group on the polyketide sidechain synthesized by LovF, work was done to investigate the promiscuity of LovD. It was found that the wild type LovD (LovDwt) could accept alternative acyl donors not provided by LovF, leading to production of lovastatin analogs both *in vitro* using purified LovD and *in vivo* when MJA was added to *E. coli* expressing LovD[111]. Simvastatin was successfully produced when an alternative acyl donor, α-dimethylbutyryl-S-methyl-mercaptopropionate (DMB-SMMP), and MJA were provided to LovD expressed in *E. coli*. This whole cell biocatalysis afforded simvastatin[110]. Deletion of *E. coli* native gene *bioh* that was responsible for DMB-SMMP hydrolysis led to increased conversion of MJA to SVA, but LovDwt had low activity towards the synthetic acyl donor[112].

The next steps were focused on improving LovD activity towards this alternative substrate. Further rational mutations to LovD increased its solubility and activity in *E. coli*[113]. Gao *et al* achieved an 11-fold increase in LovD catalytic efficiency through directed evolution, resulting in mutant LovDG7[114]. In addition, LovDG7 was further evolved by Codexis, resulting in a version that was >1000-fold more efficient at SVA synthesis than the LovDwt[115]. This evolved version of the acyltransferase, LovD9, is used to as part of the current industrial semi-synthetic method of producing simvastatin. The natural product lovastatin is produced and purified from *A. terreus*, then hydrolyzed to monacolin J acid (MJA) *in vitro*, and finally combined with DMB-SMMP and cell lysate from *E. coli* containing LovD9[116,115]. This process affords simvastatin, but requires multiple purification steps and the use of two organisms[110,114,115].

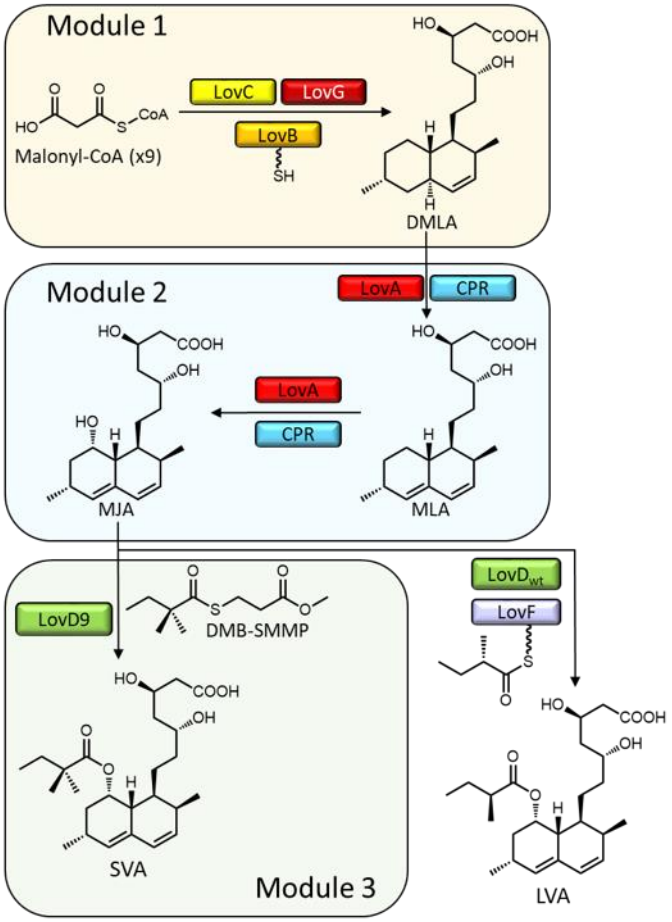

Figure 5: Lovastatin and simvastatin biosynthesis. In Module 1, LovB (polyketide synthase) and LovC (enoyl reductase) use nine malonyl-CoA units to form the polyketide backbone which is released from LovB by LovG (thioesterase), forming DMLA (dihydromonacolin L acid). In Module 2, LovA (cytochrome P450 monooxygenase) catalyzes the conversion of DMLA to MLA (monacolin L acid) and then to MJA (monacolin J acid). In between steps LovA is reduced by its partner CPR (cytochrome P450 reductase). In Module 3, the semisynthetic acyl donor DMB-SMMP (α-dimethylbutyryl-S-methyl-mercaptoproprionate) is transferred to MJA by an evolved version of LovD (acyl transferase) to form SVA (simvastatin acid). In the native pathway, LovF (polyketide synthase) provides the sidechain which is transferred by the wild type LovD to form LVA (lovastatin acid).

There is significant interest in developing a one-step fermentation process for producing simvastatin in a single organism. The native lovastatin host, *A. terreus,* was targeted but has proven resistant to engineering attempts to produce simvastatin directly[4]. *E. coli*, while successfully used to express LovD9, cannot functionally express the polyketide synthase LovB or P450 LovA necessary for *de novo* production of MJA. Refactoring the pathway to MJA in a heterologous yeast host is another potential strategy. *S. cerevisiae*

has successfully been used as a heterologous host to produce 20 mg/L MJA via expression of LovG, LovC, LovB, *A. terreus* CPR, and LovA from three separate 2μ plasmids[72]. Recently, lovastatin was produced in *Pichia pastoris*, but the limited ability of MJA to cross the *P. pastoris* cell membrane implies that this strategy could prove challenging for SVA production involving exogenously added DMB-SMMP[117]. Therefore, the research in this work focused on producing simvastatin in *S. cerevisiae*. The goal was to improve MJA production in *S. cerevisiae* through refactoring of the native *A. terreus* MJA pathway and to achieve one-pot conversion of MJA to SVA through the additional expression of LovD9 and addition of DMB-SMMP (Fig. 5).

1.9 Conclusions

A number of useful synthetic biology strategies have been developed in *S. cerevisiae* that make it a versatile tool for the discovery, characterization, and production of PKSs and their products. Yeast has been used as a host for cloning or purifying protein and as a tool for ascertaining the function of enzymes in a PKS cluster as well as for individual modules of PKSs. In addition, this organism has been engineered as a host for testing biosynthesis of unnatural "natural" products, for screening bioactive compounds to find treatments for specific conditions, and for use as an industrial production host for heterologous pathways.

With such a versatile tool as yeast available, it is remarkable that more work has not been done with fungal PKSs in yeast. The literature is abundant with examples of yeast as a heterologous host for plant and bacterial PKSs, but yeast has been surprisingly underutilized for fungal PKSs. There is a rich variety of fungal natural products left to be discovered, characterized, and engineered. The future of this field will involve a greater utilization of heterologous fungal PKS expression in yeast, especially of cryptic clusters, in

order to aid in the discovery and production of chemically diverse compounds that will have impacts in fields such as fuels and drug discovery.

Though significant improvements have been made for simvastatin synthesis, the current method of industrial production is still not ideal. Two organisms, *A. terreus* and *E. coli*, are required and multiple purification steps. As *A. terreus* has resisted engineering attempts to produce SVA directly, a new heterologous host that can combine MJA synthesis and conversion to SVA in one cell is desired. In this work, *S. cerevisiae* was engineered for production of MJA, and efficient conversion of MJA to SVA, reducing the number of steps and organisms required. Two common laboratory *S. cerevisiae* strains were compared and strain engineering was performed to increase the precursor pool and overall polyketide production. The copy number and integration loci of the P450 *lovA* were tuned and the overall stability and copy number of all three plasmids were characterized. These metabolic engineering strategies increased the production of MJA significantly. Process optimization, including more alkaline pH and optimized cell lysis, were used to increase SVA production, achieving nearly 100% conversion from MJA.

2. Materials and Methods

2.1 Strains and medium

Saccharomyces cerevisiae strains descended from BJ5464-NpgA and BY4741 were compared as biosynthetic hosts[118,119]. Strains and modifications are summarized in Table 1 and in Fig. 6**.** Engineered yeast strains were grown either in SD medium lacking leucine, tryptophan, histidine, and/or uracil where applicable, or in YPD medium (2% dextrose), supplemented with 0.2 g/L hygromycin or G418 when applicable. *E. coli* strains were used for cloning, and followed standard recombinant DNA techniques.

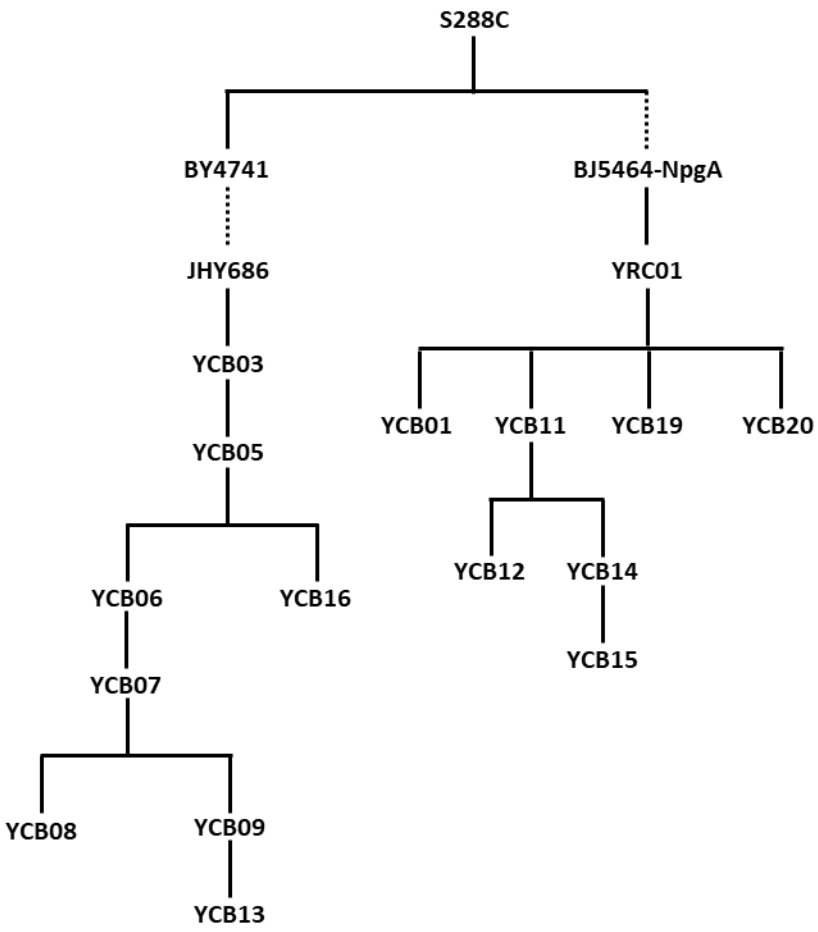

Figure 6: Yeast strain family tree. This tree indicates how the strains built in this work relate to one another. Dashed lines indicate multiple or unknown generations.

Table 1: Yeast strains used in this work.

Strain	Parent	Genotype modifications to parent	Reference
BY4741	S288C	MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0	Brachmann *et al*, 1998[119]
DHY213	BY4741	SAL1+ HAP1+ CAT5(91M) MIP1(661T) MKT1(30G) RME1(INS-308A) TAO3(1493Q)	Harvey *et al*, 2018[120]
JHY686	DHY213	prb1Δ pep4Δ ADH2p-*npgA*-ACS1t lys2Δ0	Harvey *et al*, 2018[120]
YCB03	JHY686	YPRCTy1-2::*iCas9*::leu2	This study
YCB06	YRCB03	ura3Δ0::ADH2p-*CPR*-ADH2t nte1Δ::ADH2p-*lovA*-ADH2t	This study
YCB07	YCB06	pyc2Δ::ADH2p-*lovA*-ADH2t	This study
YCB09	YCB07	HO::ADH2p-*lovA*-ADH2t	This study
BJ5464 -NpgA	S288C	MATα ura3-52 trp1 leu2-Δ1 his3-Δ200 pep4::HIS3 prb1-Δ1.6R can1 GAL	Ma *et al*, 2009[25]
YRC01	BJ5464-NpgA	ura3Δ::ADH2p-*CPR*-ADH2t	Tang *et al*, 2015[121]
YCB11	YRC01	X-2::ADH2p-*lovA*-ADH2t	This study
YCB19	YRC01	pyc2Δ::ADH2p-*lovA*-ADH2t	This study
YCB08	YCB07	yia6Δ::ADH2p-*lovA*-ADH2t	This study
YCB01	YRC01	YDRWTy1-5::ADH2p-*mlcG*-ADH2t::trp1	This study
YCB05	YCB03	ura3Δ0::ADH2p-*CPR*-ADH2t	This study
YCB16	YCB05	pah1Δ::ADH2p-lovA-ADH2t	This study
YCB12	YCB11	X-4::ADH2p-*lovA*-ADH2t	This study
YCB14	YCB11	pyc2Δ::ADH2p-*lovD9*-ADH2t	This study
YCB15	YCB14	nte1Δ::ADH2p-*lovA*-ADH2t	This study
YCB20	YRC01	pyc2Δ::ADH2p-*lovD9*-ADH2t	This study
YCB13	YCB09	X-4::ADH2p-*mlcE*-ADH2t	This study

2.2 Plasmid and strain construction

DNA restriction enzymes were from New England Biolabs, NEB. Q5® High-Fidelity DNA Polymerase (NEB) and AccuPrime™ Pfx DNA Polymerase (Invitrogen™) were used for PCR.

Gene knockouts and integrations were accomplished via CRISPR-Cas9, following the protocol outlined in mode iv from Horwitz *et al*[122]. pCRCT was a gift from Huimin Zhao (Addgene plasmid # 60,621) bearing iCas9[123]. The plasmids bearing gRNAs used antibiotic markers KanMX and HygR. For integrations and knockouts in the BJ5464-descended strains, the plasmid was transformed as three linear pieces: One piece with the iCas9 cassette and two with the backbone and gRNA cassette that overlapped using the gRNA target sequence as a homology arm. The gRNA target sequences were designed using E-CRISP and CRISPRdirect[124,125]. Genome sequences were retrieved from the *Saccharomyces* Genome Database (SGD)[126]. The gRNA cassettes were ordered as gBlocks from IDT. Oligonucleotides were also ordered from IDT. Plasmids assembled via yeast homologous recombination (YHR) during the first knockout at some loci were purified and sequenced. These whole plasmids were used for subsequent modifications at the same loci in other strains and are listed in Table 5.Constructs were confirmed by sequencing (Laragen, CA).

Production plasmids relevant to this paper are listed in Table 2. Plasmids were assembled via yeast homologous recombination, restriction-digestion cloning, or SLIC[127]. For surface expression, the aga1p sequence was amplified using cPCR with YRC01 as the template. The aga2p-linker sequence was ordered on a gBlock from IDT. The leader sequences for protein secretion were also purchased on gBlocks from IDT. The following tables are found in the Appendix. Primers are listed in Table 10. Gene and peptide sequence sources are

listed in Table 7. Sequences of the ordered gBlocks are listed in Table 6. Integration loci and gRNA target sequences are listed in Table 4. Additional plasmids used during construction are listed in Table 5. Sources of sequences used for LovD9 surface display are listed in Table 8. Signal peptide sequences and sources are listed in Table 9.

Table 2: Production plasmids used in this work

Production Plasmid	Description	Reference
plovB	2μ; AmpR; URA3; ADH2p-*lovB*-6xHis-ADH2t	YEpLovB-6His[72]
pGC	2μ; AmpR; TRP1; ADH2p-*lovC*-6xHis-ADH2t; ADH2p-*lovG*-6xHis-ADH2t	pSL05[72]
pCB21	2μ; AmpR; TRP1; ADH2p-6xHis-*mlcG*-ADH2t; ADH2p-*lovG*-6xHis-ADH2t	This study
pCB27	2μ; AmpR; HIS3; ADH2p-6xHis-*mlcG*-ADH2t; ADH2p-*lovG*-6xHis-ADH2t	This study
pCB19	2μ; AmpR; LEU2; ADH2p-*lovA*-6xHis-ADH2t; ADH2p-6xHis-*lovD9*-ADH2t	This study
pCB18	2μ; AmpR; LEU2; ADH2p-*lovA*-6xHis-ADH2t; ADH2p-6xHis-*lovDG7*-ADH2t	This study
pXW161	2μ; AmpR; Leu2; ADH2p-*lovG*-6xHis-ADH2t	Xu *et al*, 2013 [72]
pCB57	2μ; AmpR; URA3; ADH2p-*aga1p*-6xHis-ADH2t	This study
pCB58	2μ; AmpR; LEU2; ADH2p-aga2 leader-*aga2p*-(G4S)3-6xHis-lovD9-ADH2t	This study
pCB59	2μ; AmpR; LEU2; ADH2p-aga2 leader-6xHis-*lovD9*-(G4S)3-*aga2p*-ADH2t	This study
pCB66	2μ; AmpR; LEU2; ADH2p-app8 leader-6xHis-*lovD9*-myc-ADH2t	This study
pCB67	2μ; AmpR; LEU2; ADH2p-appS4 leader-6xHis-*lovD9*-myc-ADH2t	This study
pCB68	2μ; AmpR; LEU2; ADH2p-K28 leader-6xHis-*lovD9*-myc-ADH2t	This study
pCB69	2μ; AmpR; LEU2; ADH2p-synthetic leader-6xHis-*lovD9*-myc-ADH2t	This study

pCB51	2μ; AmpR; TRP1; ADH2p-6xHis-*lovD9*-PRM9t; PCK1p-*lovG*-6xHis-SPG5t; ICL1p-6xHis-*mlcG*-IDP1t	This study, promoters from Harvey *et al.*, 2018[120], terminators from Curran *et al.*, 2013 [128]
pXK32	2μ; AmpR; URA3; ADH2p-*lovF*-6xHis-ADH2t	Xie *et al.*, 2009 [73]
pCB44	2μ; AmpR; LEU2; ADH2p-6xHis-*lovDwt*-ADH2t	This study
pCB50	2μ; AmpR; TRP1; ADH2p-6xHis-*lovDwt*-PRM9t; PCK1p-*lovG*-6xHis-SPG5t; ICL1p-6xHis-*mlcG*-IDP1t	This study, promoters from Harvey *et al.*, 2018[120], terminators from Curran *et al.*, 2013[128]
pCB23	2μ; AmpR; URA3; ADH2p-*lovB*-6xHis-ADH2t; ADH2p-*lovA*-6xHis-ADH2t	This study
pCB26	2μ; AmpR; LEU2; ADH2p-6xHis-*lovDwt*-ADH2t; ADH2p-*lovF*-6xHis-ADH2t	This study
pF3	2μ; AmpR; LEU2; ADH2p-*lovF*-6xHis-ADH2t	This study
pCB15	2μ; AmpR; LEU2; ADH2p-6xHis-*lovD9*-ADH2t	This study
pCB49	2μ; AmpR; LEU2; ADH2p-*mlcE*-ADH2t	This study
pCB05	2μ; AmpR;URA3; ADH2p-6xHis-*lovD9*-ADH2t	This study

2.3 Yeast cultivation and lysis

Transformation of *S. cerevisiae* strains was accomplished by the standard lithium acetate method[129]. Transformants were selected on plates of YPD with 0.2 g/L hygromycin or G418, or on uracil, leucine, histidine and/or tryptophan dropout SD medium. Colonies of JHY686 and descendants were visible 48 hours after plating while colonies of YRC01 and descendants required 72 hours.

Single colony transformants were restreaked onto YPEG or onto SD dropout media if plasmid-bearing. Starter cultures were prepared by inoculating 2 – 3 mL of selective media with a patch of the yeast restreak. This starter culture was shaken at 28 °C and 250 rpm for 24 - 48 h. Next, the appropriate volume of starter culture was spun down and resuspended in 9 mL of YPD to reach OD_{600} = 0.1. This new culture was aliquoted to 3 culture tubes, 3 mL in each, for each condition to be tested. The cultures were shaken at 28 °C and 250 rpm for 72 hours until lysis, unless otherwise noted. If noted in the text,

DMB-SMMP and/or MJA were added to these cultures at the noted timepoints post-inoculation.

For glass bead mechanical lysis, cultures were kept on ice throughout the procedure outside of vortexing periods. 1.5 mL of 0.5 mm diameter glass beads from Biospec Products were added to the 3 mL culture, along with 250 µL of 1 M Tris-HCl buffer pH=8.7. The culture was then vortexed for 30 seconds three times, resting on ice for 1-5 min between vortexing rounds.

For chemical lysis, 250 µL of Y-PER™ Yeast Protein Extraction Reagent from ThermoFisher Scientific or the described amounts of 3-(N,N-dimethylmyristylammonio)propanesulfonate (SB3-14) from Sigma-Aldrich were added to the culture tube, along with 250 µL of 1 M Tris-HCl buffer pH=8.7.

After lysis, 4 or 16 µL 0.2 g/mL DMB-SMMP in DMSO was added to each culture. DMB-SMMP was synthesized as described previously[110]. Finally, cultures were shaken at 28 °C and 250 rpm. Samples were 100 µL, and were taken directly before lysis and 24 hours post-lysis, unless otherwise noted. In the case of YCB09, 1 µL of 0.2 g/mL DMB-SMMP in DMSO was added at 24, 36, 48, and 60 hours post inoculation into YPD and samples were taken 72 hours post inoculation.

2.4 Culture extraction and quantification

In a 1.5 mL microcentrifuge tube, 100-200 µL culture and ~40 µL 0.5 mm glass beads were combined. This was extracted twice with an organic phase of twice the culture volume of 2% trifluoroacetic acid (TFA) in ethyl acetate (EtOAc). Samples were vortexed for 30 seconds and centrifuged for 5 minutes before organic phase was removed. After extraction, the organic phase was dried in a refrigerated CentriVap Concentrator from LabConco for

1 hour. Samples were resuspended in four times the original sample volume of 0.1 M NaOH in methanol (MeOH). For separate cell and media extraction, the cell pellet was spun down and the supernatant aspirated and stored in a separate tube. The two fractions were separately extracted following the above protocol. Chromotography analyses were performed on a Shimadzu 2020 EVLC–MS (Phenomenex kinetex, 1.7 µm, 2.0 x 100 mm, C18 column) using positive and negative mode electrospray ionization with a linear gradient of 5–95% MeCN–H2O supplemented with 0.1% (v/v) formic acid in 15 min followed by 95% acetonitrile (MeCN) for 5 min with a flow rate of 0.3 mL/min. Under this protocol, DMLA eluted at 9.7 min, MLA eluted at 9.1 min, MJA eluted at 7.1 min, and SVA eluted at 13.4 min. Standard curves were produced using DMLA, MJA, and SVA. The DMLA standard curve was used to quantify MLA by assuming an equivalent mass response factor. All measurements were taken in biological triplicate unless otherwise noted.

2.5 Growth rate studies

A patch of cells from a single-colony restreak of the appropriate culture was inoculated into 2 -3 mL starting culture YPD. This starter culture was shaken at 28°C and 250 rpm for 24 h. Next, 100 µL of YPD was inoculated to OD_{600} = 0.01 in wells in a Corning96fc UV transparent plate with the applicable conditions. The plate was covered with Microseal® 'B' PCR Plate Sealing Film, adhesive, optical #msb1001 from BioRad. Growth curves were measured on a Tecan Infinite M200 Pro plate reader. The program temperature was set for 29.0°C, with 28.5°C minimum and 30.0°C maximum. The program was run for 99 cycles of 120 seconds of linear 5 mm shaking (330.5 rpm), absorbance was measured at 600 nm, then 1000 seconds of orbital 3 mm shaking (218.3 rpm). Conditions were tested in biological triplicate, and the means were reported.

2.6 Plasmid copy number measurements

Colony forming units (CFU) assays were performed by removing 5 µL samples of culture, diluting the cultures 20,000-fold in sterile water, and plating 5 µL on 60 mm diameter plates containing either YPAD or selective minimal media with uracil, leucine, or tryptophan dropouts. YPAD plates were done in technical duplicate and dropout plates were done in technical singlet. Each timepoint was done in biological triplicate.

Total DNA extraction was performed using the method from Lõoke *et al*[130]. The volume of culture sample taken was equivalent to one OD_{600} unit. An OD_{600} unit of 1 is the number of cells dissolved in 1 mL that gives $OD_{600} = 1$. Samples were spun down for 30 seconds at top speed and supernatant was aspirated. The resulting pellet was frozen at -20°C. Later the pellet was resuspended in 100 µL water and 10 µL of the resuspension was used for the extraction. The concentration of the resulting extracted DNA was measured and 2 ng was used per qPCR reaction, performed using Luna® Universal qPCR Master Mix from NEB on a Bio-Rad CFX96™ Real-Time System and C1000™ Thermal Cycler. Results were quantified using the standard curve method. The reference gene used was *alg9*. The targeted genes representing the three plasmids were *lovB*, *lovG*, and *lovD9*. The primers used for qPCR are in Table 10.

2.7 Crude protein extraction and SDS-Page/Western Blot

Eight 50 mL YPD (1% Dextrose) cultures of YRC01 + pCB05 were inoculated to $OD_{600} = 0.1$. At each timepoint, one of the cultures was spun down and the cell pellet frozen. Later all of the samples were thawed and each resuspended in 8 mL yeast lysis buffer. Cells were lysed using sonication and centrifuged to get a clean supernatant containing soluble proteins. The soluble proteins from each sample were separated by SDS-PAGE using 12% Bi-tris gels and transferred to nitrocellulose membranes. Monoclonal anti-polyHistidine

antibody produced in mouse clone HIS-1 (Sigma-aldrich) was used to detect the 6xHis tag on the N-terminus of LovD9 through standard techniques and this antibody was detected by Goat anti-Mouse IgG-Fc Fragment HRP-conjugated Bethyl A90131-P (Bethyl laboratories). Imaging was accomplished using Clarity ECL Substrate (BioRad).

2.8 Cloning of mlcE

The sequence for *mlcE* was amplified from cDNA acquired from *Penicillium citrinum*. *P. citrinum* was cultured for 72 hours in MBG3-8 liquid medium[131]. Samples of the mycelia were taken at 48 and 72 hours after inoculation and the mRNA was extracted using RiboPure™ RNA Purification Kit, yeast from Invitrogen. The cDNA was acquired from this sample using oligo$(dT)_{20}$ as the primer. The resulting cDNA was used as the template for PCR amplification of *mlcE*. The primers can be found in Table 10. The resulting product was cloned into pCB49 and sequenced using standard cloning techniques.

3. Results and Discussions

3.1 Initial production of simvastatin in YRC01

To construct an SVA production strain, the proposed biosynthetic route was split into three modules as shown in Fig. 5. Module 1 concerns the synthesis of DMLA and requires the incorporation of *lovB, lovG,* and *lovC*. Module 2 concerns the conversion of DMLA to MLA and MJA and requires incorporation of *lovA* and *CPR*. The final module requires incorporation of an evolved version of *lovD* and addition of DMB-SMMP to afford SVA. YRC01, a descendent of BJ5464-NpgA with the *A. terreus CPR* was integrated in the genome at the *ura3* locus was chosen as the initial base strain[121]. For simplicity, the initial attempt to produce SVA in YRC01 used almost entirely episomal expression. High-copy 2μ plasmids have been successfully used to express PKSs and heterologous pathways in yeast before, both for pathway elucidation and for some metabolic engineering strategies,

though typically with at most two separate high copy plasmids ranging from 6 to 15 kb[4,21,132-134]. ADH2p is a strong promoter that is repressed in the presence of glucose and activated in the presence of ethanol[135]. Reports from different groups have described that use of ADH2p for episomal expression leads to greater and more sustained expression and can even allow successful production from plasmid-based genes through fermentation in complex media[12,105,106,132]. Therefore, the ADH2 promoter and terminator were chosen for all the gene cassettes.

As discussed previously, LovDwt was first evolved in the Tang lab for increased affinity to DMB-SMMP. The resulting mutant, LovDG7, evolved further by Codexis into another LovD variant, LovD9, that has a ~ 1000-fold increase in activity towards DMB-SMMP over LovDwt[115]. However, in both circumstances selection pressures focused on improved performance in *E. coli*. Codexis also codon-optimized *lovD9* for use in *E. coli*, which could lead to decreased performance in yeast. Therefore, YRC01 was transformed with three plasmids expressing the MJA pathway and either LovD9 (pCB19) or LovDG7 (pCB18). The cultures were grown for 58 hours and exogenous DMB-SMMP was added at 46, 50, and 54 hours post inoculation to a final concentration of 0.879 mM. At 58 hours post inoculation, YRC01+pLovB+pGC+pCB19 had clearly produced more SVA than YRC01+pLovB+pGC+pCB18, so *lovD9* was used for all future strain building (Fig. 7). Our next concern was the stability and amount of LovD9 we could expect from episomal expression on 2μ plasmids. A western blot was performed on extracts from yeast expressing pCB19 and LovD9 was still present up to 108 hours after inoculation in complex media (Fig. 8).

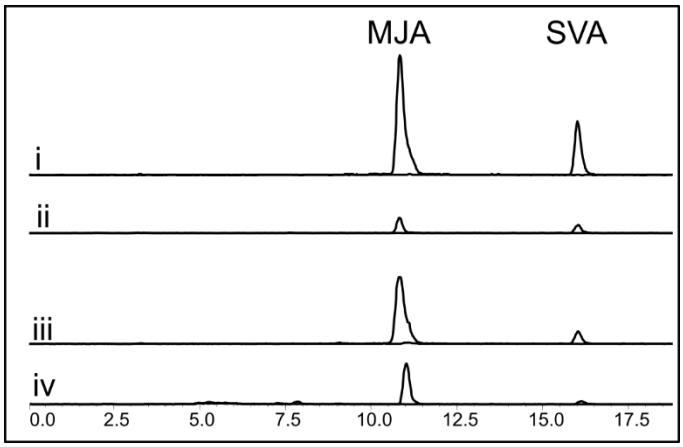

Figure 7: LC-MS traces from YRC01+pLovB+pGC+pCB19 and YRC01+pLovB+pGC+pCB18.The media and cell extracts from YRC01+pLovB+pGC+pCB19 are i and ii, respectively. The media and cell extracts from YRC01+pLovB+pGC+pCB18 are iii and iv, respectively. The cultures were grown for 58 hours and exogenous DMB-SMMP were added at 46, 50, and 54 hours post inoculation to a final concentration of 0.879 mM.

In order to optimize DMLA production, we compared *lovC* with a related homologue. The mevastatin (also called compactin) cluster of *Penicillium citrinum* has been found to contain many genes with significant homology and identical putative functions to those in the lovastatin pathway[131,136]. The enoyl reductase *mlcG,* homologue of *lovC* in *P. citrinum,* has been shown to be functionally equivalent to *lovC* in the lovastatin pathway[131,25]. DMLA production in *S. cerevisiae* using *mlcG* instead of the *A. terreus* native *lovC* led to 50% increase in DMLA titer (Fig. 9), therefore we chose to express *mlcG* in Module 1. YRC01 was transformed with the resulting set of plasmids. The resulting strain, YRC01+pLovB+pCB21+pCB19, was inoculated in YPD at OD_{600} =0.1 and grown for 72 hours. DMB-SMMP was added at 24, 36, 48, and 60 hours after inoculation to a final concentration of 1.22 mM. Samples taken at 72 hours after inoculation showed significant production of MJA, but less than 2 mg/L SVA (Fig. 10).

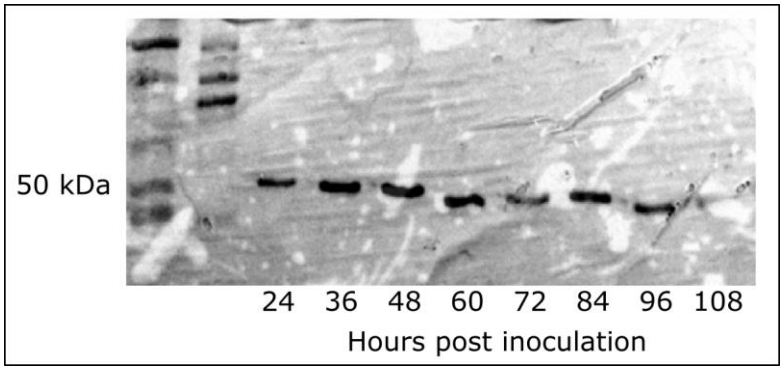

Figure 8: Western blot of LovD9 after expression from YRC01+pCB19 in YPD. Separate cultures were inoculated from the same YRC01+pCB19 seed culture and cultured in YPD for 24, 36, 48, 60, 72, 84, 96, or 108 hours before sampling. The first two columns on the left are the protein ladders.

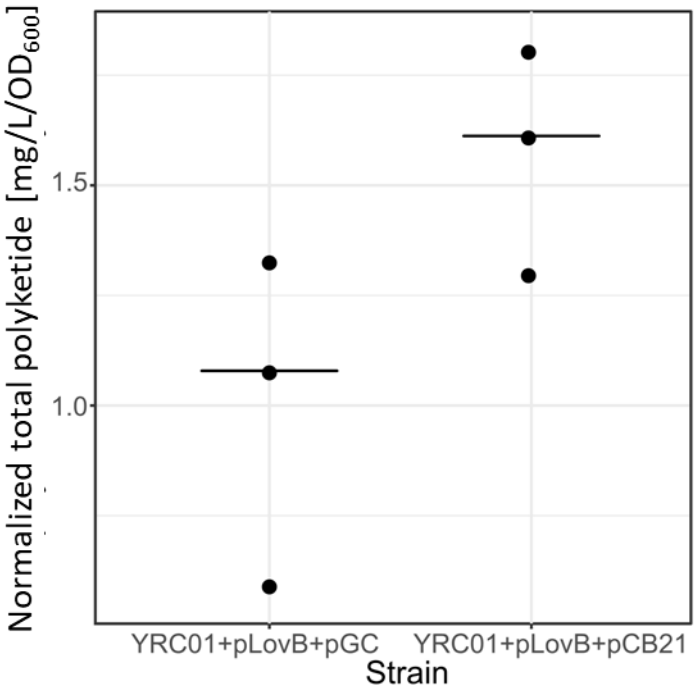

Figure 9: Comparison of DMLA production from MlcG and LovC. DMLA production in YRC01 from plovB combined with either pCB21 or pGC, normalized by the OD_{600}. Each point is a different sample and the black horizontal lines represent the median.

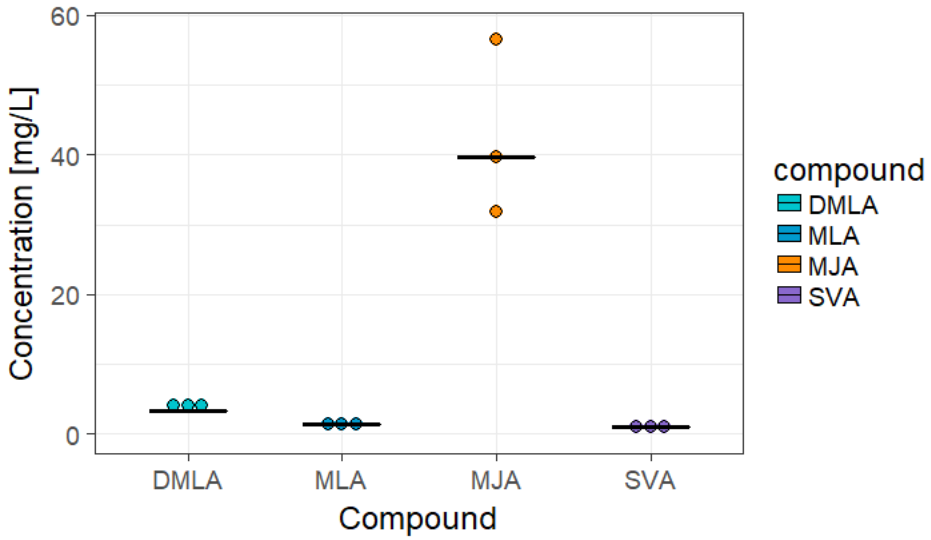

Figure 10: Breakdown of polyketide production in YRC01+pLovB+pCB21+pCB19. This culture was inoculated to OD_{600} = 0.1 in YPD and grown for 72 hours. DMB-SMMP was added at 24, 36, 48, and 60 hours after inoculation to a final concentration of 1.22 mM. Each point is a different sample and the black horizontal lines represent the median.

3.2 Integration of *mlcG* to reduce plasmid burden

Other studies have examined the stability and expression of high and low copy yeast plasmids compared to genomic integrations. Simultaneous expression of three separate genes is 15 times as frequent when integrated into three genomic loci versus expressed on three 2μ plasmids[13]. In addition, 2μ plasmids can have erratic expression levels, and in the case of multiple plasmids, high expression from one plasmid may correlate with low expression from the other plasmid[137]. Expression and copy number of high-copy plasmids can also be strongly affected by the genes encoded on the plasmid, including the promoter, the target gene and the marker used on the vector[110,115,116]. Precise control over gene expression level can be vital to successful metabolic engineering in yeast, so reducing the number of high copy plasmids used was of interest[138,139].

Module 1 was the first target for integration. LovB catalyzes over thirty separate reactions in the synthesis of DMLA, so it was anticipated that the *lovB* copy number would be the limiting factor in DMLA synthesis relative to copy numbers of its partner enzymes, *mlcG*

and *lovG* [25]. However, integration of a single copy of *mlcG* in the YDRWTy1-5 locus of YRC01 led to a 90% decrease in DMLA production when this strain, YCB01, was transformed with two plasmids expressing *lovB* (pLovB) and *lovG* (pXW161) (Fig. 11). This indicated that significantly more copies of *mlcG* would need to be integrated to reach these similar levels of production, so all of Module 1 was kept 2μ plasmids pLovB and pCB21.

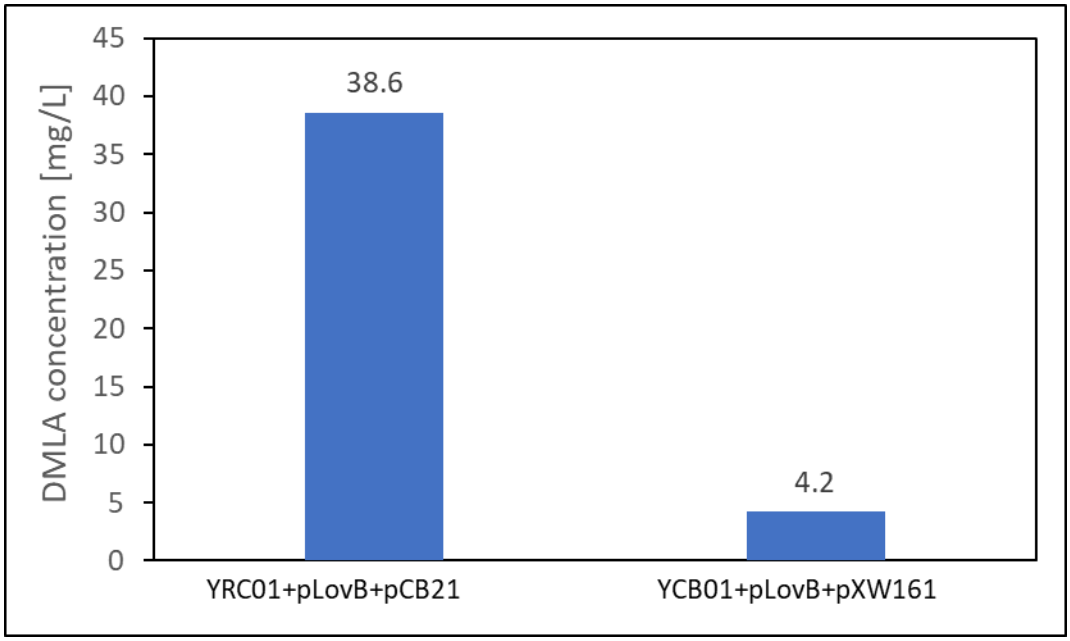

Figure 11: Comparison of DMLA production from integration or 2μ expression of *mlcG*. Production of DMLA from YRC01+pLovB with either pXW161 (2μ, *lovG*) and one integrated copy of *mlcG* and with pCB21 (2μ, *lovG, mlcG*). Data results from a single sample.

3.3 MJA pathway construction and production of SVA in *S. cerevisiae* BY4741 series

YRC01 was chosen as the initial host due to already having the necessary modifications for production of polyketides. However, YRC01 grows relatively slowly, with a doubling time of approximately 2.37 hours, and requires nearly 72 hours post-transformation for colonies to appear. As the metabolic engineering process has frequently been described as repeated cycles of design, build, test, and learn[140], potential improvements could be more

rapidly investigated by shortening the "build" portion of the cycle. In addition, the whole genome sequence for BJ5464 is not published, hindering the choice and design target loci and homology for genome editing. The whole genome sequence of BY4741 is publicly available, but this strain does not have the necessary modification for expression of polyketide synthases and a significantly higher tendency towards petites[120]. Therefore, JHY686, which was recently developed from the BY4741 series at the Stanford Genome Technology Center (SGTC), was selected as the parent strain for testing. This strain has repaired petite tendency, increased respiration growth rate, the necessary protease mutations, and integration of the phosphopantetheinyl transferase gene *npgA* to enable posttranslational modification of the expressed PKSs[120]. JHY686 has a doubling time of approximately 1.62 hours and requires only approximately 48 hours post-transformation for colonies to appear, thus significantly decreasing the time required for the "build" step of the metabolic engineering cycle. To further accommodate rapid genome edits in this strain, we integrated single copy of iCas9 in the YPRCTy1-2 locus using leu-marker mediated integration, forming YCB03[123]. Future integrations and knockouts were then accomplished via transformation of the gRNA cassette and expression backbone as two separate pieces, as discussed in mode iv of Horwitz *et al*[122] YCB03 was transformed with the Module 1 plasmids pLovB (*lovB*) and pCB27 (*mlcG, lovG*), leading to an initial DMLA titer of 10 mg/L (Fig. 12A). Though this was lower than the initial polyketide production in YRC01+pLovB+pCB21+pCB19 (Fig. 10), it was high enough that this strain would still be sufficient as a rapid testing strain for polyketide production improvements.

The next step was to incorporate Module 2, *lovA* and *CPR*, into the DMLA-producing strain to afford conversion to MLA and MJA. Like many P450s, LovA is targeted to the endoplasmic reticulum (ER)[141,24]. Recently, Trenchard *et al* showed that expression of an ER-targeted P450 from a high copy plasmid can cause ER membrane proliferation stress

and lower production of the desired heterologous product[142]. To avoid this, Module 2 was integrated into the genome of the production strains. The *A. terreus CPR* was integrated at the *ura3* locus in YCB03, forming YCB05. Knocking out *pyc2* (involved in reverse glycolysis), *nte1* (involved in lipid biosynthesis), or *yia6* (involved in transport of NAD^+ to the mitochondria) has been shown to increase the acetyl-CoA pool in yeast[63]. As acetyl-CoA is a precursor to malonyl-CoA, this strategy had potential to improve production of DMLA as well. Therefore, *nte1, pyc2*, and *yia6* were knocked out sequentially in YCB05 and replaced the *lovA* expression cassette, forming YCB06, YCB07, and YCB08, respectively. After transformation with the Module 1 plasmids, YCB06+pLovB+pCB27 increased total polyketide production to 20 mg/L, and YCB07+pLovB+pCB27 further increased polyketide production to 25 mg/L (Fig. 12, 13). YCB08, with the YIA6 knockout led to the lowest polyketide production, less than 1 mg/L (Fig. 13).

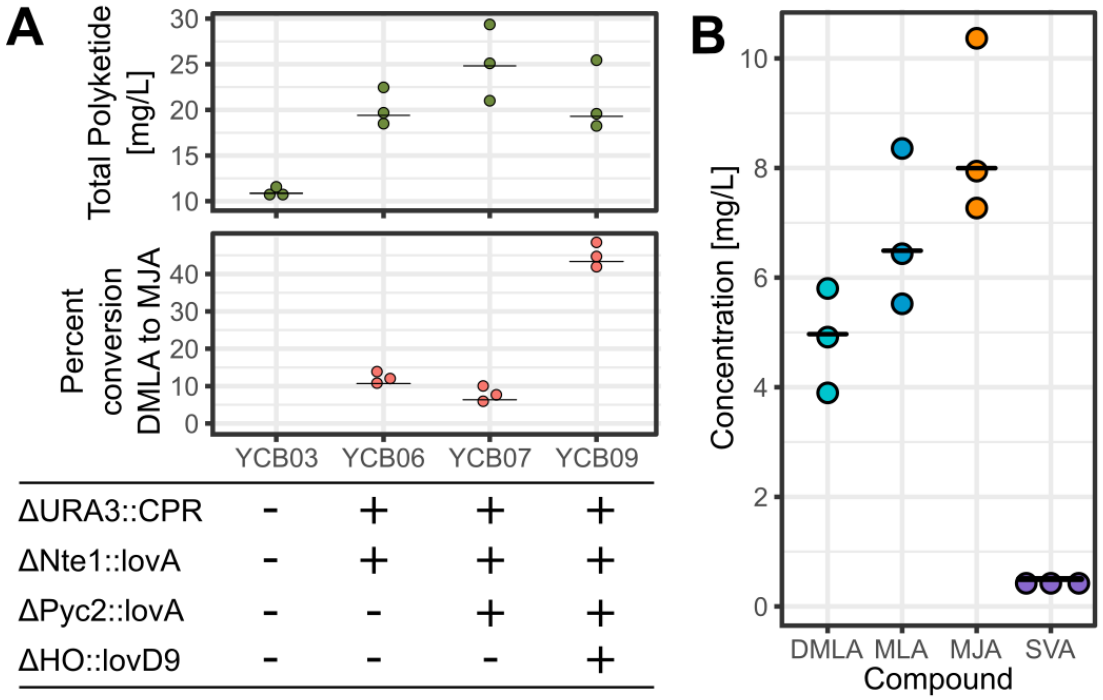

Figure 12: Polyketide production in the BY4741-descended strains. All strains contain Module 1 plasmids pLovB and pCB27 (A) Total polyketide production in green. The percent conversion of DMLA to MJA is in orange. Each point is a different sample. Black horizontal lines indicate median values. (B) Breakdown of polyketide production in YCB09. Results in triplicate. Black horizontal lines represent the median.

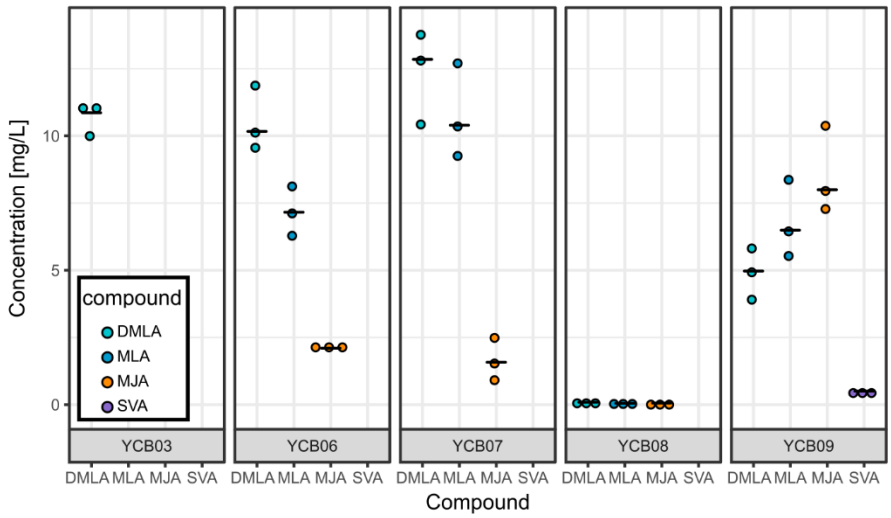

Figure 13: Breakdown of polyketide production from YCB03, YCB06, YCB07, YCB08, and YCB09. All strains were transformed with Module 1 plasmids pLovB and pCB27. DMB-SMMP was added to YC09 throughout the 72 hour culture to a final concentration of 1.22 mM. Each point is a different sample. Black horizontal lines indicate median values.

While replacing *nte1* and *pyc2* with copies of *lovA* did double total polyketide titer, conversion from DMLA to MJA was only 5-10% (Fig. 12A). Assuming the relative amounts of DMLA, MLA, and MJA were at equilibrium, it was hypothesized that further converting MJA to SVA would increase the flux of DMLA to MJA. To test this, a copy of *lovD9* was integrated in the HO locus of YCB07, forming YCB09[115]. YCB09+pLovB+pCB27 was cultured in YPD and DMB-SMMP was added at 24, 36, 48, and 60 hours post inoculation to a final concentration of 1.22 mM. While expressing LovD9 increased the percent conversion of DMLA to MJA to 44%, the final SVA titer was only 0.5 mg/L and total polyketide production decreased to 20 mg/L (Fig. 12A, B). The growth of these strains improved with increased knockouts, as is seen in Fig. 14. YCB08 had the shortest doubling time and reached the highest OD_{600} and yet produced significantly less polyketide. This could be due to more carbon being used for cell replication rather than polyketide production (Fig. 13, 14).

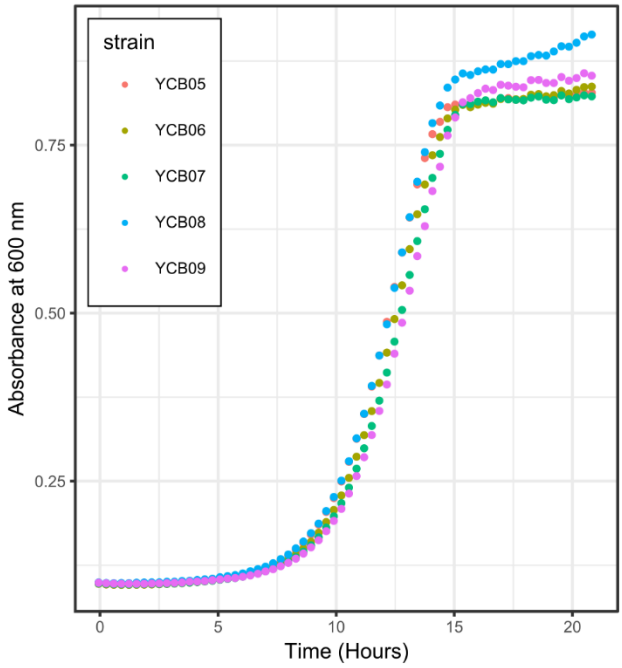

Figure 14: Growth curves of YCB05, YCB06, YCB07, YCB08, and YCB09 in YPD. All points are the means of triplicate cultures.

Another possible method of improving the conversion of DMLA to MJA could be to increase the ER available for the P450 LovA to anchor in. This has been accomplished in the past via knockout of *pah1*[143]. Therefore, *pah1* was deleted in YCB05 and replaced with a copy of *lovA*, forming YCB16. However, this strain could no longer grow on histidine-deficient media, even with plasmids that included the missing *his3* gene such as the Module 1 plasmid pCB27 (Fig. 15). Arendt *et al.* had also deleted *pah1* in a strain descended from BY4741 but did not attempt to grow this strain on histidine dropout media, so it is unknown if this deficiency also existed in their *pah1* deletion strain[143].

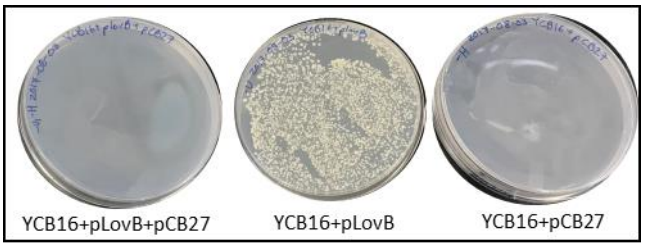

Figure 15: Growth of YCB16 on selective media. From left to right: YCB16+pLovB+pCB27 on uracil and histidine dropout media, YCB16+pLovB on uracil dropout media, and YCB16+pCB27 on histidine dropout media.

3.4 Expression of statin-specific efflux pump MlcE

Bioactive secondary metabolites like fungal polyketides are often toxic in high concentrations. In the case of statins, the benefit of lovastatin and its analogs as pharmaceuticals is their inhibition of 3-hyrdroxy-3-methylglutaryl coenzyme A reductase (HMG-CoA reductase). However, this also leads to growth inhibition in yeast[83,144]. While the toxicity of the final product presents a challenge to heterologous production in yeast, many native hosts have evolved resistance mechanisms to their own natural products[80]. One such self-resistance mechanism is the statin-specific efflux pump *mlcE* from the compactin cluster of *P. citrinum*[58,83,145,146]. Expression of a codon-optimized version of *mlcE* in *S. cerevisiae* by Ley *et al.* conferred increased resistance specifically to mevastatin, lovastatin, and simvastatin[83]. The MlcE-expressing strain had restored yeast growth through at least 800 mg/L exogenously added lovastatin, significantly above the 200 mg/L lovastatin that greatly inhibits wild type yeast growth[83]. Due to these results, *mlcE* had significant potential to increase SVA production from YCB09+pLovB+pCB27. In addition to reducing toxicity, exporting the synthesized SVA from the cells could reduce the LovD-catalyzed hydrolysis of SVA by decreasing the intracellular SVA concentration[147].

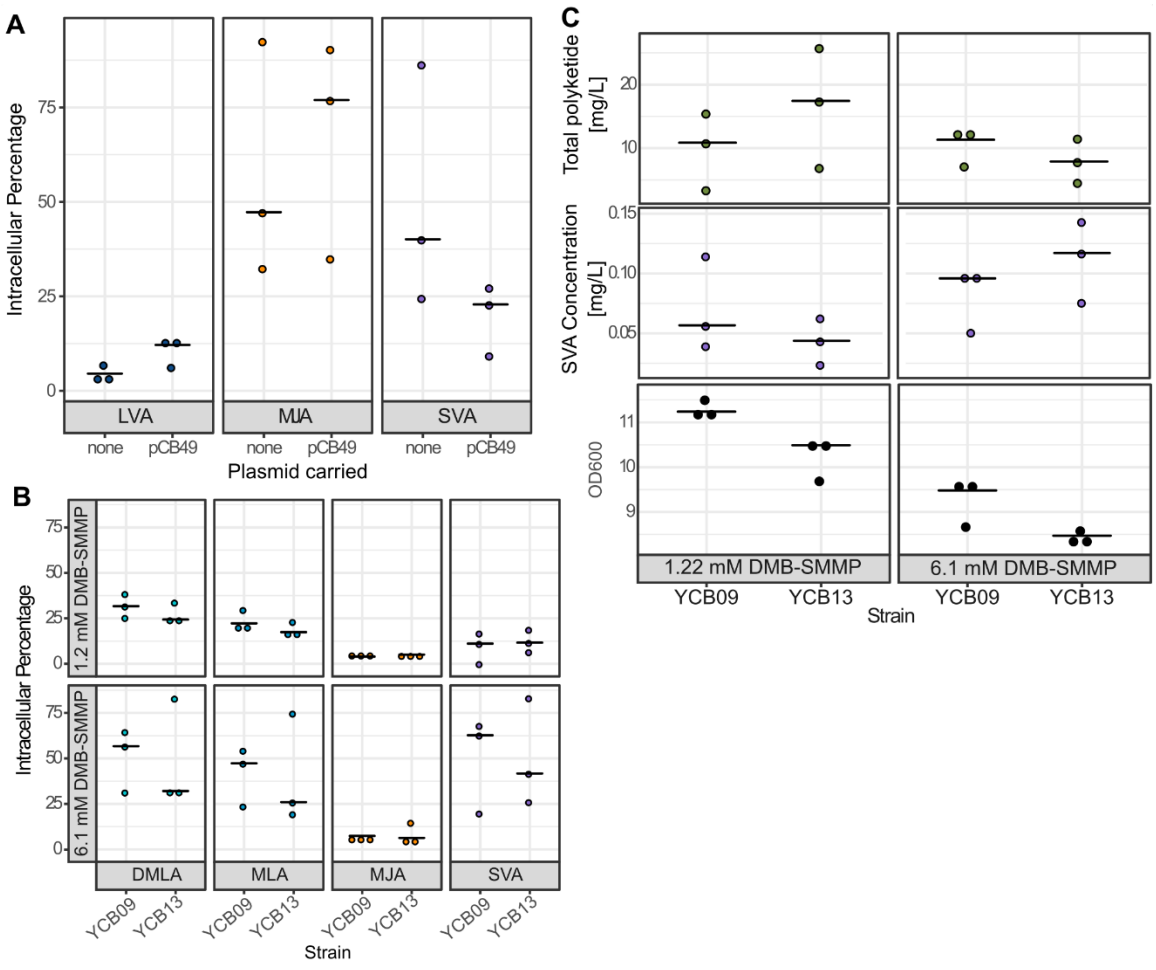

Figure 16: Polyketide production and localization results from expression of *mlcE*. (A) Percentage of LVA, MJA and SVA that was intracellular after exogenous addition to cultures of YRC01 or YRC01+pCB59. (B) Intracellular percentages of polyketides produced by YCB09+pLovB+pCB27 or YCB13+pLovB+pCB27 with addition of DMB-SMMP to a final concentration of 1.22 mM or 6.1 mM. (C) Polyketide production and OD_{600} of YCB09+pLovB+pCB27 or YCB13+pLovB+pCB27 with addition of DMB-SMMP to a final concentration of 1.22 mM or 6.1 mM. Each point is a different sample. Black horizontal lines indicate median values.

Before incorporating *mlcE* into an SVA producing yeast strain, the promiscuity of MlcE towards the SVA precursor MJA was tested, as well as the effect of MlcE at the current levels of polyketide production in YCB09+pLovB+pCB27. If MlcE also exported MJA out of the cells, then expressing *mlcE* behind the same promoter as the MJA pathway genes could lead to decreased conversion from MJA to SVA. First, the sequence of *mlcE* was retrieved from the cDNA of *P. citrinum* and was cloned into an expression cassette with ADH2p and ADH2t. A plasmid, pCB49, expressing this cassette was transformed into

YRC01 and this strain was inoculation in YPD and cultured for 72 hours before sampling. At 24, 36, and 48 hours after inoculation, MJA , LVA, and SVA were exogenously added to a final concentration of 30 mg/L MJA, 3 mg/L LVA, and 3 mg/L SVA. The percentage of the exogenously added compound that was extracellular in YRC01 versus YRC01+pCB49 was compared in Fig. 16A. Though LVA and MJA did not show a significant different in their localization patterns after the addition of pCB49, the intracellular percentage of exogenously added SVA decreased by approximately 15% in YRC01+pCB49 compared to YRC01.

The next step was to test if MlcE would increase *de novo* production of SVA. A copy of the *mlcE* cassette was integrated in the X-4 locus in YCB09, forming YCB13. Polyketide production was compared in YCB09+pLovB+pCB27 and YCB13+pLovB+pCB27 with two different levels of DMB-SMMP concentrations. At 24, 36, 48, and 60 hours after inoculation, DMB-SMMP was added to a final concentration of 1.22 mM or 6.1 mM. In the both sets of cultures, the percentage of intracellular DMLA, MLA, MJA, and SVA showed no significant difference between YCB09+pLovB+pCB27 and YCB13+pLovB+pCB27 (Fig. 16B). Comparison between the different concentrations of DMB-SMMP added indicated that the intracellular percentage of SVA increased by 20-40% with increased DMB-SMMP in both YCB09+pLovB+pCB27 and YCB13+pLovB+pCB27, though all strains and conditions produced less than 0.5 mg/L SVA (Fig. 16B,C). This experiment also provided insight into the effect of adding DMB-SMMP throughout the production culture growth time. Increasing the final concentration of DMB-SMMP from 1.22 mM to 6.1 mM decreased the total polyketide titer of YCB13+pLovB+pCB27 from 17 mg/L to 8 mg/L YCB13, though YCB09+pLovB+pCB27 produced 11-12 mg/L total polyketide under both conditions (Fig. 16C). Even more striking was the effect on the maximum OD_{600} reached by the cells. For both YCB09+pLovB+pCB27 and YCB13+pLovB+pCB27, the OD_{600} at 72

hours post inoculation dropped by 17-19% when the concentration of DMB-SMMP was increased (Fig. 16D). These results indicate that DMB-SMMP may inhibit yeast growth. The final OD_{600} of YCB13+pLovB+pCB27 was lower under both DMB-SMMP conditions, indicating that expression of MlcE or integration into the X-4 locus may also inhibit yeast growth.

Though high copy expression of MlcE from a 2μ plasmid did decrease the intracellular percentage of exogenously added SVA by 15%, integration of *mlcE* into an SVA production strain did not lead to a significant increase in SVA titers or a change in the intracellular percentage of any of the produced polyketides. There are several possible explanations for this discrepancy. First, Ley *et al.* used a version of *mlcE* that was codon-optimized for expression in *S. cerevisiae* while the version used in this work was taken directly from the cDNA of *P. citrinum*[83]. Codon-optimization significantly increased the abundance of LovA in *S. cerevisiae*, so the differences in codons could explain the small effect on localization from the wild-type version of *mlcE*[24]. In addition, the amount of SVA produced by YCB09+pLovB+pCB27 was below 0.5 mg/L SVA. It is possible that at significantly higher levels of SVA production expression of MlcE could lead to a more significant difference in the final SVA titer.

Based on these initial results from testing in the more rapid "build" JHY686 descended strains, we further focused on each of the modules. First, increasing the total amount of polyketide produced. Second, improving conversion through the pathway to MJA, and third, improving conversion from MJA to SVA.

3.5 Optimization of Total Polyketide Production and Conversion to MJA

Utilizing the successful strategies gleaned from strain modifications in JHY686, polyketide production was returned to YRC01. Comparison of these strains showed that

YRC01+pLovB+pCB21 produced twice as much polyketide as YCB03+pLovB+pCB27, indicating that incorporating the successful changes from the JHY686 tests could lead to significantly more total polyketide (Fig. 17). JHY686 has a doubling time of 1.62 hours compared to the 2.37 hour doubling time of YRC01, so this discrepancy in polyketide production could be due to JHY686-descended strains having higher carbon flux towards growth than polyketide production. As in the JHY686 series, CRISPR-Cas9 was used for genomic modifications in YRC01. To avoid any marker-mediated integrations and to reduce any effect caused by genomic expression of iCas9 had on polyketide production, in this case iCas9 and the gRNA cassette were expressed from a single plasmid. To accomplish this, four linear pieces of DNA were transformed into the strain of interest. In addition to the donor DNA with appropriate homology for the targeted locus, strains targeted for genome edits were transformed with three linear pieces of DNA: the iCas9 cassette and two pieces composing the of the backbone and gRNA cassette for targeting the desired locus, all with the appropriate homology for recombination into one plasmid.

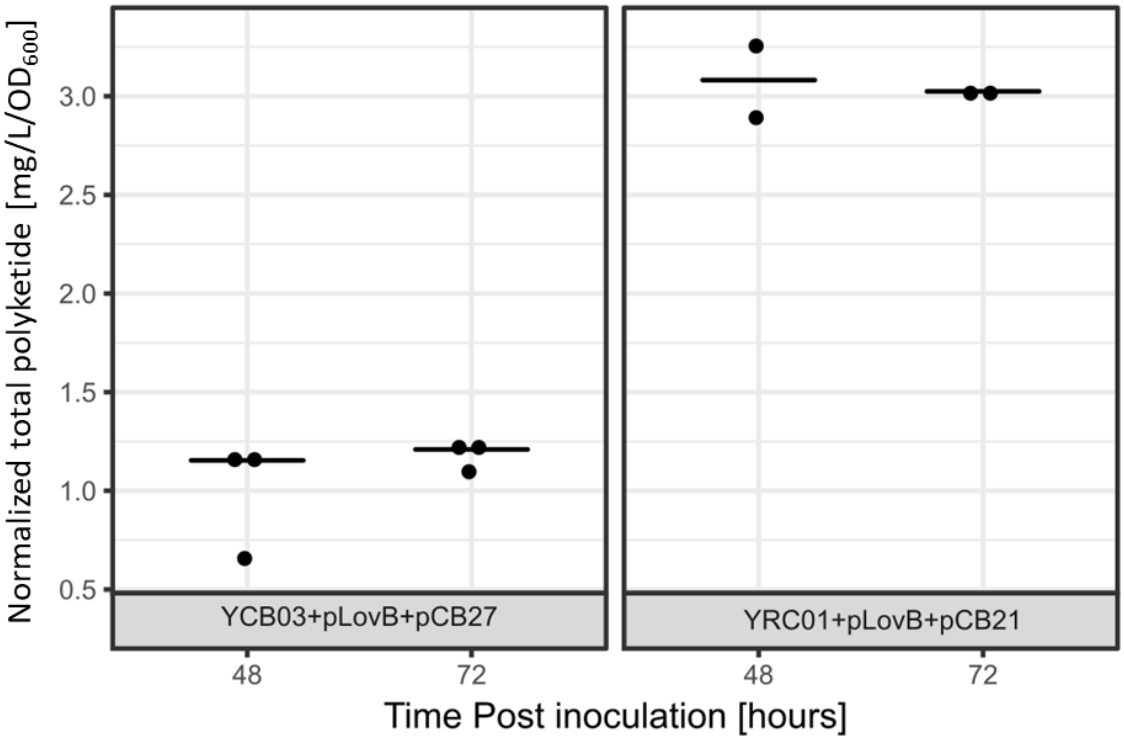

Figure 17: Comparison of polyketide production from YCB03 and YRC01. Cultures of YCB03+pLovB+pCB27 and YRC01+pLovB+pCB21 were grown in YPD and the polyketide titer, normalized by the OD_{600}, was measured at 48 and 72 hours after inoculation. Each point is a different sample and the black horizontal lines represent the median.

One copy of *lovA* was integrated in the X-2 locus[148] of YRC01 with no associated gene deletion, yielding strain YCB11. After transformation with the Module 1 plasmids, YCB11+pLovB+pCB21 had a final titer of 30 mg/L polyketide compared to the 65 mg/L produced by YRC01+pLovB+pCB21 (Fig. 18). This decrease was of similar magnitude to the decrease seen when pCB19, a 2μ plasmid expressing *lovA*, was transformed into YRC01 along with the Module 1 plasmids. The percent conversion from DMLA to MJA was less than 75% in YCB11+pLovB+pCB21, so a second copy of *lovA* was integrated in the X-4 locus of YCB11, yielding YCB12. This strain produced only 6 mg/L total polyketide, with only 50% conversion from DMLA to MJA, an even further decrease from YCB11+pLovB+pCB21. In the JHY686 series, no similar drop in production was seen when

copies of *lovA* were integrated with simultaneous deletions of *pyc2* and *nte1*, so *pyc2* and *nte1* were deleted next in YCB11 to attempt to rescue production (Fig. 18).

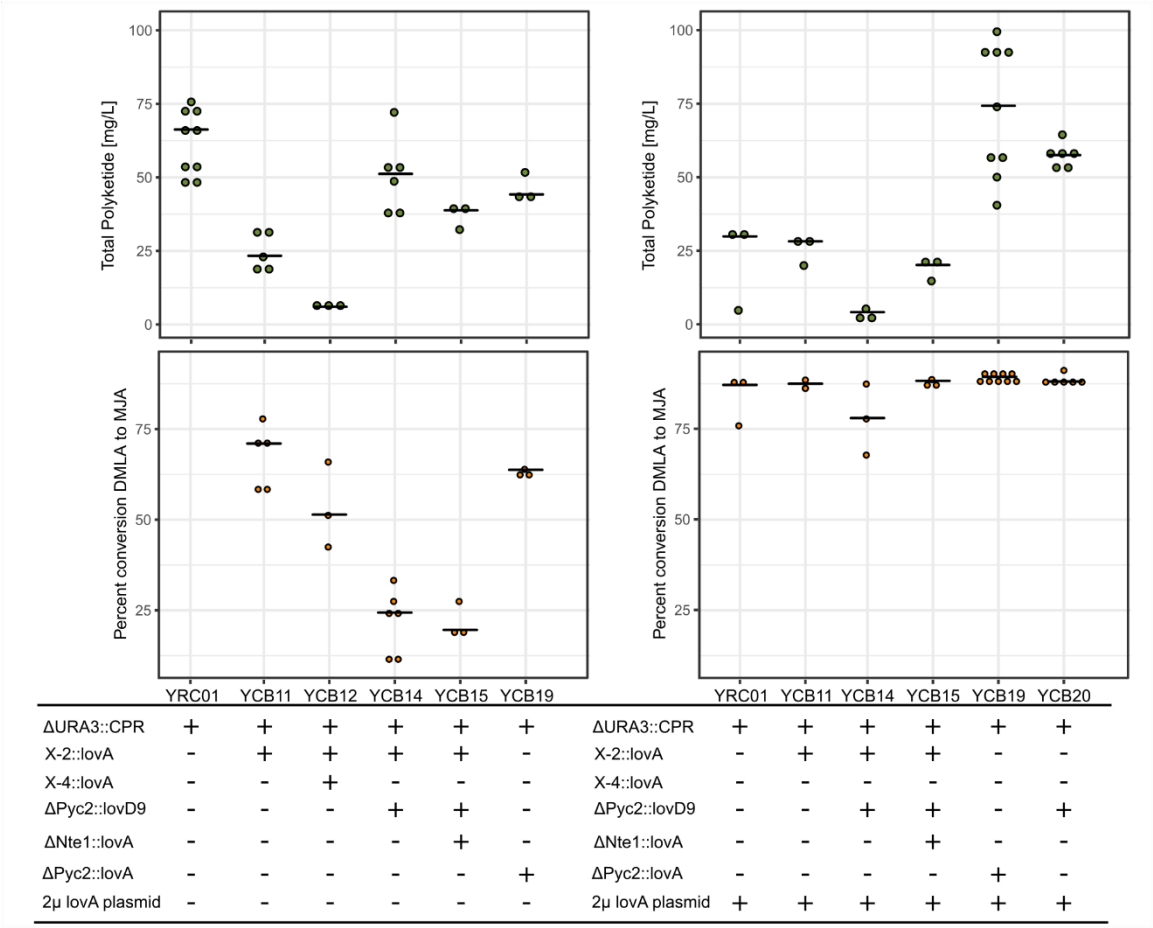

Figure 18: Polyketide production in the BJ5464-descended strains. All strains contained the Module 1 plasmids, pLovB and pCB21. For strains that contain the 2μ lovA plasmid, this was pCB60 for YCB20 and pCB19 for all other strains. Total polyketide production in green. Percent conversion from DMLA to MJA is in orange. Each point is a different sample. Black horizontal lines indicate median values.

YCB14 was the result of deletion of *pyc2* and simultaneous integration of *lovD9* in that locus in YCB11. This doubled total polyketide production relative to YCB11 (Fig. 18). The percent conversion of DMLA to MJA decreased significantly, but that could be due to the limits of one copy of *lovA* in the presence of a much higher concentration of DMLA. To improve conversion from DMLA to MJA, and to continue improving total polyketide production, *nte1* was deleted and another copy of *lovA* was integrated in that locus,

yielding YCB15. After being transformed with the Module 1 plasmids, this strain produced similar amounts of polyketide to YCB14 and the percent conversion of DMLA to MJA was not increased (Fig. 18). It was hypothesized that the X-2::*lovA* integration may have contributed to the lower production and conversion to MJA. Returning to YRC01, *pyc2* was deleted and replaced it with *lovA* or *lovD9*, yielding strains YCB19 and YCB20, respectively. When transformed with the Module 1 plasmids, YCB19+pLovB+pCB21 produced a similar amount of total polyketide as YCB14+pLovB+pCB21 and YCB15+pLovB+pCB21, ~40 mg/L, but had a 63% percent conversion from DMLA to MJA, significantly higher than was seen in YCB14+pLovB+pCB21 and YCB15+pLovB+pCB21. This discrepancy in MJA titer may be explained by the choice of loci for integration of *lovA*. One challenge of integration is tuning the correct copy number and the variable expression level from different genomic loci, which can be important for optimal pathway production and cell health[149,142]. Studies quantitatively comparing expression levels and integration effects at loci throughout the yeast genome have been performed, showing that expression level can vary by as much as 8.7-fold between loci and that the use of multiple integration sites dispersed through the genome increases strain stability[138,139,150,151]. It is possible that transcription and gene expression is higher at the *pyc2* locus than at either the *nte1* or X-2 loci, which could explain why higher conversion of DMLA to MJA was seen in YCB19+pLovB+pCB21. Though the knockout of *pyc2* rescued total polyketide production through increase of the acetyl-CoA precursor pool, the highest percent conversion of DMLA to MJA in the strains episomally maintaining only the Module 1 plasmids was 65%. Growth curves comparing YRC01, YCB14, YCB15, and YCB19 demonstrated a pattern similar to that seen in the JHY686 strains in that strains with knockouts of *pyc2* and *nte1* grew to a higher final OD_{600} (Fig. 19). In addition, YCB15 grew to a significantly higher final OD_{600} than YRC01, YCB14, or YCB19 but also produced the least polyketide when transformed with the Module 1 plasmids. This supports the hypothesis that improving

strain growth also improved production to a point, but further improvements to growth led to lower polyketide production due to less available carbon.

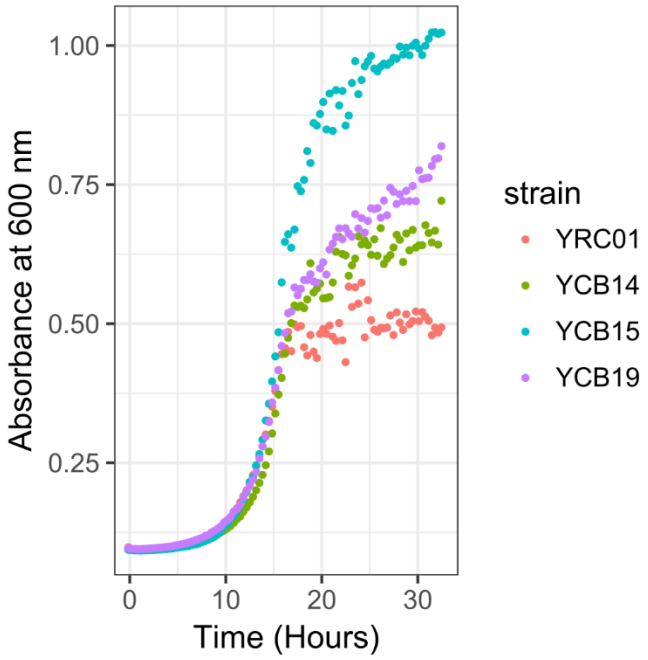

Figure 19: Growth curves of YRC01, YCB14, YCB15, and YCB19 in YPD. Each point is the mean average of three replicates.

A third plasmid, pCB19 (*lovD9* and *lovA*), was transformed into YRC01, YCB11, YCB14, YCB15, and YCB19. YCB20 was transformed with the Module 1 plasmids and pCB60 (*lovA*). This increased the percent conversion of DMLA to MJA up to nearly 90% for all the strains except for YCB14, which had slightly over 75% conversion from DMLA to MJA (Fig. 18). In some strains, like YCB11 and YCB19, the addition of pCB19 did not lead to a decrease in total polyketide production. However, YRC01 and YCB15 both had a ~50% decrease in total polyketide production and YCB14 decreased by nearly 90%. The addition of pCB19 to these strains showed that any addition of *lovA* to YRC01, whether integrated or episomally expressed, led to decreased total polyketide production. The decrease in production from YCB14 and YCB15 could be due to the integration of *lovD9* in the *pyc2*

locus, but given that pCB19 had a copy of *lovA* and a copy of *lovD9*, this is likely not the cause. YCB19+plovB+pCB21+pCB19 led to the highest polyketide production with 90% conversion of DMLA to MJA at 72 hours post inoculation. A time course was run of this strain to confirm at what time MJA production was maximized and indicated that MJA titer peaked at 75 mg/L, 72 hours post inoculation (Fig. 20)

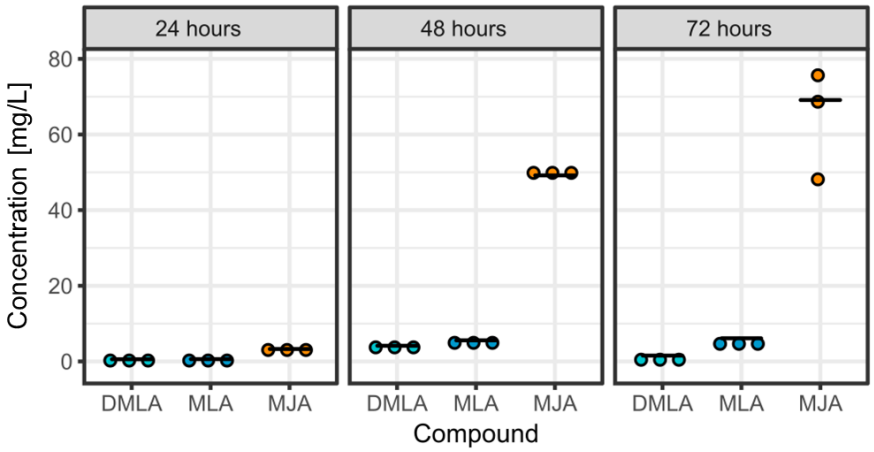

Figure 20: Breakdown of polyketide production in YCB19+plovB+pCB21+pCB19 at 24, 48, and 72 hours post inoculation into YPD. Each point is a different sample. Black horizontal lines represent median values.

Increasing the number of 2μ plasmids maintained in the SVA production strain from two to three could potentially reduce plasmid stability and hamper future fermentation, so other combinations of episomal expression of the necessary genes were pursued. There were concerns that cloning multiple gene cassettes with the same promoter and terminator into a single 2μ vector would increase the loss of plasmid sections *in vivo* due to homologous recombination. Recently, Harvey *et al* developed a set of ADH2p-like promoters that have diverse sequences and are also induced in the presence of ethanol and suppressed in the presence of glucose[120]. In addition, expression-enhancing terminators with divergent sequences have also been developed[128]. Utilization of these sequence-divergent promoters and terminators would allow cloning of multiple gene cassettes onto a single plasmid without the risk of homologous recombination. For an initial test of this method, plasmid pCB51 was constructed with the following gene cassettes: ADH2p-

lovD9-PRM9,; PCK1p-*lovG*-SPG5t, and ICL1p-*mlcG*-IDP1t. YCB11 was transformed with pLovB and pCB51, then cultured in YPD for 72 hours and DMB-SMMP was added at 24, 36, 48, and 72 hours to a final concentration of 1.22 mM. Total polyketide production from YCB11+pLovB+pCB51 was 50% lower than that from YCB11+pLovB+pCB21 and varied significantly between samples (Fig. 18, Fig. 21). The SVA titer from this strain was below 1 mg/L, a similar amount to what was seen in previous SVA-producing strains (Fig. 21).

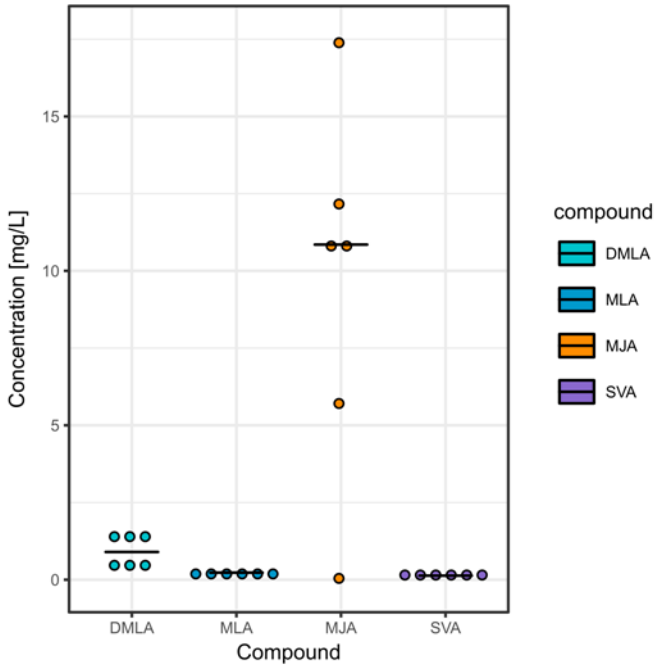

Figure 21: Polyketide production from YCB11+plovB+pCB51. Samples were cultured in YPD for 72 hours and DMB-SMMP was added at 24, 36, 48, and 72 hours to a final concentration of 1.22 mM. Each point is a different sample. Black horizontal lines represent median values.

3.6 Production of lovastatin in *S. cerevisiae*

SVA titers from YRC01+pLovB+pCB21+pCB19, YCB09+pLovB+pCB27, and YCB11+pLovB+pCB21 were all below 1 mg/L (Fig. 10, Fig. 12, Fig. 21). One possible explanation was that the exogenously added DMB-SMMP was not entering the cells, thus resulting in low SVA titers. To test this hypothesis, a LVA production strain was

constructed. For LVA production, the intracellular polyketide synthase LovF provides the acyl side chain and no exogenous addition of DMB-SMMP is required. For the first test, LovF and LovDwt were expressed from two separate 2μ plasmids, pXK32 and pCB44, in YRC01 and MJA was added to the strain at 24, 36, 48, and 60 hours post inoculation to a final concentration of 27 mg/L at 48 hours post inoculation and 54 mg/L at 72 hours post inoculation. Samples taken at 48 and 72 hours after inoculation showed production of 2.5 mg/L LVA (Fig. 22). Though low, this was higher statin titer than any of the SVA production strains reached and had potential to increase further if the MJA was made *in vivo* instead of being added exogenously. The LVA production pathway requires the expression of two polyketide synthases, LovB and LovF, so two new plasmids were built to limit the total number of plasmids in the production strain: pCB23 (*lovB* and *lovA*) and pCB26 (*lovF* and *lovDwt*). YRC01+pCB23+pCB21 led to less than 2 mg/L total polyketide compared to over 75 mg/L polyketide from YRC01+pLovB+pCB21 (Fig. 22). Addition of pCB26 decreased total polyketide production to less than 1 mg/L, and led to only 0.027 mg/L LVA. In the SVA-producing strain YRC01+pLovB+pCB21+pCB19 (Fig. 10), pLovB is 15.3 kilobases (kb), pCB21 is 8.8 kb, and pCB19 is 11.1 kb. In contrast, pCB23 is 17.4 kb and pCB26 is 17.2 kb. Notably, only one plasmid from YRC01+pLovB+pCB21+pCB19 was greater than 15 kb, and this plasmid only expressed LovB. In YRC01+pCB23+pCB21+pCB26, two plasmids were greater than 17 kb, and each of them was expressing a PKS and another smaller gene. The decline in production levels between these two similar expression systems and products was most likely due to the increased plasmid burden from the increase in plasmid size and genes expressed from a single plasmid.

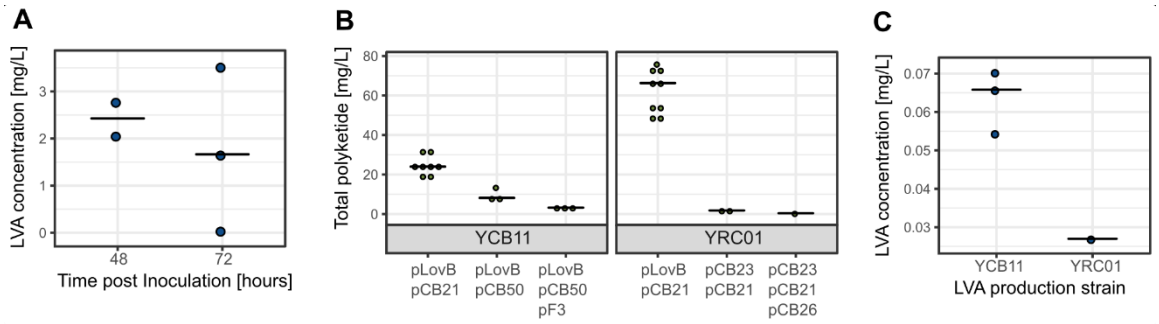

Figure 22: LVA production. (A) LVA produced from YRC01+pXK32+pCB44 after 48 or 72 hours growth in YPD with addition of MJA at24, 36, 48, and 60 hours post inoculation to a final concentration of 27 mg/L at 48 hours post inoculation and 54 mg/L at 72 hours post inoculation. (B) Total polyketide production from YCB11+pLovB+pCB21, YCB11+pLovB+pCB50, YCB11+pLovB+pCB50+pF3, YRC01+pLovB+pCB21, YRC01+pCB23+pCB21, and YRC01+pCB23+pCB21+pCB26. (C) LVA titers from YCB11+pLovB+pCB50+pF3 and YRC01+pCB23+pCB21+pCB26. Each point is a different sample. Black horizontal lines represent median values.

In order to allot two plasmids to only expression of the PKSs *lovB* and *lovF*, YCB11 with an integrated copy of *lovA* was also used as a production host. Again using the ADH2p-like promoters and synthetic terminators, pCB50 was constructed with the following cassettes: ADH2p-*lovDwt*-PRM9t, PCK1p-*lovG*-SPG5t, and ICL1p-*mlcG*-IDP1t[120,128]. Total polyketide production from YCB11+pLovB+pCB50 was 10 mg/L, a 60% decrease from the 25 mg/L produced by YCB11+pLovB+pCB21 (Fig. 22). This was a similar drop in total polyketide production as was seen when pCB21 was substituted with pCB51 (Fig. 21), indicating that the change in regulatory elements and increased plasmid size is the cause of this decrease. Potentially the ADH2p-like promoters led to less optimal expression levels of MlcG and LovG, or the copy number of pCB50 and pCB51 was significantly less than that of pCB21. A third plasmid, pF3 (*lovF*) was transformed into the strain to complete the LVA production pathway. YCB11+pLovB+pCB50+pF3 produced only 3.2 mg/L total polyketide, and only 0.06 mg/L LVA (Fig.22). The initial interest in building a *de novo* LVA production strain was for the insight it could afford to SVA synthesis. As it became apparent that significant engineering would be needed to increase the total

polyketide production and draw meaningful conclusions from the percent conversion from MJA to LVA, pursuit of this line of investigation was halted.

3.7 Characterization of plasmid stability and copy numbers

YCB19+pLovB+pCB21+pCB19 represents the best strain to date in the production of MJA, the immediate precursor to simvastatin. However, plasmid stability remained a concern. High-copy 2μ plasmids have been successfully used to express PKSs and heterologous pathways in yeast before but are still notorious for varying copy numbers and instability[63,25,152,153]. Understanding the stability of the three plasmids expressed in YCB19+pLovB+pCB21+pCB19 over the 72 hour production culture would give insight into the potential success of future fermentation of this strain give a new lens to examine polyketide production over time. The copy number of the plasmids would also offer an initial estimate of the number of integrated copies of each pathway gene to reach a similar production level.

To characterize the stability of this plasmid expression system, a time course was run using the YCB19+plovB+pCB21+pCB19 strain to measure copy numbers at 24, 48, and 72 hours after inoculation (Fig. 20, Table 3). The percentage of cells carrying each plasmid was similar at 24 and 72 hours after inoculation, indicating that the cells stably maintained the plasmids for this period of culturing. Notably, each plasmid was only carried by approximately 40-50% of the cells at each time point, resulting in a significantly lower copy number of each plasmid when averaged over all cells. The percentage of the cells carrying all three of the plasmids simultaneously was therefore 40-50% or less. This indicates that similar levels of production and conversion to MJA could potentially be achieved by integrating less than ten copies of *lovA*, rather than the twenty or more copies of pCB19 maintained by the cells.

Table 3: Plasmid copy number characterization. The percent of cells carrying plasmids was calculated by comparing the colony forming units (CFU) on the appropriate selective media plates to the CFU on complex media plates. The copy number averaged over all cells was calculated by qPCR using the standard curve method, and the copy number per cell carrying plasmid was calculated by taking the average copy number for all the cells and dividing it by the fraction of cells carrying the plasmid of interest.

	Time post inoculation	Percent of cells carrying plasmid	Copy number per cell carrying plasmid	Copy number averaged over all cells
pLovB	24 hours	39.3 ± 12.7	18.4 ± 5.7	6.8 ± 1.9
	48 hours	47.5 ± 12.2	22.8 ± 3.3	10.4 ± 1.0
	72 hours	44.1 ± 4.1	22.5 ± 5.3	9.7 ± 1.9
pCB21	24 hours	40.7 ± 13.5	75.6 ± 38.7	25.6 ± 8.4
	48 hours	43.3 ± 11.0	108.6 ± 20.7	45.1 ± 7.7
	72 hours	37.4 ± 5.0	81.5 ± 19.3	30.0 ± 5.5
pCB19	24 hours	51.7 ± 10.8	18.1 ± 3.1	9.5 ± 3.0
	48 hours	52.8 ± 12.7	25.2 ± 7.5	12.4 ± 0.7
	72 hours	44.7 ± 12.4	28.6 ± 6.8	12.1 ± 2.4

3.8 Development of an *in situ* lysis process to overcome substrate and enzyme localization differences

With high conversion to MJA achieved, the final step was to improve conversion of MJA to SVA. In the initial SVA production strains, YRC01+pLovB+pCB21+pCB19 and YCB09+pLovB+pCB27, DMB-SMMP was exogenously added throughout the 72-hour production run and less than 1 mg/L SVA was produced. In addition, the total polyketide production decreased by 20% compared to YCB07+pLovB+pCB27 (Fig. 12A,B). Several causes for this low conversion were considered. First, growth curves of *S. cerevisiae* in the presence of DMB-SMMP indicated that it inhibits yeast growth, which could contribute to the lower total polyketide production (Fig. 23). Second, LovD9 has optimal activity at pH

8.0 to 9.5 while the intracellular and extracellular pH of *S. cerevisiae* grown in YPD are significantly lower, so the lowered activity could be due to suboptimal pH conditions[154]. Third, as DMB-SMMP must be exogenously added, the relative localizations of MJA, DMB-SMMP, and LovD9 might be one of the reasons that SVA production was so low. Finally, bioactive secondary metabolites like fungal polyketides are often toxic at high concentrations. Simvastatin has been shown to inhibit yeast growth, although at concentrations closer to 200 mg/L of exogenously added SVA[83]. This raised concerns for the challenges in producing large quantities of SVA in yeast, especially when adding DMB-SMMP which is already toxic to the cells. These concerns were all addressed in turn.

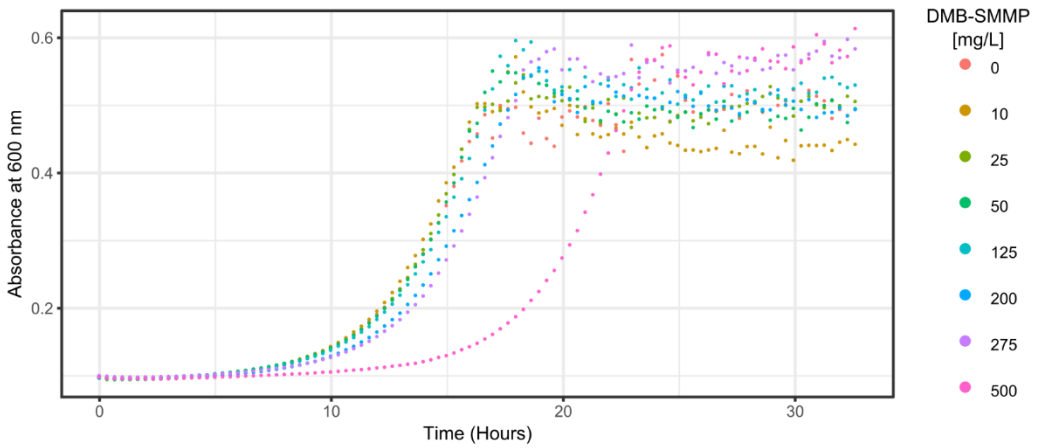

Figure 23: Growth inhibition of DMB-SMMP. Growth curves of YRC01 in YPD when exposed to increasing concentrations of DMB-SMMP. Each point is the mean average of three replicates.

Growth curves of *S. cerevisiae* in the presence of DMB-SMMP indicated that it inhibits yeast growth (Fig. 23). To further test the effect of DMB-SMMP on growth and polyketide production, YCB14+pLovB+pCB21 was cultured for 72 hours in YPD. DMB-SMMP was added to one set of cultures at 24, 36, 48, and 60 hours after inoculation to a final concentration of 6.1 mM. The cultures without added DMB-SMMP grew to a final OD_{600} that twice as much as the culture that was fed DMB-SMMP (Fig. 24). In addition, the total polyketide produced was 13-fold higher in the cultures that were not exposed to DMB-

SMMP. When total polyketide production was normalized by OD_{600}, the cultures that were exposed to 6.1 mM DMB-SMMP produced only 15% as much per cell as the cultures with no DMB-SMMP added (Fig. 24). To avoid DMB-SMMP toxicity in YCB19+pLovB+pCB21+pCB19, DMB-SMMP was added to a final concentration of 1.22 mM only at 72 hours after inoculation, when the concentration of MJA was the highest, and samples were taken 24 hours later. The resulting titer of SVA was only 0.5 mg/L (Fig. 25), similar to the cultures in which DMB-SMMP was added throughout (Fig. 10, 12, 21).

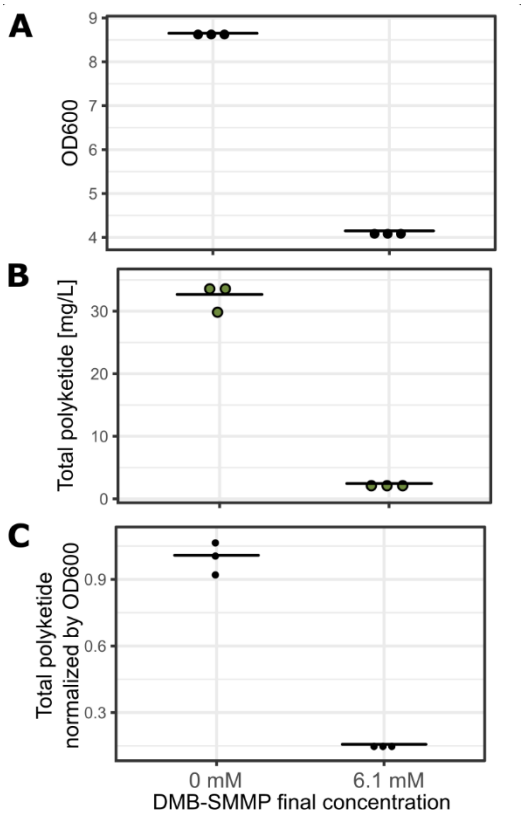

Figure 24: Effect of DMB-SMMP on YCB14+pLovB+pCB21. Polyketide production and growth from YCB14+pLovB+pCB21 after 72 hours culture in YPD. DMB-SMMP was added to one set of cultures at 24, 36, 48, and 60 hours after inoculation to a final concentration of 6.1 mM. (A) OD_{600}. (B) Total polyketide titer. (C) Polyketide titer normalized by the OD_{600}.

The next process modification was pH optimization. The intracellular pH of *S. cerevisiae* is 6-7 while the extracellular pH is 5-6[154]. However, LovD9 has optimal activity at pH 8.0

to 9.5 because of the requirement of a general base for catalysis[115]. Therefore, Tris-HCl buffer at pH 8.7 was added with DMB-SMMP to the cultures at 72 hours post-inoculation. Twenty-four hours later, polyketide concentrations were measured. Increasing the culture pH led to a five-fold increase in SVA titer to 5.9 mg/L (Fig. 25). However, the percent conversion from MJA to SVA was still less than 15%.

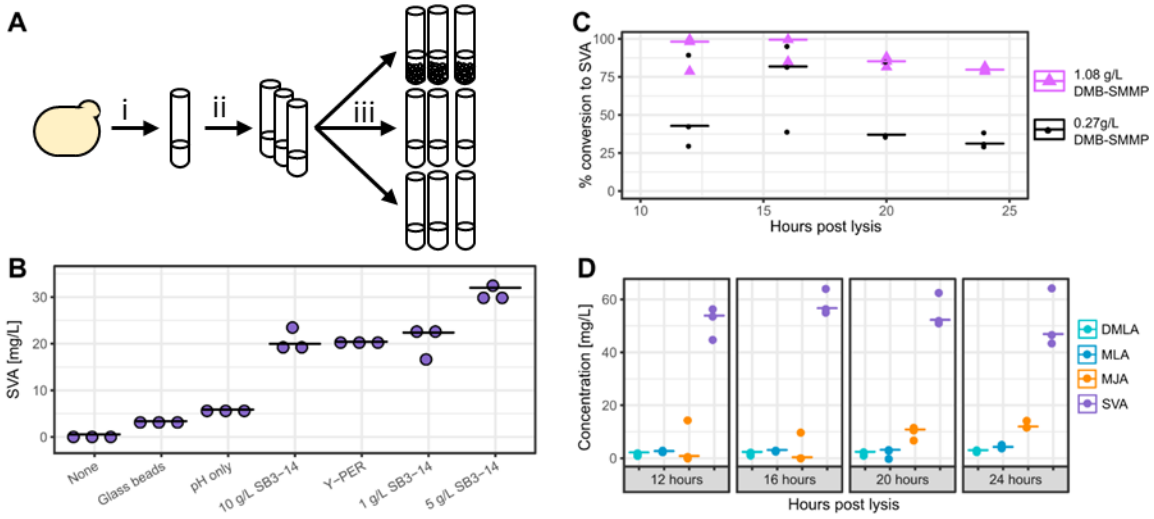

Figure 25: SVA production from YCB19+pLovB+pCB21+pCB19. (A) Schematic showing the production protocol. In step i, the yeast was picked into selective media. After 24 hours, triplicate cultures of YPD were inoculated from the seed in step ii. After 72 hours of culturing, the pH was adjusted, DMB-SMMP was added, and then the appropriate lysis method was used, as shown in step iii. Cultures were shaken for up to 24 hours more while samples were taken. (B) SVA concentration 24 hours after lysis and DMB-SMMP feeding using different lysis methods. The pH of all the cultures was raised to 8.7 except for the one labeled none. Each point is a different sample. The horizontal black line represents the median. (C) Time course of percent conversion of MJA to SVA after addition of 2 g/L of SB3-14 and varied concentrations of DMB-SMMP. Results in triplicate. The horizontal lines represent the median values. (D) Breakdown of polyketide production of the highest SVA producing conditions from the time course in (C).

To further increase conversion from MJA to SVA, localization of the *de novo* MJA was characterized. Separate analysis of the cell pellet versus the media indicated that 98% of the produced MJA had been secreted to the extracellular space (Fig. 26). As DMB-SMMP was added exogenously, it is reasonable to assume that a significant portion of this substrate would also be in the media. With both substrates extracellular and the LovD9 expressed intracellularly, conversion to SVA was therefore expectedly low. Once possible

solution is to remove the cell wall barrier at the time of MJA to SVA conversion. Cell-free reactions have been used to improve production of natural products, but these systems have required spinning down cultures and resuspending in a lysis buffer or otherwise concentrating the extract, which introduces significant process steps and scale-up barriers[155,156]. In addition, the produced extracellular MJA would not be spun down with the cell pellet and would be lost using traditional cell-free methods.

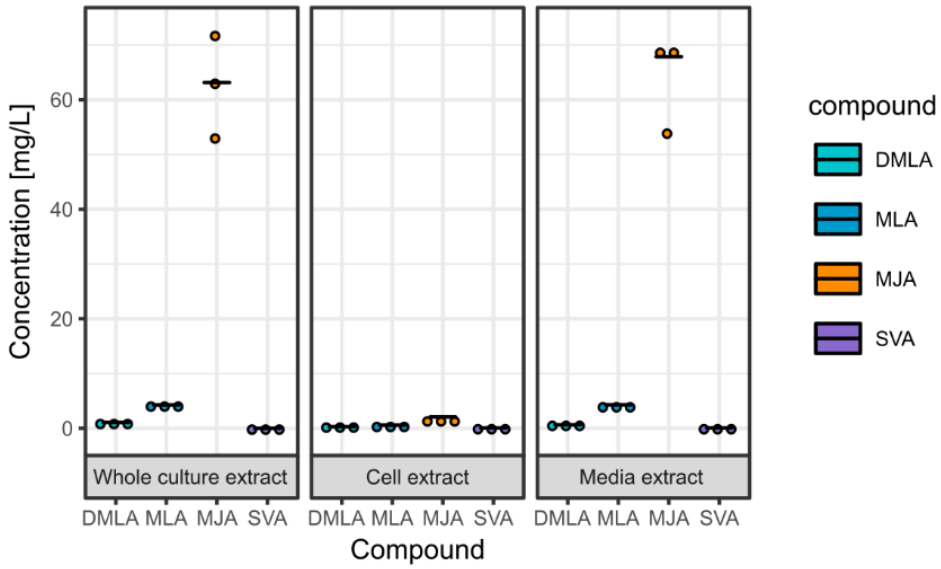

Figure 26: Polyketide localization and breakdown of YCB19+pLovB+pCB21+pCB19. Overall polyketide concentrations are compared to that from the cell pellet and the media fractions. Cultures were grown for 72 hours in YPD before addition of DMB-SMMP to a final concentration of 1.22 mM. Samples were taken 24 hours after addition of DMB-SMMP. Each point represents a different sample. The black horizontal lines represent the median.

For this reason, the cells were lysed *in situ* at 72 hours post inoculation when the MJA titer was highest (Fig. 25, 26). The cells were grown in minimal media for 24 hours before being inoculated into YPD at OD_{600} = 0.1. After 72 hours of growth, the cells were lysed, buffer was added to increase the pH to 8.7, and DMB-SMMP was added to a final concentration of 1.22 mM. Several lysis methods were tested, all performed with the increased media pH and addition of 1.22 mM DMB-SMMP. Initially, the cells were lysed *in situ* using glass beads, however, decreased conversion relative to the whole cells was observed (Fig. 25B).

Lysis efficiency was then tested using Y-PER and the cost-effective detergent 3-(N,N-Dimethylmyristylammonio)propanesulfonate (SB3-14)[157,158]. Y-PER and SB3-14 were both found to lead to improved SVA production, with 5 g/L SB3-14 giving the highest titer of SVA at 32 mg/L (Fig. 25). The cultures tested were also restreaked onto YPD to compare lysing effectiveness (Fig. 27). The highest SVA titer, 32 mg/L, resulted from lysis with 5 g/L SB3-14 which fully lysed the cells and led to no growth of the restreaks or colonies. The cells were also fully lysed from 10 g/L SB3-14, but this led to only 20 mg/L SVA. To optimize production, a further lysis efficiency test of 1, 2, 3, 4, and 5 g/L SB3-14 was run which indicated that 2 g/L SB3-14 was the minimum concentration required to lyse the cells completely (Fig. 27).

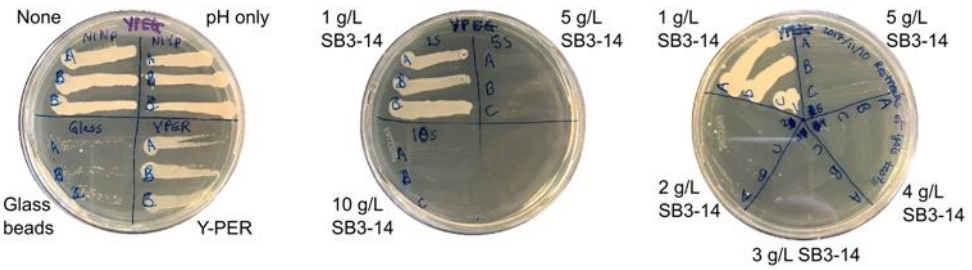

Figure 27: Lysing effectiveness. YCB19+pLovB+pCB21+pCB19 was grown for 72 hours in YPD before attempted lysis. Twenty-four hours later, 10 µL of each culture was streaked out on YPEG plates. All conditions were tested in triplicate. The pH of all the cultures was adjusted except for the one labeled none. The lysis method is listed next to the applicable restreaks. Conditions tested: No change (none), pH adjustement only, glass beads lysis, addition of Y-PER, or lysis with 1, 2, 3, 4, 5, or 10 g/L SB3-14.

Next, conversion of MJA to SVA was monitored over time. For YCB19+pLovB+pCB21+pCB19 that was not lysed, SVA titer peaked between 11 and 24 hours post lysis at 1.6 mg/L (Fig. 28). Two time courses were run for YCB19+pLovB+pCB21+pCB19 cultures that were lysed with 2 g/L SB3-14. The first time course indicated that SVA titer peaked between 12 and 24 hours post lysis (Figure 29). A second targeted time course confirmed that the highest titer of SVA occurred 16 hours post lysis (Fig. 25C,D). At longer time, the titer of SVA decreased, while that of MJA increased. This is expected as LovD9 can catalyze the reverse reaction which is hydrolysis of SVA to

MJA[110,112]. Adding an increased amount of DMB-SMMP to a final concentration of 4.88 mM led to production of 55 mg/L SVA and over 95% conversion from MJA (Fig. 25C,D). This represents the highest level of SVA produced in yeast using fermentation approaches and a 110-fold increase from the initial SVA production reached. These results indicate that *in situ* lysis is a viable new technique that could be applied to the production of other semisynthetic compounds using a yeast host.

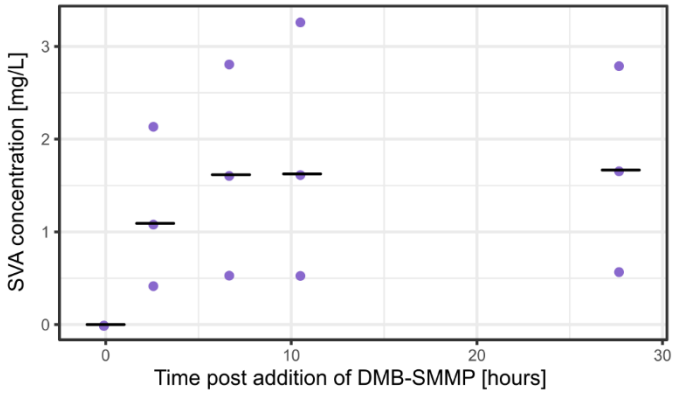

Figure 28: Time course of SVA production from YCB19+pLovB+pCB21+pCB19 without lysis. Cultures were grown in YPD for 72 hours at which point DMB-SMMP was added to a final concentration of 1.22 mM and the pH was increased to 8.8. Each point represents a different sample. The black horizontal lines represent the median.

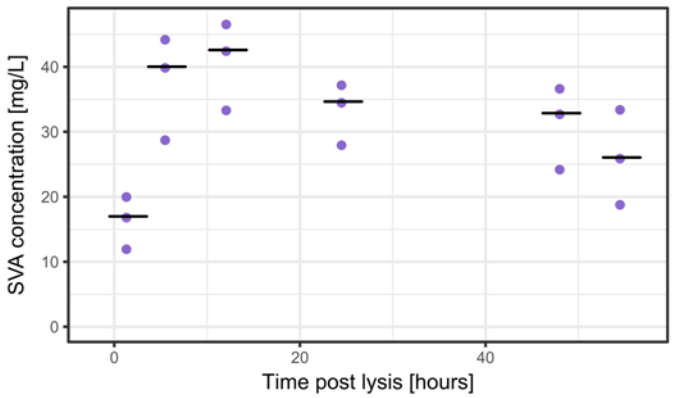

Figure 29: Time course of SVA production from YCB19+pLovB+pCB21+pCB19 after lysis with 2 g/L SB3-14. Cultures were grown in YPD for 72 hours at which point DMB-SMMP was added to a final concentration of 1.22 mM, SB3-14 was added to a final concentration of 2 g/L, and the pH was increased to 8.7. Each point represents a different sample. The black horizontal lines represent the median.

3.9 Surface expression and export of LovD9

Though the *in situ* lysis protocol for conversion of MJA to SVA resolved the localization differences between the substrates and enzyme and allowed nearly 100% conversion from MJA to SVA, after lysis the cells were no longer able to replicate and could not be used to continue production indefinitely. In addition, though low-cost, the detergent adds another step to the process that could increase the cost of the produced SVA. An alternative solution to the localization differences between the substrates and enzymes is to secrete LovD9 to the media.

Surface expression of enzymes in yeast has been used as a metabolic engineering strategy, especially for substrates that cannot or should not enter the cell[25,121]. Boder *et al* first developed the a-agglutinin mating protein for surface expression of fusion proteins in 1997, and it has since been used extensively for yeast cell surface expression[159,160,161,162]. The *S. cerevisiae* native protein Aga1p is coexpressed along with a fusion of the *S. cerevisiae* native protein Aga2p and the protein of interest. After translation, the Aga1p subunit is exported to the cell surface where it is covalently linked to cell wall glucans[163]. The Aga2p fusion protein is secreted into the extracellular space. The Aga2p subunit can be fused to the C- or N-terminus of the protein of interest. In either construct form, the native Aga2p signal peptide must be included at the N-terminus of the protein of interest to direct the protein through the native secretory pathway of yeast[161]. After secretion, the Aga2p and Aga1p subunits are anchored to one another by a pair of disulfide bonds[161]. To ensure proper folding and activity of the fused protein of interest, frequently a linker peptide chain is also inserted between the Aga2p subunit and the protein of interest[164,165,166]. The activity of the surface-displayed fusion protein can be strongly affected by which terminus Aga2p is fused to[161]. Crystal structures of LovD9 indicated that both termini are near to the surface of the enzyme, so there was no clear choice for which terminus to fuse to Aga2p

in order to least disrupt the folding and activity[115]. Therefore, two versions of the LovD9 fusion protein were constructed; one with the Aga2p subunit fused to the N-terminus and the other with the Aga2p subunit fused to the C-terminus. Both fusion proteins were constructed with the native Aga2p leader sequence at the N-terminus and a $(G_4S)_3$ linker between LovD9 and Aga2p[166] (Fig. 30A).

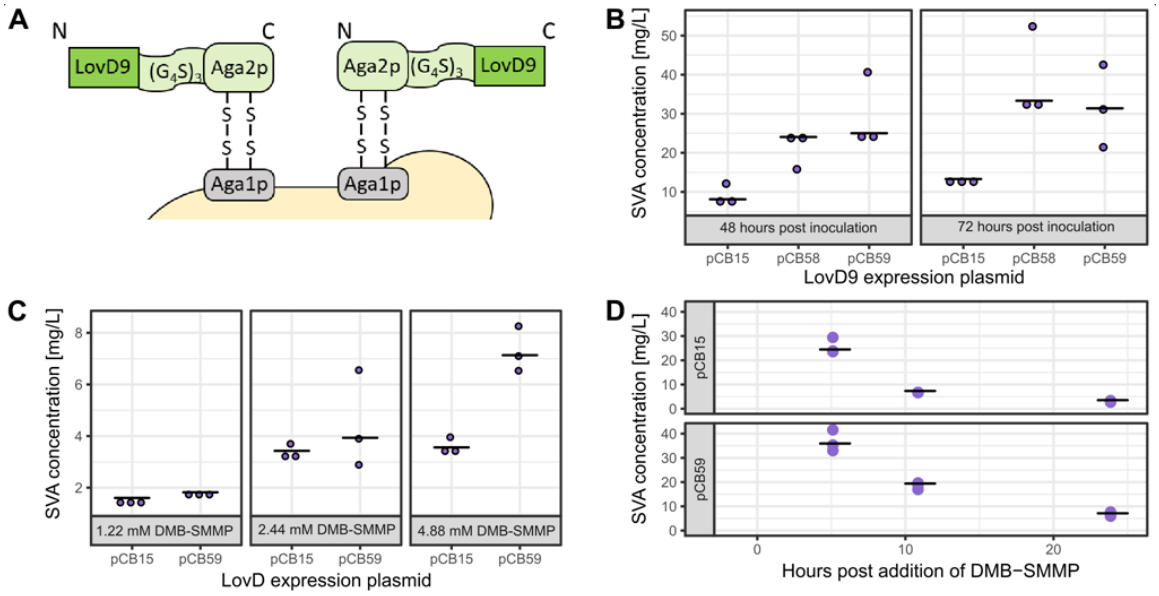

Figure 30: Surface display of LovD9. (A) Cartoon of the recombinant versions of LovD9 used for surface display. Aga1p is expressed separately and covalently bonds to the cell wall. Aga2p is anchored to Aga1p by two disulfide bonds and is expressed at either the C-terminus or N-terminus of LovD9. The Aga2p subunit and LovD9 are connected by a $(G_4S)_3$ linker. (B) Cultures of YRC01+pCB57 and either pCB15, pCB58, or pCB59 were grown for 48 or 72 hours in YPD before MJA and DMB-SMMP were added to a final concentration of 1.22 mM each and the pH increased to 8.7. Samples were taken 24 hours later. The time of the substrate addition is noted below each graph. (C) Cultures of YRC01+pCB57 and either pCB15 or pCB59 were grown for 72 hours before the pH was increased to 8.7, MJA was added to a final concentration of 0.3 mM, and DMB-SMMP was added to a final concentration of 1.22, 2.44, or 4.88 mM. Samples were taken 24 hours later. (D) Time course of SVA production from YRC01+pCB57 and either pCB15 or pCB59. Cultures were grown for 72 hours before the pH was increased to 8.7, MJA was added to a final concentration of 0.3 mM, and DMB-SMMP was added to a final concentration of 4.88 mM. Each point represents a different sample. The black horizontal lines represent the median.

YRC01 was transformed with a plasmid expressing Aga1p (pCB57) along with a plasmid either expressing cytosolic LovD9 (pCB15), the N-terminal Aga2p-fusion LovD9 (pCB58), or the C-terminal Aga2p-fusion LovD9 (pCB59). The resulting strains were grown for 48 or 72 hours before MJA and DMB-SMMP were added to a final concentration of 1.22 mM each and the pH increased to 8.7. Samples were taken 24 hours after the addition of the

substrates. Both versions of the fusion LovD9 led to significantly higher production of SVA than the cytosolic LovD9 (Fig. 30B). All the strains also had higher conversion from MJA to SVA when the substrates were added 72 hours rather than 48 hours after inoculation, indicating that there may be more LovD9 available at that time. Since both fusion proteins performed at a similar level, future experiments only compared the C-terminal Aga2p-fusion LovD9 (expressed from pCB59) with the cytosolic LovD9 (expressed from pCB15). In the next experiment, the ratio of DMB-SMMP to MJA was tuned. YRC01+pCB57+pCB15 and YRC01+pCB57+pCB59 were cultured for 72 hours. At this time, MJA was added to a final concentration of 100 mg/L (0.30 mM). In addition, the pH was increased to 8.7 and DMB-SMMP was added to each culture to a final concentration of 1.22, 2.44, or 4.88 mM. Samples were taken 24 hours later. As the amount of DMB-SMMP added increased, not only did the titer of SVA increase, but the difference in SVA production from pCB59 versus pCB15 increased as well (Fig. 30C). While production from the surface-expressed LovD9 seemed to increase proportionally with DMB-SMMP concentration, the increase in SVA production from cytosolic LovD9 plateaued at a lower concentration of exogenously added DMB-SMMP. Together this evidence indicated that the C-terminal Aga2p-fusion LovD9 was being secreted to the media. A production time course of strains YRC01+pCB57+pCB59 and YRC01+pCB57+pCB15 after addition of 4.88 mM DMB-SMMP indicated that the highest level of SVA production for both strains was reached between zero and six hours post addition of substrates (Fig. 30D). Surface expression of LovD9 led to 35 mg/L SVA while cytosolic expression of LovD9 led to 25 mg/L SVA at six hours post substrate addition, and higher titers could potentially be reached by further optimizing the sample time. Interestingly, the SVA titer from whole cells peaked significantly earlier than the SVA titer from the lysis method, in which the SVA titer peaked at 16 hours post lysis and DMB-SMMP addition.

Despite the early success of the surface display feeding studies, there were still significant challenges associated with incorporating LovD9 surface display into a strain that produces MJA *de novo*. While the cytosolic LovD can be directly replaced with the C-terminal Aga2p-fusion LovD9 in YCB19+pLovB+pCB21+pCB19, Aga1p must also be expressed, adding another gene that must be incorporated into the MJA production strain. In addition, it is possible that the fusion of Aga2p to LovD9 decreased its activity by interfering with the correct folding. For these reason, direct secretion of LovD9 to the media without displaying the secreted protein on the cell surface was also explored. In the *S. cerevisiae* secretory pathway, the signal peptide or leader sequence is a peptide signal at the N-terminus of a secretory protein composed of the pre- and pro-leader sequences[167]. After the leader sequence is translated, the signal peptide directs the target protein to the ER, where the pre-leader peptide is cleaved and the target protein is fully translated. Next, the target protein is transported to the Golgi apparatus where the pro-leader peptide is cleaved and finally the target protein is secreted out of the cell[167]. There are many different leader sequences that have been used for secretion of recombinant proteins in *S. cerevisiae,* including native, heterologous, and synthetic signal peptides[161,168]. The choice of leader sequence can have a strong effect on recombinant protein secretion, and the most advantageous leader sequence can vary depending on the target protein[168].

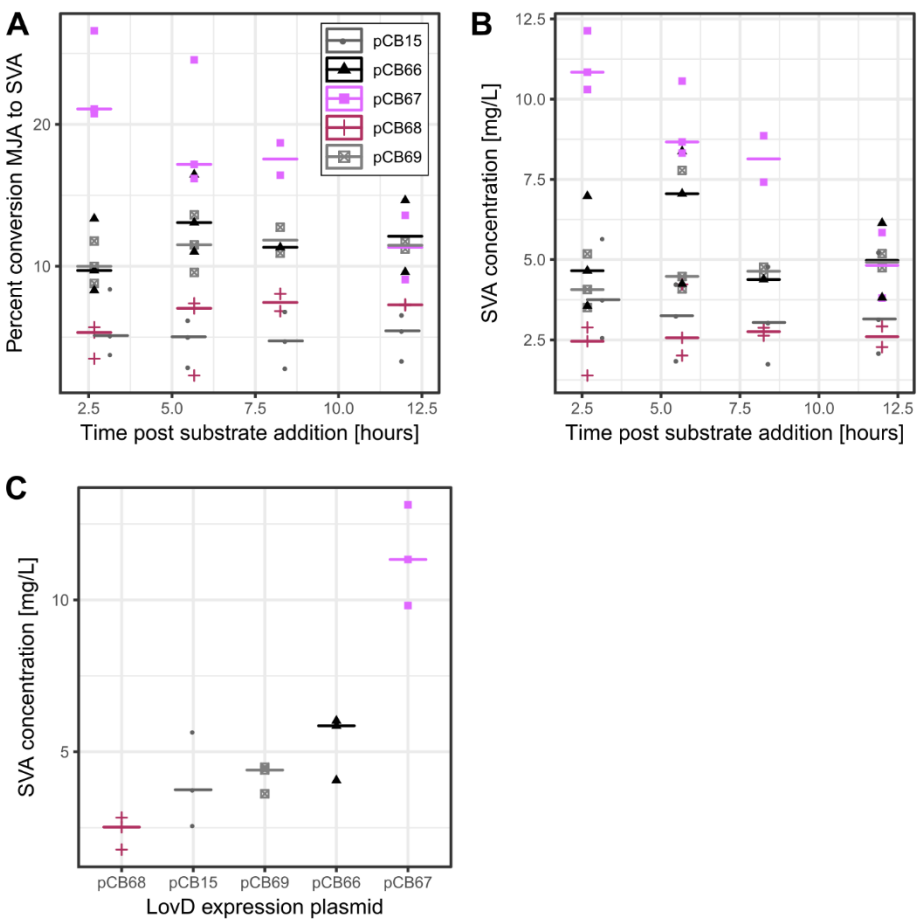

Figure 31: SVA production from secretion of LovD9. YRC01 was transformed with a plasmid expressing either cytosolic LovD9 (pCB15), App8-LovD9 (pCB66), AppS4-LovD9 (pCB67), K28-LovD9 (pCB68), or Syn-LovD9 (pCB69) and grown for 72 hours in YPD before the pH was increased to 8.7, DMB-SMMP was added to a final concentration of 4.88 mM, and MJA was added to a final concentration of 0.30 mM. (A) Time course of percent conversion of MJA to SVA. (B) Time course of SVA titer. (C) SVA titers from 2.5 hours after substrate addition. Each point represents a different sample. The horizontal lines represent the median.

Due to the variation in response to different leader sequences, LovD9 export was tested utilizing four leader sequences that were developed for protein export in *S. cerevisiae*. Two of the leader sequences, appS4 and app8 were selected from a library of mutant alpha mating factor 1 leader sequences because of the high levels of secretion achieved using these leader sequences to secrete structurally unrelated proteins[167]. A synthetic leader sequence Yap3-TA57 (syn) was also tested since it showed higher secretion from *S. cerevisiae* when compared to the native alpha factor leader[168]. The fourth leader sequence

tested was the signal peptide (K28) recently developed from the yeast K28 virus toxin to secrete Green Fluorescent Protein (GFP) in fours species of yeast[169,170,171]. The sequences of these signal peptides can be found in Table 9. Each of the leader sequence were cloned onto the N-terminus of *lovD9* and onto plasmids with the ADH2 promoter and terminator. YRC01 was transformed with a plasmid expressing either cytosolic LovD9 (pCB15), App8-LovD9 (pCB66), AppS4-LovD9 (pCB67), K28-LovD9 (pCB68), or Syn-LovD9 (pCB69) and grown for 72 hours in YPD before the pH was increased to 8.7, DMB-SMMP was added to a final concentration of 4.88 mM, and MJA was added to a final concentration of 0.30 mM. A time course of polyketide titers showed that the highest percent conversion from MJA to SVA occurred at 2.5 or 6 hours after substrate addition, depending on the version of LovD9 expressed (Fig. 31A,B). A comparison of SVA production at the 2.5 hour time point showed that cytosolic LovD9 led to 3.5 mg/L SVA and 5% conversion from MJA to SVA (Fig. 31C). All the leader sequences led to higher SVA titers except for K28-LovD9, which led to only 2.5 mg/L SVA. The two rationally engineered leader sequences performed the best, with AppS4-LovD9 leading to the nearly 17% conversion from MJA to SVA and 11 mg/L SVA. Even higher titers may be reached through further optimization of sample timing. This successful protein secretion is a good first step for developing a fermentable and continuously growing SVA production strain. From the feeding studies, surface display of LovD9 led to the highest titers of SVA. However, *de novo* production of MJA coupled with LovD9 secretion or surface display introduces more variables to SVA production which could affect which version of LovD9 is the optimal choice. This proof-of-concept has introduced another strategy that can be used to synthesize semi-synthetic analogs of natural products through fermentation.

4. Conclusion

In this study, the LVA biosynthetic pathway was refactored from *A. terreus* in *S. cerevisiae* and found that higher polyketide production was reached in the BJ5464-descended strains than in the faster growing BY4741-descended strains. Incorporation of the P450 *lovA*, whether integrated or episomally expressed, decreased total polyketide production. However, total polyketide production can be increased by knockouts of *pyc2* and *nte1*, and could compensate for the drop in titer upon *lovA* overexpression. High copy expression of *lovA* was necessary near complete conversion of DMLA to MJA. An analysis of plasmid copy number over culture time indicated the required copy number of each gene to reach similar levels of production from an integrated pathway. The initial SVA production from the top MJA-producing strain was 0.5 mg/L, which was increased 5-fold via pH optimization and 110-fold via optimized *in situ* lysis, reaching 55 mg/L SVA and nearly 100% conversion from MJA. This work marks the first time SVA has been produced using a single heterologous host, *S. cerevisiae*, and the first time this sort of *in situ* lysis has been used for production of a semisynthetic natural product. This provides a novel strategy that can be used by future researchers for production of semisynthetic polyketides in yeast.

Due to the majority of the *de novo* produced MJA localizing in the media, as well as the exogenously added DMB-SMMP, surface display and secretion of LovD9 were explored as alternative strategies to cell lysis. The initial studies using LovD9 surface display and exogenously added substrates indicated that LovD9 surface display doubled the titer of SVA reached by cytosolic LovD9. Similar studies testing four different secretion signal peptides showed that App8-LovD9 tripled the titer of SVA reached by cytosolic LovD9 in similar conditions. This preliminary work opens an exciting area of possibilities for one-pot production of SVA in *S. cerevisiae* without the limitations imposed by cell-free systems.

5. Appendix A: Supplementary tables

Table 4: The loci used for integrations in this work, as well as the homology used and gRNA target sequences if applicable

Locus	reference	Upstream homology	downstream homology	gRNA target
ura3Δ	Tang, M. C. *et al.*[121]	TTGCGAGGCATATTTATGGTGAAGGATAAG	TAGAAATCATTACGACCGAGATTCC	ACCATCAAAGAAGGTTAATG
Nte1	Cardenas, J. & Da Silva, N. A.[63]	TACAGGTTATAAGGATCTTAAAAATCTGGGAAACTTGA	ATACAGTAAAAAACAATGCAATACCTAAATCTATCTAC	GTTAGTGCAAGCTTTCAAAC
Pyc2	Cardenas, J. & Da Silva, N. A.[63]	TTGATTACATACATAAACAAGCCCCTCTTTTCTTCCAA	AGTGTTTGATTACTTTCTATTTGGCATTTAAATATACAT	GTTGTCTCTTCACCAGCAGA
HO	Baganz, F. et al[172]	TTGTTGAAGCATGATGAAGCGTTCTAAACGCACTATTCATCATTA	ATCTTGAGATGGCGTATTTCTACTCCAGCATTCTAGTTAAGAA	GGTCCAGTACTCGCAGGAAA
X-2	Mikkelsen, M. D. et al.[148]	ATCCTCTCATACCATATTAAGTAAATTGCCTCCATTT	AGCGGAGGAATAGTATGATAAATCT	CCTGCGAGTTTCTCTGCCCG
YPRCTy1-2	Kim, J. M. et al[173]	CTGTATACATAATATGATAGCCTTTACCAACAATGGAC	AATTATTCATCAAATACATCTCGATATCCATATTTTGGTTGCAAAGAAATAAATACCGG	N/A; leu2 marker-mediated integration was used for this locus
YDRWTy1-5	Kim, J. M. et al[173]	CATAGTAGACACAGAGTTGTATTTGCGCTTCTGAGCGATGCTTCCGAGATTGTTGAAGCA	ATAGGTTTGGGAACCTATGTTTTTTATTGTACCAAACGCATCCCTCTTCAATAATCTTCG	N/A; trp1 marker-mediated integration was used for this locus
Yia6	Cardenas, J. & Da Silva, N. A.[63]	ATAGATCCCAACCTTTATGTGAGCCAGCCCGAAGGGCCCT	GCATCGAACAGCCCACCAAATGCTTTGCGCACACATTATAT	GGGAAGTAGCCCAGGACAAT
X-4	Mikkelsen, M. D. et al.[148]	TGGTAAGCCGCCGTTTATAAACAGGGAAGATGTCCTT	ATGGTACAGGATACGCTAATTGCGC	GTCAAGGGAGGCACAGAGCA

pah1	Arendt, *et al.*[143]	ATTTCGAAGAAG TATGTAATTACC AAATAGCTCAGA AGAACACTTGGT ACGTAACAATCA	CACGATGGCCTACTCT TCACTTGTCCTTGCTC TTCCTTCTACCCTCAC CGTGCGCTGCTT	TAGTGGAGCA TCCAGACGGA

*Table 5:*Additional plasmids used for cloning or as PCR templates.

Plasmid name	Description	Reference
pRC152	2µ; AmpR; TRP1; ADH2p-*CPR*-ADH2t	Tang, M. C. et al[121]
pCB09	2µ; AmpR; URA3; ADH2p-*lovA*-FLAG-ADH2t; ADH2p-6xHis-*lovD9*-ADH2t	This study
pRS424	2µ; AmpR; TRP1; KanMX	Stanford Genome Technology Center gift
pRS425	2µ; AmpR; LEU2; HygR	Stanford Genome Technology Center gift
XW02	2µ; AmpR; LEU2; ADH2p-ADH2t	Xu, W. et al[118]
XW55	2µ; AmpR; URA3; ADH2p-ADH2t	Xu, W. et al[118]
pCRCT	2µ; AmpR; URA3; TEF1p-*iCas9*-ADH2t	Bao, Z. et al[123], Addgene plasmid # 60,621
pCB30	2µ; AmpR; KanMX; URA3_gRNA_cassette	This study
pCB32	2µ; AmpR; HygR; Nte1_gRNA_cassette	This study
pCB45	2µ; AmpR; HygR; X-2_gRNA_cassette; TEF1p-*iCas9*-ADH2t	This study
pCB52	2µ; AmpR; KanMX; Pyc2_gRNA_cassette; TEF1p-*iCas9*-ADH2t	This study
pXP320	2µ; AmpR; HIS3	Fang, F. et al[150]

Table 6: gBlocks ordered from IDT that were used in this work

gBlock	sequence
gRNA_del_URA3_1	agcttgtctgtaagcggaTCTCCACGGGACCCACAGTCGTAGATGCGTcacgtgTCTTTGAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGAAACTTGATGTTTTCTTTCGAGTATATACAAGGTGATTACATGTACGTTTGAAGTACAACTCTAGATTTTGTAGTGCCCTCTTGGGCTAGCGGTAAAGGTGCGCATTTTTTCACACCCTACAATGTTCTGTTCAAAAGATTTTGGTCAAACGCTGTAGAAGTGAAAGTTGGTGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCACCATCAAAGAAGGTTAATGGTTTTAGAGCTAGAAATAGCAAGTTAA

	AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGGTG CTTTTTTTGTTTTTTATGTCTgtttaaacAATGTATGAAACCCTGTATGGAGAGT GATTctggcgtttttccatagg
gRNA_Nte1_1	agcttgtctgtaagcggaTCTCCACGGGACCCACAGTCGTAGATGCGTcacgtgTCTTT GAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGAAACTTGAT GTTTTCTTTCGAGTATATACAAGGTGATTACATGTACGTTTGAAGTACAACT CTAGATTTTGTAGTGCCCTCTTGGGCTAGCGGTAAAGGTGCGCATTTTTTCA CACCCTACAATGTTCTGTTCAAAAGATTTTGGTCAAACGCTGTAGAAGTGAA AGTTGGTGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAAT GATCGTTAGTGCAAGCTTTCAAACGTTTTAGAGCTAGAAATAGCAAGTTAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGGTGC TTTTTTTGTTTTTTATGTCTgtttaaacAATGTATGAAACCCTGTATGGAGAGTG ATTctggcgtttttccatagg
gRNA_Pyc2_1	agcttgtctgtaagcggaTCTCCACGGGACCCACAGTCGTAGATGCGTcacgtgTCTTT GAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGAAACTTGAT GTTTTCTTTCGAGTATATACAAGGTGATTACATGTACGTTTGAAGTACAACT CTAGATTTTGTAGTGCCCTCTTGGGCTAGCGGTAAAGGTGCGCATTTTTTCA CACCCTACAATGTTCTGTTCAAAAGATTTTGGTCAAACGCTGTAGAAGTGAA AGTTGGTGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAAT GATCGTTGTCTCTTCACCAGCAGAGTTTTAGAGCTAGAAATAGCAAGTTAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGGTGC TTTTTTTGTTTTTTATGTCTgtttaaacAATGTATGAAACCCTGTATGGAGAGTG ATTctggcgtttttccatagg
gRNA_HO_1	agcttgtctgtaagcggaTCTCCACGGGACCCACAGTCGTAGATGCGTcacgtgTCTTT GAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGAAACTTGAT GTTTTCTTTCGAGTATATACAAGGTGATTACATGTACGTTTGAAGTACAACT CTAGATTTTGTAGTGCCCTCTTGGGCTAGCGGTAAAGGTGCGCATTTTTTCA CACCCTACAATGTTCTGTTCAAAAGATTTTGGTCAAACGCTGTAGAAGTGAA AGTTGGTGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAAT GATCGGTCCAGTACTCGCAGGAAAGTTTTAGAGCTAGAAATAGCAAGTTAA AATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGGTG CTTTTTTTGTTTTTTATGTCTgtttaaacAATGTATGAAACCCTGTATGGAGAGT GATTctggcgtttttccatagg
gBlock_tag1	GCGCGGaacaaaatgaaattgaaaactgttagatctgctgttttgtcttctttgtttgcttctcaagttttgggtc aaccaattgatgatactgaatctcaaactacttctgttaatttgatggctgatgatactgaatctgcttttgctactca aactaattctggtggtttggatgttgttggtttgatttctatggctaaaagagaagctggtgaaccaaaaATGA GATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCATTAGCT GCTCCAGctAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAG CTGTCATCGaTTACTcAGATTTAGAAGGGGATTTCGATGcTGCTGcTTTGCCA TTaTCCAACAGCACAAATAACGGGTTATcGTcTAcAAATACTACTATTGCCAG CATTGCTGCTAAAGAAGAAGGGGTACAGCTGGATAAAAGAGAGGCT
gBlock_tag2	GCCGCATGAGATTTCCTTCAATTTTTACTGCAGTTgTATTCGCAGCATCCTCC GCATTAGCTGCTCCAGcCAACACTACAgCAGAAGATGAAACGGCACAAATTC CGGCTGAAGCTGTCATCGGTTACTTAGgTTTAGAAGGGGATTcCGATGTTGC TGcTTTGCCATTaTCCgACAGCACAAATAACGGGTcATTGTcTAcAAATACTAC TATTGCCAGCATTGCTGCTAAAGAAGAAGGGGTACAGCTGGATAAAAGAGA AGCTatggagagcgtttcctcattattCaacatttttcaacAATCATGGTTAACTAcAAATCGT TAGTTCTAGCACTATTAAGTGTTTCAAATCTCAAATATGCACGGGGTTGATC CTGTTTGTAGAGACTTGACAAGAACCTTCGAAAAAGCAATCTATGCCCAATA CCAACAAGGTGAACAAAAACTCATCTCAGAAGAGGATCTGTAAcgatttatag

Nterminus_aga2_gBlock	gcatacaatcaactatcaactattaactatatcgtaatacATGCAGTTACTTCGCTGTTTTTCAA TATTTTCTGTTATTGCTTCAGTTTTAGCACAGGAACTGACAACTATATGCGA GCAAATCCCCTCACCAACTTTAGAATCGACGCCGTACTCTTTGTCAACGACT ACTATTTTGGCCAACGGGAAGGCAATGCAAGGAGTTTTTGAATATTACAAAT CAGTAACGTTTGTCAGTAATTGCGGTTCTCACCCCTCAACAACTAGCAAAGG CAGCCCCATAAACACACAGTATGTTTTTGGTGGTGGTGGATCTGGTGGtGGt GGttCtGGtGGtGGtGGttCtCATCATCACCATCACCACGGTAGTAACATTGATG CCGCTGTCGCCGCCGATCCTGTAGTCTTGATGGAAAC
Cterminus_aga2_gBlock	gcatacaatcaactatcaactattaactatatcgtaatacATGCAGTTACTTCGCTGTTTTTCAA TATTTTCTGTTATTGCTTCAGTTTTAGCACATCATCACCATCACCACGGTAGT AACATTGATGCCGCTGTCGCCGCCGATCCTGTAGTCTTGATGGAAACCAAGA ACCTTCGAAAAAGCAATCTATGCCCAATACCAACAAGGTGGTGGTGGTGGA TCTGGTGGtGGtGGttCtGGtGGtGGtGGttCtCAGGAACTGACAACTATATGCGA GCAAATCCCCTCACCAACTTTAGAATCGACGCCGTACTCTTTGTCAACGACT ACTATTTTGGCCAACGGGAAGGCAATGCAAGGAGTTTTTGAATATTACAAAT CAGTAACGTTTGTCAGTAATTGCGGTTCTCACCCCTCAACAACTAGCAAAGG CAGCCCCATAAACACACAGTATGTTTTTTAACACGTGttAAACATGCCTTCAcg atttatagttttc

Table 7: Genes names and sequence sources used in this work

Gene	source
lovB	Ma, S. M. et al.[25]
lovG	Xu, W. et al.[72]
lovC	Ma, S. M. et al.[25]
mlcG	Ma, S. M. et al.[25]
CPR	Barriuso, J. et al.[24]
lovA	Barriuso, J. et al.[24]
lovD9	Jimenez-Oses, G. et al.[115]
lovDG7	Gao et al.[114]
lovDwt	Gao et al.[114]
mlcE	Abe et al.[131]

Table 8: Sequence sources used for components of LovD9 surface display

Gene or sequence names	source
Aga1p	Boder et al.[159]
Aga2p leader	Boder et al.[159]
Aga2p mature peptide	Boder et al.[159]
(G4S)3	Wang, et al. [166]

Table 9: Sequences and sources of leader sequences used for secretion of LovD9

Signal Peptide name	reference	leader sequence
app8 alpha factor	Rakestraw et al.[167]	ATGAGATTTCCTTCAATTTTTACTGCAGTTTTAT TCGCAGCATCCTCCGCATTAGCTGCTCCAGctAA CACTACAACAGAAGATGAAACGGCACAAATTCC GGCTGAAGCTGTCATCGaTTACTcAGATTTAGAA GGGGATTTCGATGcTGCTGcTTTGCCATTaTCCA ACAGCACAAATAACGGGTTATcGTcTAcAAATAC TACTATTGCCAGCATTGCTGCTAAAGAAGAAGG GGTACAGCTGGATAAAAGAGAGGCT
appS4 alpha factor	Rakestraw et al.[167]	ATGAGATTTCCTTCAATTTTTACTGCAGTTgTAT TCGCAGCATCCTCCGCATTAGCTGCTCCAGcCAA CACTACAgCAGAAGATGAAACGGCACAAATTCC GGCTGAAGCTGTCATCGGTTACTTAGgTTTAGA AGGGGATTcCGATGTTGCTGcTTTGCCATTaTCC gACAGCACAAATAACGGGTcATTGTcTAcAAATA CTACTATTGCCAGCATTGCTGCTAAAGAAGAAG GGGTACAGCTGGATAAAAGAGAAGCT
Yeast K28 killer toxin	Eiden-Plach et al. [169]	atggagagcgtttcctcattattCaacatttttcaacAATCATGG TTAACTAcAAATCGTTAGTTCTAGCACTATTAAG TGTTTCAAATCTCAAATATGCACGGGGT
Synthetic signal peptide	Liu et al.[168]	aacaaaatgaaattgaaaactgttagatctgctgttttgtcttctttgttt gcttctcaagttttgggtcaaccaattgatgatactgaatctcaaactac ttctgttaatttgatggctgatgatactgaatctgcttttgctactcaaact aattctggtggtttggatgttgttggtttgatttctatggctaaaagagaa gctggtgaaccaaaa

Table 10: Primers used in this work

Primer name	sequence	Purpose	Template/binding site
CB0299	CCAATTGTTTAATCCGG GCTGGTT	qPCR for alg9 forward	Yeast genome
CB0300	CAGTGGACAGATAGCGT AGAGAGT	qPCR for alg9 reverse	Yeast genome
CB0176	cagttggcccttggctgata	qPCR for lovB	pLovB
CB0177	tgacagctgacgtgtctcg	qPCR for lovB	pLovB
CB0162	tctgtccctggtatcaagacgag	qPCR for lovG	pCB21
CB0301	cctcctgccactcggtatc	qPCR for lovG, pairs with CB162	pCB21
CB0302	GCATCTCCACACATTAA CTACGGTG	qPCR for lovDsimh for	pCB19

CB0303	CACCGTCCAAATCTTCTAAGGCTATAATACC	qPCR for lovDsimh rev	pCB19
CB0011	TTGCGAGGCATATTTATGGTGAAGGATAAGgcaaaacgtaggggcaaacaaac	Amplify cassette (ADH2p-CPR-ADH2t) for integration into URA3 locus of JHY686, use pRC152 as template	pRC152
CB0012	TAGAAATCATTACGACCGAGATTCCgggagcaaaaagtagaatattatcttttattcgtg	Amplify cassette (ADH2p-CPR-ADH2t) for integration into URA3 locus of JHY686 , pRC152 as template	pRC152
URA check R	GTGAAGTCATTGACACAGTCTG	used for cPCR check of locus to make sure it was integrated correctly	Yeast genome
CB0031	GAGGGCGGATTACTACCGT	genome check forward (upstream of URA3 deletion site, used for cPCR to check integration of CPR)	Yeast genome
CB0032	AAGCCTTGTCCCAAGGCAG	genome check reverse (downstream of URA3 deletion site, used for cPCR to check integration of CPR)	Yeast genome
CB0033	gccttcacgatttatagtttccattatc	ADH2t check forward (used for cPCR check that CPR integrated successfully)	ADH2 terminator
CB0036	GAGGGTGAACCCACCGA	Sequence CPR cassette, forward	*A. terreus CPR*
CB0037	GGATTCCACGACGGCAGT	Sequence CPR cassette, reverse	*A. terreus CPR*
CB0038	CCAGGCCAGGAAATCTTCTTCC	Sequence CPR cassette, reverse	*A. terreus CPR*
CB0039	AGCGACAAGGACTACTTCCAC	Sequence CPR cassette, forward	*A. terreus CPR*
CB0040	TACAGGTTATAAGGATCTTAAAAATCTGGGAAACTTGACAAAACGTAGGGGCAAACAAAC	prepare LovA cassette for integration into Nte1 locus, could	ADH2 promoter

		also be used for any other ADH2p - ADH2t cassette	
CB0041	ATACAGTAAAAAACAATGCAATACCTAAATCTATCTACACAGTGAATTCGAGCTCGGTAC	prepare LovA cassette for integration into Nte1 locus, could also be used for any other ADH2p - ADH2t cassette	ADH2 terminator
CB0044	CAAACATACCCAATATGGCAGAGATTC	NteI forward primer for cPCR check of integration into NteI locus	Yeast genome
CB0045	TAAAAGCACGCGACCGAAAAG	NteI locus reverse primer for cPCR check of integration into Nte1 locus	Yeast genome
CB0046	CAGCTAAAGGTAAATGGTACGATGC	forward primer binds to lovA, used for cPCR check of integration of lovA into Nte1 locus or Pyc2 locus	*A. terreus lovA*
CB0055	ACAGGTTATAAGGATCTTAAAAATCTGGG	Sequence lovA integration at Nte1	*A. terreus lovA*
CB0056	GCCATAGTCTTAATTGTCTAGCTTGG	Sequence lovA integration at Nte1	*A. terreus lovA*
CB0057	CCAGTGTTTGGATCTAGTAACTAGAA	Sequence lovA integration at Nte1	*A. terreus lovA*
CB0058	GGAGCTTGAAAGGTAATGATTGAAATATTAGAG	Sequence lovA integration at Nte1	*A. terreus lovA*
CB0042	TTGATTACATACATAAACAAGCCCCTCTTTTCTTCCAAcaaaacgtaggggcaaacaaac	prepare LovA cassette for integration into Pyc2 locus, could also be used for any other ADH2p - ADH2t cassette	ADH2 promoter
CB0043	AGTGTTTGATTACTTTCTATTTGGCATTTAAATATACATcagtgaattcgagctcggtac	prepare LovA cassette for integration into Pyc2 locus, could also be used for any other ADH2p - ADH2t cassette	ADH2 terminator
CB0046	CAGCTAAAGGTAAATGGTACGATGC	forward primer binds to lovA, used for cPCR check of integration of lovA	*A. terreus lovA*

		into Nte1 locus or Pyc2 locus	
CB0047	GTCTAAGAAGAAAACCTACAGTCGATCA	reverse primer of cPCR check of Pyc2 loci	Yeast genome
CB0048	TCGGTGCACCAATGGC	forward primer of cPCR check of Pyc2 loci, binds to Pyc2	Yeast genome
CB0060	GCAGGGAGTGTTAAGTAAGAAGTAC	sequencing integration at Pyc2 KO	Yeast genome
CB0203	ATCGTTGTCTCTTCACCAGCAGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG	making gRNA_Pyc2_1	pCB30, pCB32
CB0204	TAAAACTCTGCTGGTGAAGAGACAACGATCATTTATCTTTCACTGCGGAG	making gRNA_Pyc2_1	pCB30, pCB32
CB0007	GAAGAGAGTTGTCACCAAGGCC	amplify HO locus for sequencing	Yeast genome
CB0008	TCTTGACAAGAAGAGGTTTACTCTCTAGG	amplify HO locus for sequencing	Yeast genome
CB0072	TTGTTGAAGCATGATGAAGCGTTCTAAACGCACTATTCATCATTAATTTAGCGGCCGCAA	forward primer to integrate lovD in HO locus. Binds to ADH2p, has HO homology	ADH2 promoter
CB0073	ATCTTGAGATGGCGTATTTCTACTCCAGCATTCTAGTTAAGAACctcgaggggagcaaaa	reverse primer to integrate lovD in HO locus. Binds to ADH2t, has HO homology	ADH2 terminator
CB0075	TTGTGCAACGTATTCTCTTAACAAAGG	upstream reverse primer for sequencing lovD integ in HIS3 and HO loci	*lovD9*
CB0076	GTGCGCCATGCCTGT	downstream forward primer for sequencing lovD integ in HIS3 and HO loci	*lovD9*
CB0079	ACAGCTATTGCTACTCAAATGAGGT	upstream forward primer for sequencing lovD integ in HO locus	Yeast genome
CB0080	AACAAATCAGTGCCGGTAACG	downstream reverse primer for	Yeast genome

		sequencing lovD integ in HO locus	
CB0081	TGTCATATTGTCGAAGTGGTCACAA	forward primer for cPCR to check HO locus. Binds to parent.	Yeast genome
CB0239	ATGATCGGTCCAGTACTCGCAGGAAAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG	make gRNA_HO_1	pCB30, pCB32
CB0240	CTCTAAAACTTTCCTGCGAGTACTGGACCGATCATTTATCTTTCACTGCGGAG	make gRNA_HO_1	pCB30, pCB32
CB0117	ATCCTCTCATACCATATTAAGTAAATTGCCTCCATTTgcaaaacgtaggggcaaacaaac	for making ADH2p-lovA-ADH2t donor DNA for X-2 locus	ADH2 promoter
CB0118	AGCGGAGGAATAGTATGATAAATCTgggagcaaaaagtagaatattatcttttattcgtg	for making ADH2p-lovA-ADH2t donor DNA for X-2 locus	ADH2 promoter
CB0121	ATGATCCCTGCGAGTTTCTCTGCCCGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG	for making ori with gRNA_X-2_1 when paired with iCas9-pCB30 3	pCB30, pCB32
CB0122	CTCTAAAACCGGGCAGAGAAACTCGCAGGGATCATTTATCTTTCACTGCGGAG	for making antibiotic marker piece with gRNA_X-2_1 when paired with iCas9-pCB30 1	pCB30, pCB32
CB0123	GACGTAAAGTCAGGCAAGGC	for cPCR check	Yeast genome
CB0124	TCGGCGCTTAGTTTCGGA	for cPCR check	Yeast genome
CB0125	ACTCCTATAGTATAATATCGCCACTGACC	for sequencing	Yeast genome
SHR2-TEF1 F	CACCTTTCGAGAGGACGATGCCCGTGTCTAAATGATTCGACCAGCCTAAGAATGTTCAACCATAGCTTCAAAATGTTTCTACTCC	forward primer to amplify iCas9 from pCRCT for integration into YPRCTy1-2 with leu marker	pCRCT
YPRCTy1-2 ADH2t R	CCGGTATTTATTTCTTTGCAACCAAAATATGGATATCGAGATGTATTTGATGAATAATTatgagaaatatcgaggactcg	reverse primer to amplify iCas9 from pCRCT for integration into YPRCTy1-2 with leu marker	pCRCT

YPRCTy1-2 pXP F	CGAGAGAACTTCTAGTATATCTGTATACATAATATGATAGCCTTTACCAACAATGGACAGTGCCAAGCTTGCATGCC	forward primer to amplify leu marker from XW02 for integration into YPRCTy1-2 with iCas9	*leu2* cassette
SHR2 pXP R	GTTGAACATTCTTAGGCTGGTCGAATCATTTAGACACGGGCATCGTCCTCTCGAAAGGTGAGCTCGGTACCCGGG	reverse primer to amplify leu marker from XW02 for integration into YPRCTy1-2 with iCas9	*leu2* cassette
iCas9-back F	gcattgatttgagtcagctaggagg	forward primer to confirm iCas9 integration at YPRCTy1-2	*iCas9*
YPRC Ty1-2 check R	GAGGGTTGAACTCACGATCTTGC	reverse primer to confirm iCas9 integration at YPRCTy1-2	*iCas9*
iCas9 seq 1a	gattgtaggtttgtaccaactgg	sequence iCas9 and leu integration into JHY686	*iCas9*
iCas9 seq 2a	ggatggtactgaggaattattggtg	sequence iCas9 and leu integration into JHY686	*iCas9*
iCas9 seq 2b	ggcgtttgcgaatctctcc	sequence iCas9 and leu integration into JHY686	*iCas9*
iCas9 seq 2c	gcttcattaggtacctaccatg	sequence iCas9 and leu integration into JHY686	*iCas9*
iCas9 seq 2d	cttggtccacatacatgtctc	sequence iCas9 and leu integration into JHY686	*iCas9*
iCas9 seq 3a	ggctctgccaagcaaatatg	sequence iCas9 and leu integration into JHY686	*iCas9*
iCas9 seq3 F	gaaataggcaaagcaaccgc	sequence iCas9 and leu integration into JHY686	*iCas9*
YPRC Ty1-2 check R	GAGGGTTGAACTCACGATCTTGC	sequence iCas9 and leu integration into JHY686	*iCas9*
iCas9 seq1 R	cgttgtcaaaggtccgttgc	sequence iCas9 and leu integration into JHY686	*iCas9*
CB0027	cacgtgACGCATCTACGACTGTGGGTCCCGTGGAGAtccgcttacagacaagct	Prepare gRNA plasmid backbone	pCB30, pCB32

CB0028	gtttaaacAATGTATGAAAC CCTGTATGGAGAGTGAT Tctggcgtttttccataggc	Prepare gRNA plasmid backbone	pCB30, pCB32
CB0029	Atccgcttacagacaagct	short plasmids for initial amplification of gRNA plasmid backbone template	pCB30, pCB32
CB0030	ctggcgtttttccataggc	short plasmids for initial amplification of gRNA plasmid backbone template	pCB30, pCB32
CB0215	tgaatcgagtccctcgatatttctcat actagttctagagcgcgccaacaaa tatattgc	iCas9-pCB30 1 (use to make marker fragment of Cas9 gRNA plasmid)	pCB30, pCB32
CB0035	cctatggaaaaacgccagAAT	Amplify gRNA cassette from storage plasmid (reverse)	pCB30, pCB32
CB0065	gtttaaacAATGTATGAAAC CCTGTATGGAG	forward primer amplifies Hygromycin B resistance backbone from hygromycin storage plasmid and G418 storage plasmid (pCB30)	pCB30, pCB32
CB0066	cacgtgACGCATCTACGAC	reverse primer amplifies Hygromycin B resistance backbone from hygromycin storage plasmid and G418 storage plasmid (pCB30)	pCB30, pCB32
CB0212	CTGCATAGCTTCAAAAT GTTTCTACTCC	iCas9ec F, use to make iCas9 cassette	pCRCT
CB0213	tctagaactagtatgagaaatatcg agggactc	iCas9ec R, use to make iCas9 cassette	pCRCT
CB0214	AAGAGTAAAAAAGGAGT AGAAACATTTTGAAGCT ATGCAGacgaaagggcctcgtg	iCas9-pCB30 3 (use to make the 2μ fragment of Cas9 gRNA plasmid)	pCRCT
CB0215	tgaatcgagtccctcgatatttctcat actagttctagagcgcgccaacaaa tatattgc	iCas9-pCB30 1 (use to make marker fragment of Cas9 gRNA plasmid)	pCRCT

iCas9-pCB303	AAGAGTAAAAAAGGAGTAGAAACATTTTGAAGCTATGCAGacgaaagggcctcgtg	pairs with CB0122	pCRCT
GC_to_mlcG_cass_for	TAAATGATTCGACCAGCCTAAGAATGTTCAACttcgccctatagtgagtcgtattacaat	Forward primer, amplifying backbone to add homology for replacing lovC with mlcG (building pCB21)	pSL05
GC_to_mlcG_cass_rev	GCCTATTGATGATCTGGCGGAATGTCTGCCGTGCCATAatcctctagagtcgacctgcag	Reverse primer, amplifying backbone to add homology for replacing lovC with mlcG (building pCB21)	pSL05
mlcG_cass_to_GC_for	TATGGCACGGCAGACATTCCGCCAGATCATCAATAGGCagctggagctcggatcc	Forward primer, amplifying mlcG cassette to add homology for replacing lovC with mlcG	pCB20
mlcG_cass_to_GC_rev	GTTGAACATTCTTAGGCTGGTCGAATCATTTAGacgttgtaaaacgacggcca	Reverse primer, amplifying mlcG cassette to add homology for replacing lovC with mlcG	pCB20
mlcG_for_His_to CB07	gtaataccaTATGCATCATCACCATCACCACatgggcgttgccatgactgaagggagtttc	Forward primer, adding homology to mlcG to clone with AHD2p	pSMa125,
mlcG_rev_CB07	gataatgaaaactataaatcgtgaaggcatttaaacggaaaaccgaaccacaatcttctc	Reverse primer, adding homology to mlcG to clone with AHD2t	pSMa125,
CB07_for_to_mlcG	gagaagattgtggttcggttttccgttaaatgccttcacgatttatagtttcattatc	Forward primer, amplified backbone to build mlcG cassette	pLovB
CB07_rev_to_mlcG	gaaactcccttcagtcatggcaacgcccatGTGGTGATGGTGATGATGCATAtggtatta	Reverse primer, amplified backbone to build mlcG cassette	pLovB
CB09_cass_NaeI_for	gccagccggcctaaagggaacaaaagctggagc	adding NotI site to lovA cassette, forward. This created matching sticky end for	pCB09

		pCB15, which was digested with NotI and BamHI	
LovA_cass_BamH_rev	atatggatcccgttgtaaaacgacggcca	adding BamHI site to lovA cassette, reverse. This created matching sticky end for pCB15, which was digested with NotI and BamHI	pCB09
LovA_NheI_for	GGCAGCTAGCATGACAGTTGATGCTTTGACA	forward, added NheI to lovA. This created compatable sticky end to match with XW55 backbone, which was digested with NheI and PmlI	pCB09
LovA_pmlI_rev	CCACCACGTGTCAGATCTTATCGTCGTCATCC	reverse, added PmlI to lovA. This created compatible sticky end to match with XW55 backbone, which was digested with NheI and PmlI	pCB09
LovD9_For	aaCATATGCATCATCACCATCACCACGGTAGTAACATTGATGCCGCTG	forward, added NdeI to lovD9 . This created compatible end to match with XW02 backbone, which was digested with NdeI and PmeI	pCB05
LovD9_Rev	aaCACGTGTTAACCTTGTTGGTATTGGGCATAGATTG	reverse, added PmeI to lovD9. This created compatible sticky end to match with XW02 backbone, which was digested with NdeI and PmeI	pCB05
CB0005	tcaagcatcagtctcaggcacag	used to amplify mlcE gene	*P. citrinum* cDNA
CB0006	atgtcagaacctctacccct	used to amplify mlcE gene	*P. citrinum* cDNA
CB0009	gcaaaacgtaggggcaaacaaac	Amplify any cassette with ADH2p	pCB05

CB0010	gggagcaaaaagtagaatattatctttattcgtg	Amplify any cassette with ADH2t	pCB05
CB0063	ATAGATCCCAACCTTTATGTGAGCCAGCCCGAAGGGCCCTgcaaaacgtaggggcaaaca	prepare LovA cassette for integration into yia6 locus, could also be used for any other ADH2p - ADH2t cassette	2016-07-12 digest pCB18 with NotI to get donor DNA template
CB0064	ATATAATGTGTGCGCAAAGCATTTGGTGGGCTGTTCGATGcagtgaattcgagctcggt	prepare LovA cassette for integration into yia6 locus, could also be used for any other ADH2p - ADH2t cassette	2016-07-12 digest pCB18 with NotI to get donor DNA template
CB0065	gtttaaacAATGTATGAAACCCTGTATGGAG	forward primer amplifies Hygromycin B resistance backbone from hygromycin storage plasmid and G418 storage plasmid (pCB30)	Storage plasmid for Hyg backbon, pRS425
CB0066	cacgtgACGCATCTACGAC	reverse primer amplifies Hygromycin B resistance backbone from hygromycin storage plasmid and G418 storage plasmid (pCB30)	Storage plasmid for Hyg backbon, pRS425
CB0067	GCCATTCATTTGTTCGCACCT	reverse primer for cPCR to check YIA6 locus. Binds to child and to parent.	2016-07-12 yia6 locus after lovA cassette integ
CB0068	GTTGATCATGGCATCGTCCG	forward primer for cPCR to check YIA6 locus. Binds to parent.	S288C_YIL006W_YIA6_flanking
CB0069	GTAAGCCTTATCGAACACAATCCA	forward primer to sequence integration in YIA6 locus. Binds to parent and child	2016-07-12 yia6 locus after lovA cassette integ
CB0105	tttttccgtttgtttgcccctac	leu backbone rev	pCB15 - lovD simh leu amp

CB0106	CACGTGttAAACATGCCTTCAcg	leu backbone for	pCB15 - lovD simh leu amp
CB0107	tataaatcgTGAAGGCATGTTTaaCACGTGttaaccctgctggtactgcg	lovDwt rev to pCB15	pCB24
CB0108	ataaaagataatattctacttttgctcccctcg	leu backbone rev for going with lovF cassette	pCB15 - lovD simh leu amp
CB0109	atcccctgattctgtggataacc	use to get backbone for assembling lovDwt, mlcG, lovG plasmid	CB21-mlcG and lovG on plasmid
CB0110	acgcaggaaagaacatgtgag	use to get backbone for assembling lovDwt, mlcG, lovG plasmid	CB21-mlcG and lovG on plasmid
CB0111	aatacggttatccacagaatcagggatcgactcactatagggcgaattg	use to get insert for assembling lovDwt, mlcG, lovG plasmid	CB26 - assembled wt lovD + lovF leu carb
CB0112	ggccttttgctcacatgttctttcctgcgtAGATTACTCTAACGCCTCAGCC	use to get insert for assembling lovDwt, mlcG, lovG plasmid	CB26 - assembled wt lovD + lovF leu carb
CB0117	ATCCTCTCATACCATATTAAGTAAATTGCCTCCATTTgcaaaacgtaggggcaaacaaac	for making ADH2p-lovA-ADH2t donor DNA for X-2 locus	pCB40
CB0118	AGCGGAGGAATAGTATGATAAATCTgggagcaaaaagtagaatattatcttttattcgtg	for making ADH2p-lovA-ADH2t donor DNA for X-2 locus	pCB40
CB0121	ATGATCCCTGCGAGTTTCTCTGCCCGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG	for making ori with gRNA_X-2_1 when paired with iCas9-pCB30 3	pCB32
CB0122	CTCTAAAACCGGGCAGAGAAACTCGCAGGGATCATTTATCTTTCACTGCGGAG	for making antibiotic marker piece with gRNA_X-2_1 when paired with iCas9-pCB30 1	pCB32
CB0123	GACGTAAAGTCAGGCAAGGC	for cPCR check of X-2 locus	X-2 after ADH2p-lovA-ADH2t integ
CB0124	TCGGCGCTTAGTTTCGGA	for cPCR check of X-2 locus	X-2 Chr X 194944-195980
CB0125	ACTCCTATAGTATAATATCGCCACTGACC	for sequencing X-2 locus	X-2 after ADH2p-lovA-ADH2t integ
CB0126	TGGTAAGCCGCCGTTTATAAACAGGGAAGATGTCCTTgcaaaacgtaggggcaaacaaac	for making ADH2p-lovA-ADH2t donor DNA for X-4 locus	pCB40

CB0127	GCGCAATTAGCGTATCCTGTACCATgggagcaaaaagtagaatattatcttttattcgtg	for making ADH2p-lovA-ADH2t donor DNA for X-4 locus	pCB40
CB0128	ATGATCGACTAATGGCAGCCGTCGTTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG	making gRNA_X-4_1	pCB30
CB0129	CTCTAAAACAACGACGGCTGCCATTAGTCGATCATTTATCTTTCACTGCGGAG	making gRNA_X-4_1	pCB30
CB0130	ACCACCAAGTTTGGTTCAGC	cPCR for X-4 locus	X-4
CB0131	CTTTGTCAAGGGAGGCACAGA	cPCR for X-4 locus	X-4
CB0132	ATGATCGTCAAGGGAGGCACAGAGCAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG	making gRNA_X-4_2	pCB30
CB0133	CTCTAAAACTGCTCTGTGCCTCCCTTGACGATCATTTATCTTTCACTGCGGAG	making gRNA_X-4_2	pCB30
CB0134	TCAGCTTCTACAAGTGACTCGAG	for sequencing integration in X-4 locus	X-4
CB0143	GCCTGGTAAAGTTGTGTGCTAGTGTCTCCCGTCTTCTGTCttaaccctgctggtactgcg	prep lovD wt for 8 piece YHR, pairs with CB0150	pCB44
CB0144	ACCAAACTCACGCAACTAATTATTCCATAATAAAATAACAACatgcgctaccaagcatct	prep lovG for 8 piece YHR, pairs with CB0145	pCB21
CB0145	TAACCTTCTTGGTAATAGCGCGATGAAACAACGTCTTTGcacaaactcagtggtggt	prep lovG for 8 piece YHR, pairs with CB0144	pCB21
CB0146	GAAAACTCTTAGCATAACATAACAAAAAGTCAACGAAAACATCATCACCATCACCACatg	prep mlcG for 8 piece YHR, pairs with CB0147	pCB21
CB0147	AAAGTGGTAGATTGGGCTACGTAAATTCGAttaaacggaaaaccgaaccaca	prep mlcG for 8 piece YHR, pairs with CB0146	pCB21
CB0148	cacggaaatgttgaatactcatactcttc	prep trp marker for 8 piece YHR, pairs with CB0149	pCB21
CB0149	cctttagctgttctatatgctgcca	prep trp marker for 8 piece YHR, pairs with CB0148	pCB21

CB0150	taCAATCAACTATCAACTATTAACTATATCGTAATACCatgggatccatcattgatgctg	prep lovD wt for 8 piece YHR, pairs with CB0143	pCB44
CB0151	gtattacgatatagttaatagttgatagttgattgtatgc	ADH2p reverse	pCB44
CB0152	ATGCCTTCAcgatttatagtttcattatcaa	ADH2t forward	pCB44
CB0153	catacaatcaactatcaactattaactatatcgtaatacatgtcagaacctctaccccct	mlcE forward with homology to ADH2p	pCB46
CB0154	gaaaactataaatcgTGAAGGCATGTTTaaCACGTGttcaagcatcagtctcaggcac	mlcE reverse with homology to ADH2t	pCB46
CB0157	tgtttcttcttggtaaatagaatatcaagctacaa	prep lovD simh for 8 piece YHR, pairs with CB0158, bonds to ADH2p	pCB15
CB0158	AAAGTTGTGTGCTAGTGTCTCCCGTCTTCTGTCTTAACCTTGTTGGTATTGGGCATAGAT	prep lovD simh for 8 piece YHR, pairs with CB0157	pCB15
CB0159	ctgaaagaccagcccttcgag	bonds to lovF, pairs with CB0009	pCB25, pXK32
CB0160	accacactagactctgcagttca	bonds to lovF, pairs with CB0010	pCB25, pXK32
CB0161	gccggtaaattgtatacactcttgg	binds to mlcE, used for sequencing	Theoretical ADH2p-mlcE-ADH2t with leu marker
CB0162	tctgtccctggtatcaagacgag	binds to lovG, use for sequencing	theoretical lovG, mlcG, lovD wt plasmid
CB0163	tctcgcttcaacagcgagtg	binds to lovD wt, use for sequencing	theoretical lovG, mlcG, lovD wt plasmid
CB0164	ccatccgttgctccggg	binds to lovD wt, use for sequencing	theoretical lovG, mlcG, lovD wt plasmid
CB0165	TGAGTCATGTCAGCTCTTCTGG	binds to lovD simh, use for sequencing	theoretical lovG, mlcG, lovD simh plasmid
CB0166	GTTAACCCACACTTCTGGTTTGTC	binds to lovD simh, use for sequencing	theoretical lovG, mlcG, lovD simh plasmid
CB0167	cacctctgacttgagcgtc	binds to ecoli ori, used for changing plasmid marker	pCB49
CB0168	ttgacagcttatcatcggatcgatc	binds to 2μ ori and empty backbone,	pCB21

		used for changing plasmid marker	
CB0194	GACAGAAGACGGGAGACACTAGC	PRM9t FOR	pCB50
CB0195	TTTTCGTTGACTTTTTGTTATGTTATGCTAAGAG	ICL1p R	pCB50
CB0196	GCAAAGACGTTGTTTCATCGC	SPG5t F	pCB50
CB0197	GATGGTAATGATCCGAACTTGGG	IDP1t R	pCB50
CB0198	TCGAATTTACGTAGCCCAATCTACC	IDP1t F	pCB50
CB0199	gcgcggaacccctatttg	AmpR promoter F	pCB50
CB0200	GGTATTACGATATAGTTAATAGTTGATAGTTGATTG	ADH2 R	pCB50
CB0201	TCTAAAACCCACGCATAACCAAAGGAATGATCATTTATCTTTCACTGCGGAG	making gRNA_Nte1_2	pCB48
CB0202	ATGATCATTCCTTTGGTTATGCGTGGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG	making gRNA_Nte1_2	pCB48
CB0203	ATCGTTGTCTCTTCACCAGCAGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG	making gRNA_Pyc2_1	pCB45
CB0204	TAAAACTCTGCTGGTGAAGAGACAACGATCATTTATCTTTCACTGCGGAG	making gRNA_Pyc2_1	pCB45
CB0212	CTGCATAGCTTCAAAATGTTTCTACTCC	iCas9ec F, use to make iCas9 cassette	pCRCT
CB0213	tctagaactagtatgagaaatatcgagggactc	iCas9ec R, use to make iCas9 cassette	pCRCT
CB0214	AAGAGTAAAAAAGGAGTAGAAACATTTTGAAGCTATGCAGacgaaagggcctcgtg	iCas9-pCB30 3 (use to make the 2µ fragment of Cas9 gRNA plasmid)	pCB30/pCB32
CB0215	tgaatcgagtccctcgatatttctcatactagttctagagcgcgccaacaaatatattgc	iCas9-pCB30 1 (use to make marker fragment of Cas9 gRNA plasmid)	pCB30/pCB32
CB0235	ATGATCGATTGAGGGGGGCTTGATGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG	Use to make gRNA_PAH1_0 from Arendt et al paper	pCB32

CB0236	CTCTAAAACCATCAAGCCCCCCTCAATCGATCATTTATCTTTCACTGCGGAG	Use to make gRNA_PAH1_0 from Arendt et al paper	pCB32
CB0241	TGTTTCAGGTGATTGAGGGGGGCTTGAtttaaaCGCTCATCACGGGCGACACAAGCAAT	make donor DNA for PAH1 KO; P23	P23 and P24 Donor DNA
CB0242	ATTGCTTGTGTCGCCCGTGATGAGCGtttaaaTCAAGCCCCCCTCAATCACCTGAAACA	make donor DNA for PAH1 KO; P24	P23 and P24 Donor DNA
CB0244	CTTCAGTTGTATGTCTGTGTCATGC	Use to amplify PAH1 locus to send in for sequencing	after deletionS288C_YMR165C_PAH1_flanking
CB0245	CGCACGGTGAGGGTAGAA	Use to amplify PAH1 locus to send in for sequencing	after deletionS288C_YMR165C_PAH1_flanking
CB0256	tggatggcgatgataggcg	sequence mlcE	2017-05-23 X-4 after mlcE integ
CB0263	CCAAATAGCTCAGAAGAACACTTGGTACGTAACAATCACAAAACGTAGGGGCAAACAAAC	Add homology for PAH1 locus to any ADH2p cassette	pCB40
CB0264	AGGACAAGTGAAGAGTAGGCCATCGTGgggagcaaaaagtagaatattatcttttattcg	Add homology to PAH1 locus to any ADH2t cassette	pCB40
CB0265	TATTTACCTTGACATTTCGAAGAAGTATGTAATTACCAAATAGCTCAGAAGAACACTTGG	add more PAH1 homology to a PCR product that already has some homology	ADH2p-lovA-ADH2t cassette with little PAH1 homology
CB0266	GTTGAAAGCAGCGCACGGTGAGGGTAGAAGGAAGAGCAAGGACAAGTGAAGAGTAGGCCA	add more PAH1 homology to a PCR product that already has some homology	ADH2p-lovA-ADH2t cassette with little PAH1 homology
CB0269	CCTTTGACACATCTTCTTACCCG	amplify Nte1 locus for sequencing	KO nte1 with extra homology ADH2p-lovDsimh
CB0270	TAGTGGTAGCACTACAACTGGT	amplify Nte1 locus for sequencing	KO nte1 with extra homology ADH2p-lovDsimh
CB0271	TCTAAAACTCCGTCTGGATGCTCCACTAGATCATTTATCTTTCACTGCGGAGAAGTTTCG	add gRNA_PAH1_1 to hyg or G418 backbone	pCB32

CB0272	ATGATCTAGTGGAGCATCCAGACGGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG	add gRNA_PAH1_1 to hyg or G418 backbone	pCB32
CB0274	GACGTAATACCTGTTGCAGCA	upstream; cPCR piece from PAH1 locus for sequencing	after lovA integ S288C_YMR165C_PAH1_flanking
CB0275	CATAACACTACATCACTGGGAGAC	downstream; cPCR piece from PAH1 locus for sequencing	after lovA integ S288C_YMR165C_PAH1_flanking
CB0279	ATGCATCATCACCATCACCACG	Amplify lovDg7 or lovDsimh	pCB14, pCB15
CB0280	TTAACCTTGTTGGTATTGGGCATAGATTG	Amplify lovDg7 or lovDsimh	pCB14, pCB15
CB0286	ATGACATTATCTTTCGCTCATTTTACc	Amplify AGA1p from yeast genome	S288C_YNR044W_AGA1_genomic
CB0287	TTAACTGAAAATTACATTGCAAGCAAc	Amplify AGA1p from yeast genome	S288C_YNR044W_AGA1_genomic
CB0288	atcaactatcaactattaactatatcgtaatacATGACATTATCTTTCGCTCATTTTACC	add ADH2p homology to AGA1p	S288C_YNR044W_AGA1_genomic
CB0289	cttgataatgaaaactataaatcgtgaaggcatTTAACTGAAAATTACATTGCAAGCAAC	add ADH2t homology to AGA1p	S288C_YNR044W_AGA1_genomic
CB0290	CATCATCACCATCACCACGGTA	binds to begininning of lovDsimh, not including start codon	pCB15
CB0291	gcatacaatcaactatcaactattaactatatcgtaataca	Forward primer for N-terminus G block (aga2p and lovDsimh fusion)	g block for N-terminus aga2
CB0292	GTTTCCATCAAGACTACAGGATCGG	Reverse primer for N-terminus G block (aga2p and lovDsimh fusion)	g block for N-terminus aga2
CB0293	ACCTTGTTGGTATTGGGCATAG	binds to end of lovDsimh, not including stop codon	pCB15
CB0294	CAAGAACCTTCGAAAAAGCAATCTATGC	forward primer, binds to lovDsimh, use for c terminus aga2p plasmid construction	g block for C-terminus aga2
CB0295	gaaaactataaatcgTGAAGGCATGTTTaaCA	reverse primer, binds to lovDsimh,	g block for C-terminus aga2

		use for c terminus aga2p plasmid construction	
CB0296	TAACACGTGttAAACATG CCTTCAc	binds to gap between lovDsimh and ADH2t, use for c terminus aga2p plasmid construction	pCB15
CB0297	ATCAAACTTGTTGTCTG CGAGG	sequence aga1p	aga1p plasmid
CB0310	TTACAGATCCTCTTCTGA GATGAGTTTTTGTTCAC CTTGTTGGTATTGGGCA TA	primer for adding c-myc tag to C terminus of lovDsimh	pCB58
CB0311	CAAAAACTCATCTCAGA AGAGGATCTGTAAATGC CTTCAcgatttatagttttcattat c	primer for adding c-myc tag to C terminus of lovDsimh	pCB58
CB0312	ttttcaatttcattttgttgtattacga tatagttaatagttgatagttgattgt atgc	add homology for synthetic leader sequence	pCB58
CB0313	tgatttctatggctaaaagagaagc tggtgaaccaaaaCATCATCA CCATCACCACG	add homology for synthetic leader sequence	pCB58
CB0314	ctatcaactattaactatatcgtaat acaacaaaatgaaattgaaaactgt tagatctgc	add homology to synthetic leader sequence	2018-03-21 gBlock_tag1
CB0315	ATCAATGTTACTACCGT GGTGATGGTGATGATGtt ttggttcaccagcttctcttttagc	add homology to synthetic leader sequence	2018-03-21 gBlock_tag1
CB0316	TAAAAATTGAAGGAAAT CTCATgtattacgatatagttaat agttgatagttgattgtat	add homology for app8 leader sequence	pCB58
CB0317	CTGCTAAAGAAGAAGGG GTACAGCTGGATAAAAG AGAGGCTCATCATCACC ATCACCAC	add homology for app8 leader sequence	pCB58
CB0318	ctatcaactattaactatatcgtaat acATGAGATTTCCTTCAA TTTTTACTGCAGT	add homology to app8 leader sequence	2018-03-21 gBlock_tag1
CB0319	ATCAATGTTACTACCGT GGTGATGGTGATGATGA GCCTCTCTTTTATCCAGC TGTAC	add homology to app8 leader sequence	2018-03-21 gBlock_tag1
CB0320	GCTGCTAAAGAAGAAGG GGTACAGCTGGATAAAA GAGAAGCTCATCATCAC CATCACCAC	add homology for appS4 leader sequence	pCB58

CB0321	GCATCAATGTTACTACC GTGGTGATGGTGATGAT GAGCTTCTCTTTTATCCA GCTGTACC	add homology to appS4 leader sequence	2018-03-21 gBlock_tag2
CB0322	taatgaggaaacgctctccatgtatt acgatatagttaatagttgatagttg attgtatg	add homology for K28 leader sequence	pCB58
CB0323	GCACTATTAAGTGTTTC AAATCTCAAATATGCAC GGGGTCATCATCACCAT CACCACGGT	add homology for K28 leader sequence	pCB58
CB0324	caatcaactatcaactattaactata tcgtaatacatggagagcgtttcctc attattCa	add homology to K28 leader sequence	2018-03-21 gBlock_tag2
CB0325	CGGCATCAATGTTACTA CCGTGGTGATGGTGATG ATGACCCCGTGCATATT TGAGATTT	add homology to K28 leader sequence	2018-03-21 gBlock_tag2

6. References

1 Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *Journal of natural products* **75**, 311-335, doi:10.1021/np200906s (2012).

2 Harvey, A. L. Natural products in drug discovery. *Drug Discov Today* **13**, 894-901, doi:10.1016/j.drudis.2008.07.004 (2008).

3 Hertweck, C. The biosynthetic logic of polyketide diversity. *Angewandte Chemie* **48**, 4688-4716, doi:10.1002/anie.200806121 (2009).

4 Barrios-Gonzalez, J. & Miranda, R. U. Biotechnological production and applications of statins. *Appl Microbiol Biotechnol* **85**, 869-883, doi:10.1007/s00253-009-2239-6 (2010).

5 Baxter, A. *et al.* Squalestatin 1, a potent inhibitor of squalene synthase, which lowers serum cholesterol in vivo. *The Journal of biological chemistry* **267**, 11705-11708 (1992).

6 Udagawa, T. *et al.* Cytochalasin E, an epoxide containing Aspergillus-derived fungal metabolite, inhibits angiogenesis and tumor growth. *The Journal of pharmacology and experimental therapeutics* **294**, 421-427 (2000).

7 He, B., Wang, Y., Zheng, Y., Chen, W. & Zhu, Q. Synthesis and cytotoxic evaluation of acylated brefeldin a derivatives as potential anticancer agents. *Chemical biology & drug design* **82**, 307-316, doi:10.1111/cbdd.12154 (2013).

8 Lee, S. A., Kim, Y. J. & Lee, C. S. Brefeldin a induces apoptosis by activating the mitochondrial and death receptor pathways and inhibits focal adhesion kinase-mediated cell invasion. *Basic & clinical pharmacology & toxicology* **113**, 329-338, doi:10.1111/bcpt.12107 (2013).

9 Bentley, R. Mycophenolic Acid: a one hundred year odyssey from antibiotic to immunosuppressant. *Chemical reviews* **100**, 3801-3826 (2000).

10 Petersen, A. B., Ronnest, M. H., Larsen, T. O. & Clausen, M. H. The chemistry of griseofulvin. *Chemical reviews* **114**, 12088-12107, doi:10.1021/cr400368e (2014).

11 Roze, L. V., Hong, S. Y. & Linz, J. E. Aflatoxin biosynthesis: current frontiers. *Annual review of food science and technology* **4**, 293-311, doi:10.1146/annurev-food-083012-123702 (2013).

12 Dechert-Schmitt, A. M., Schmitt, D. C., Gao, X., Itoh, T. & Krische, M. J. Polyketide construction via hydrohydroxyalkylation and related alcohol C-H functionalizations: reinventing the chemistry of carbonyl addition. *Nat Prod Rep* **31**, 504-513, doi:10.1039/c3np70076c (2014).

13 Wong, F. T. & Khosla, C. Combinatorial biosynthesis of polyketides--a perspective. *Curr Opin Chem Biol* **16**, 117-123, doi:10.1016/j.cbpa.2012.01.018 (2012).

14 Tibrewal, N. & Tang, Y. Biocatalysts for natural product biosynthesis. *Annual review of chemical and biomolecular engineering* **5**, 347-366, doi:10.1146/annurev-chembioeng-060713-040008 (2014).

15 Brakhage, A. A. Regulation of fungal secondary metabolism. *Nature reviews. Microbiology* **11**, 21-32, doi:10.1038/nrmicro2916 (2013).

16 Schumann, J. & Hertweck, C. Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. *J Biotechnol* **124**, 690-703, doi:10.1016/j.jbiotec.2006.03.046 (2006).

17 Siddiqui, M. S., Thodey, K., Trenchard, I. & Smolke, C. D. Advancing secondary metabolite biosynthesis in yeast with synthetic biology tools. *FEMS Yeast Res* **12**, 144-170, doi:10.1111/j.1567-1364.2011.00774.x (2012).

18 Anyaogu, D. C. & Mortensen, U. H. Heterologous production of fungal secondary metabolites in Aspergilli. *Front Microbiol* **6**, 77, doi:10.3389/fmicb.2015.00077 (2015).

19 Mattanovich, D. *et al.* Recombinant protein production in yeasts. *Methods in molecular biology* **824**, 329-358, doi:10.1007/978-1-61779-433-9_17 (2012).

20 Nevoigt, E. Progress in metabolic engineering of Saccharomyces cerevisiae. *Microbiology and molecular biology reviews : MMBR* **72**, 379-412, doi:10.1128/MMBR.00025-07 (2008).

21 Kealey, J. T., Liu, L., Santi, D. V., Betlach, M. C. & Barr, P. J. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proceedings of the National Academy of Sciences* **95**, 505-509 (1998).

22 Pompon, D., Louerat, B., Bronine, A. & Urban, P. Yeast expression of animal and plant P450s in optimized redox environments. *Methods in enzymology* **272**, 51-64 (1996).

23 Tsunematsu, Y., Ishiuchi, K., Hotta, K. & Watanabe, K. Yeast-based genome mining, production and mechanistic studies of the biosynthesis of fungal polyketide and peptide natural products. *Nat Prod Rep* **30**, 1139-1149, doi:10.1039/c3np70037b (2013).

24 Barriuso, J. *et al.* Double oxidation of the cyclic nonaketide dihydromonacolin L to monacolin J by a single cytochrome P450 monooxygenase, LovA. *J Am Chem Soc* **133**, 8078-8081, doi:10.1021/ja201138v (2011).

25 Ma, S. M. *et al.* Complete reconstitution of a highly reducing iterative polyketide synthase. *Science (New York, N.Y.)* **326**, 589-592, doi:10.1126/science.1175602 (2009).

26 Zabala, A. O., Chooi, Y. H., Choi, M. S., Lin, H. C. & Tang, Y. Fungal polyketide synthase product chain-length control by partnering thiohydrolase. *ACS chemical biology* **9**, 1576-1586, doi:10.1021/cb500284t (2014).

27 Athlin, L., Domellof, L. & Norberg, B. Therapeutic concentrations of griseofulvin do not affect yeast cell phagocytosis by monocytes. *Dermatologica* **174**, 122-127 (1987).

28 Desmoucelles, C., Pinson, B., Saint-Marc, C. & Daignan-Fornier, B. Screening the yeast "disruptome" for mutants affecting resistance to the immunosuppressive drug, mycophenolic acid. *The Journal of biological chemistry* **277**, 27036-27044, doi:10.1074/jbc.M111433200 (2002).

29 Kuranda, K., Francois, J. & Palamarczyk, G. The isoprenoid pathway and transcriptional response to its inhibitors in the yeast Saccharomyces cerevisiae. *FEMS Yeast Res* **10**, 14-27, doi:10.1111/j.1567-1364.2009.00560.x (2010).

30 Leszczynska, A. *et al.* Investigating the effects of statins on cellular lipid metabolism using a yeast expression system. *Plos One* **4**, e8499, doi:10.1371/journal.pone.0008499 (2009).

31 Maciejak, A. *et al.* The effects of statins on the mevalonic acid pathway in recombinant yeast strains expressing human HMG-CoA reductase. *BMC Biotechnol* **13**, 68, doi:10.1186/1472-6750-13-68 (2013).

32 Shah, N. & Klausner, R. D. Brefeldin A reversibly inhibits secretion in Saccharomyces cerevisiae. *The Journal of biological chemistry* **268**, 5345-5348 (1993).

33 Fasullo, M., Sun, M. & Egner, P. Stimulation of sister chromatid exchanges and mutation by aflatoxin B1-DNA adducts in Saccharomyces cerevisiae requires MEC1 (ATR), RAD53, and DUN1. *Molecular carcinogenesis* **47**, 608-615, doi:10.1002/mc.20417 (2008).

34 Kelly, E. J., Erickson, K. E., Sengstag, C. & Eaton, D. L. Expression of human microsomal epoxide hydrolase in Saccharomyces cerevisiae reveals a functional role in aflatoxin B1 detoxification. *Toxicological sciences : an official journal of the Society of Toxicology* **65**, 35-42 (2002).

35 Yabe, K. *et al.* Production of M-/GM-group aflatoxins catalyzed by the OrdA enzyme in aflatoxin biosynthesis. *Fungal genetics and biology : FG & B* **49**, 744-754, doi:10.1016/j.fgb.2012.06.011 (2012).

36 Shen, B. Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Curr Opin Chem Biol* **7**, 285-295 (2003).

37 Keatinge-Clay, A. T. The structures of type I polyketide synthases. *Nat Prod Rep* **29**, 1050-1073, doi:10.1039/c2np20019h (2012).

38 Du, L. & Lou, L. PKS and NRPS release mechanisms. *Nat Prod Rep* **27**, 255-278, doi:10.1039/b912037h (2010).

39 Chooi, Y. H. & Tang, Y. Navigating the fungal polyketide chemical space: from genes to molecules. *The Journal of organic chemistry* **77**, 9933-9953, doi:10.1021/jo301592k (2012).

40 Yu, D., Xu, F., Zeng, J. & Zhan, J. Type III polyketide synthases in natural product biosynthesis. *IUBMB life* **64**, 285-295, doi:10.1002/iub.1005 (2012).

41 Yu, J. *et al.* Clustered pathway genes in aflatoxin biosynthesis. *Appl Environ Microbiol* **70**, 1253-1262 (2004).

42 Beck, J., Ripka, S., Siegner, A., Schiltz, E. & Schweizer, E. The multifunctional 6-methylsalicylic acid synthase gene of Penicillium patulum. Its gene structure relative to that of other polyketide synthases. *European journal of biochemistry* **192**, 487-498 (1990).

43 Cox, R. J. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Organic & biomolecular chemistry* **5**, 2010-2026, doi:10.1039/b704420h (2007).

44 Boettger, D. & Hertweck, C. Molecular diversity sculpted by fungal PKS-NRPS hybrids. *Chembiochem : a European journal of chemical biology* **14**, 28-42, doi:10.1002/cbic.201200624 (2013).

45 Smith, D. J. *et al.* Beta-lactam antibiotic biosynthetic genes have been conserved in clusters in prokaryotes and eukaryotes. *The EMBO journal* **9**, 741-747 (1990).

46 Keller, N. P., Turner, G. & Bennett, J. W. Fungal secondary metabolism - from biochemistry to genomics. *Nature reviews. Microbiology* **3**, 937-947, doi:10.1038/nrmicro1286 (2005).

47 Smith, D. J., Burnham, M. K., Edwards, J., Earl, A. J. & Turner, G. Cloning and heterologous expression of the penicillin biosynthetic gene cluster from penicillum chrysogenum. *Bio/technology* **8**, 39-41 (1990).

48 Bok, J. W. *et al.* Fungal artificial chromosomes for mining of the fungal secondary metabolome. *BMC genomics* **16**, 343, doi:10.1186/s12864-015-1561-x (2015).

49 van Leeuwen, J., Andrews, B., Boone, C. & Tan, G. Construction of Multifragment Plasmids by Homologous Recombination in Yeast. *Cold Spring Harbor protocols* **2015**, pdb top084111, doi:10.1101/pdb.top084111 (2015).

50 Larionov, V. *et al.* Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc Natl Acad Sci U S A* **93**, 491-496 (1996).

51 Shao, Z., Zhao, H. & Zhao, H. DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic acids research* **37**, e16, doi:10.1093/nar/gkn991 (2009).

52 Oldenburg, K. R., Vo, K. T., Michaelis, S. & Paddon, C. Recombination-mediated PCR-directed plasmid construction in vivo in yeast. *Nucleic acids research* **25**, 451-452 (1997).

53 Kouprina, N. & Larionov, V. TAR cloning: insights into gene function, long-range haplotypes and genome structure and evolution. *Nature reviews. Genetics* **7**, 805-812, doi:10.1038/nrg1943 (2006).

54 Feng, Z., Kim, J. H. & Brady, S. F. Fluostatins produced by the heterologous expression of a TAR reassembled environmental DNA derived type II PKS gene cluster. *J Am Chem Soc* **132**, 11902-11903, doi:10.1021/ja104550p (2010).

55 Yamanaka, K. *et al.* Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc Natl Acad Sci U S A* **111**, 1957-1962, doi:10.1073/pnas.1319584111 (2014).

56 Mao, X. M. *et al.* Efficient Biosynthesis of Fungal Polyketides Containing the Dioxabicyclo-octane Ring System. *J Am Chem Soc* **137**, 11904-11907, doi:10.1021/jacs.5b07816 (2015).

57 Winter, J. M. *et al.* Biochemical and Structural Basis for Controlling Chemical Modularity in Fungal Polyketide Biosynthesis. *J Am Chem Soc* **137**, 9885-9893, doi:10.1021/jacs.5b04520 (2015).

58 Winter, J. M. *et al.* Identification and characterization of the chaetoviridin and chaetomugilin gene cluster in Chaetomium globosum reveal dual functions of an iterative highly-reducing polyketide synthase. *Journal of the American Chemical Society* **134**, 17900-17903 (2012).

59 Lin, H. C. *et al.* The fumagillin biosynthetic gene cluster in Aspergillus fumigatus encodes a cryptic terpene cyclase involved in the formation of beta-trans-bergamotene. *J Am Chem Soc* **135**, 4616-4619, doi:10.1021/ja312503y (2013).

60 Ishiuchi, K. *et al.* Establishing a new methodology for genome mining and biosynthesis of polyketides and peptides through yeast molecular genetics. *Chembiochem : a European journal of chemical biology* **13**, 846-854, doi:10.1002/cbic.201100798 (2012).

61 Pahirulzaman, K. A., Williams, K. & Lazarus, C. M. A toolkit for heterologous expression of metabolic pathways in Aspergillus oryzae. *Methods in enzymology* **517**, 241-260, doi:10.1016/B978-0-12-404634-4.00012-7 (2012).

62 Kakule, T. B., Lin, Z. & Schmidt, E. W. Combinatorialization of fungal polyketide synthase-peptide synthetase hybrid proteins. *J Am Chem Soc* **136**, 17882-17890, doi:10.1021/ja511087p (2014).

63 Cardenas, J. & Da Silva, N. A. Metabolic engineering of Saccharomyces cerevisiae for the production of triacetic acid lactone. *Metab Eng* **25**, 194-203, doi:10.1016/j.ymben.2014.07.008 (2014).

64 Jones, E. W. Tackling the protease problem in Saccharomyces cerevisiae. *Methods in enzymology* **194**, 428-453 (1991).

65 Lee, K. K., Da Silva, N. A. & Kealey, J. T. Determination of the extent of phosphopantetheinylation of polyketide synthases expressed in Escherichia coli

and Saccharomyces cerevisiae. *Anal Biochem* **394**, 75-80, doi:10.1016/j.ab.2009.07.010 (2009).

66 Wattanachaisaereekul, S., Lantz, A. E., Nielsen, M. L., Andresson, O. S. & Nielsen, J. Optimization of heterologous production of the polyketide 6-MSA in Saccharomyces cerevisiae. *Biotechnol Bioeng* **97**, 893-900, doi:10.1002/bit.21286 (2007).

67 Li, Y., Chooi, Y. H., Sheng, Y., Valentine, J. S. & Tang, Y. Comparative characterization of fungal anthracenone and naphthacenedione biosynthetic pathways reveals an alpha-hydroxylation-dependent Claisen-like cyclization catalyzed by a dimanganese thioesterase. *J Am Chem Soc* **133**, 15773-15785, doi:10.1021/ja206906d (2011).

68 Zhou, H., Qiao, K., Gao, Z., Vederas, J. C. & Tang, Y. Insights into radicicol biosynthesis via heterologous synthesis of intermediates and analogs. *The Journal of biological chemistry* **285**, 41412-41421, doi:10.1074/jbc.M110.183574 (2010).

69 Xu, Y. *et al.* Diversity-oriented combinatorial biosynthesis of benzenediol lactone scaffolds by subunit shuffling of fungal polyketide synthases. *Proc Natl Acad Sci U S A* **111**, 12354-12359, doi:10.1073/pnas.1406999111 (2014).

70 Chooi, Y. H. *et al.* An in planta-expressed polyketide synthase produces (R)-mellein in the wheat pathogen Parastagonospora nodorum. *Appl Environ Microbiol* **81**, 177-186, doi:10.1128/AEM.02745-14 (2015).

71 Kennedy, J. *et al.* Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science (New York, N.Y.)* **284**, 1368-1372 (1999).

72 Xu, W. *et al.* LovG: the thioesterase required for dihydromonacolin L release and lovastatin nonaketide synthase turnover in lovastatin biosynthesis. *Angewandte Chemie* **52**, 6472-6475, doi:10.1002/anie.201302406 (2013).

73 Xie, X., Meehan, M. J., Xu, W., Dorrestein, P. C. & Tang, Y. Acyltransferase mediated polyketide release from a fungal megasynthase. *J Am Chem Soc* **131**, 8388-8389, doi:10.1021/ja903203g (2009).

74 Wang, M., Beissner, M. & Zhao, H. Aryl-aldehyde formation in fungal polyketides: discovery and characterization of a distinct biosynthetic mechanism. *Chem Biol* **21**, 257-263, doi:10.1016/j.chembiol.2013.12.005 (2014).

75 Krivoruchko, A. & Nielsen, J. Production of natural products through metabolic engineering of Saccharomyces cerevisiae. *Curr Opin Biotechnol* **35**, 7-15, doi:10.1016/j.copbio.2014.12.004 (2015).

76 Wattanachaisaereekul, S., Lantz, A. E., Nielsen, M. L. & Nielsen, J. Production of the polyketide 6-MSA in yeast engineered for increased malonyl-CoA supply. *Metab Eng* **10**, 246-254, doi:10.1016/j.ymben.2008.04.005 (2008).

77 Choi, J. W. & Da Silva, N. A. Improving polyketide and fatty acid synthesis by engineering of the yeast acetyl-CoA carboxylase. *J Biotechnol* **187**, 56-59, doi:10.1016/j.jbiotec.2014.07.430 (2014).

78 Lian, J., Si, T., Nair, N. U. & Zhao, H. Design and construction of acetyl-CoA overproducing Saccharomyces cerevisiae strains. *Metab Eng* **24**, 139-149, doi:10.1016/j.ymben.2014.05.010 (2014).

79 Chen, Y., Daviet, L., Schalk, M., Siewers, V. & Nielsen, J. Establishing a platform cell factory through engineering of yeast acetyl-CoA metabolism. *Metab Eng* **15**, 48-54, doi:10.1016/j.ymben.2012.11.002 (2013).

80 Martín, J. F., Casqueiro, J. & Liras, P. Secretion systems for secondary metabolites: how producer cells send out messages of intercellular communication. *Current opinion in microbiology* **8**, 282-293 (2005).

81 Rodriguez, S., Kirby, J., Denby, C. M. & Keasling, J. D. Production and quantification of sesquiterpenes in Saccharomyces cerevisiae, including extraction, detection and quantification of terpene products and key related metabolites. *Nat Protoc* **9**, 1980-1996, doi:10.1038/nprot.2014.132 (2014).

82 do Valle Matta, M. A., Jonniaux, J. L., Balzi, E., Goffeau, A. & van den Hazel, B. Novel target genes of the yeast regulator Pdr1p: a contribution of the TPO1 gene in resistance to quinidine and other drugs. *Gene* **272**, 111-119 (2001).

83 Ley, A., Coumou, H. C. & Frandsen, R. J. N. Heterologous expression of MlcE in Saccharomyces cerevisiae provides resistance to natural and semi-synthetic statins. *Metabolic Engineering Communications* **2**, 117-123, doi:https://doi.org/10.1016/j.meteno.2015.09.003 (2015).

84 Butler, M. S., Blaskovich, M. A. & Cooper, M. A. Antibiotics in the clinical pipeline in 2013. *The Journal of antibiotics* **66**, 571-591, doi:10.1038/ja.2013.86 (2013).

85 Poust, S., Hagen, A., Katz, L. & Keasling, J. D. Narrowing the gap between the promise and reality of polyketide synthases as a synthetic biology platform. *Curr Opin Biotechnol* **30**, 32-39, doi:10.1016/j.copbio.2014.04.011 (2014).

86 McDaniel, R. *et al.* Multiple genetic modifications of the erythromycin polyketide synthase to produce a library of novel "unnatural" natural products. *Proc Natl Acad Sci U S A* **96**, 1846-1851 (1999).

87 Xu, Y. *et al.* Rational reprogramming of fungal polyketide first-ring cyclization. *Proc Natl Acad Sci U S A* **110**, 5398-5403, doi:10.1073/pnas.1301201110 (2013).

88 Shen, W., Mao, H., Huang, Q. & Dong, J. Benzenediol lactones: a class of fungal metabolites with diverse structural features and biological activities. *European journal of medicinal chemistry* **97**, 747-777, doi:10.1016/j.ejmech.2014.11.067 (2015).

89 Zhou, H., Zhan, J., Watanabe, K., Xie, X. & Tang, Y. A polyketide macrolactone synthase from the filamentous fungus Gibberella zeae. *Proc Natl Acad Sci U S A* **105**, 6249-6254, doi:10.1073/pnas.0800657105 (2008).

90 Xu, Y. *et al.* Thioesterase domains of fungal nonreducing polyketide synthases act as decision gates during combinatorial biosynthesis. *J Am Chem Soc* **135**, 10783-10791, doi:10.1021/ja4041362 (2013).

91 Gao, Z. *et al.* Investigation of fungal iterative polyketide synthase functions using partially assembled intermediates. *J Am Chem Soc* **135**, 1735-1738, doi:10.1021/ja4001823 (2013).

92 Klein, J. *et al.* Yeast synthetic biology platform generates novel chemical structures as scaffolds for drug discovery. *ACS Synth Biol* **3**, 314-323, doi:10.1021/sb400177x (2014).

93 Agarwal, V. & Moore, B. S. Fungal polyketide engineering comes of age. *Proc Natl Acad Sci U S A* **111**, 12278-12279, doi:10.1073/pnas.1412946111 (2014).

94 Khaldi, N. *et al.* SMURF: Genomic mapping of fungal secondary metabolite clusters. *Fungal genetics and biology : FG & B* **47**, 736-741, doi:10.1016/j.fgb.2010.06.003 (2010).

95 Medema, M. H. *et al.* antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic acids research* **39**, W339-346, doi:10.1093/nar/gkr466 (2011).

96 Priebe, S., Linde, J., Albrecht, D., Guthke, R. & Brakhage, A. A. FungiFun: a web-based application for functional categorization of fungal genes and proteins. *Fungal genetics and biology : FG & B* **48**, 353-358, doi:10.1016/j.fgb.2010.11.001 (2011).

97 Yin, W. B. *et al.* Discovery of cryptic polyketide metabolites from dermatophytes using heterologous expression in Aspergillus nidulans. *ACS Synth Biol* **2**, 629-634, doi:10.1021/sb400048b (2013).

98 Bergmann, S. *et al.* Genomics-driven discovery of PKS-NRPS hybrid metabolites from Aspergillus nidulans. *Nat Chem Biol* **3**, 213-217, doi:10.1038/nchembio869 (2007).

99 Kupfer, D. M. *et al.* Introns and splicing elements of five diverse fungi. *Eukaryotic cell* **3**, 1088-1100, doi:10.1128/EC.3.5.1088-1100.2004 (2004).

100 Cacho, R. A., Tang, Y. & Chooi, Y. H. Next-generation sequencing approach for connecting secondary metabolites to biosynthetic gene clusters in fungi. *Front Microbiol* **5**, 774, doi:10.3389/fmicb.2014.00774 (2014).

101 Zhang, M. Q. Computational prediction of eukaryotic protein-coding genes. *Nature reviews. Genetics* **3**, 698-709, doi:10.1038/nrg890 (2002).

102 Kang, H. S. & Brady, S. F. Mining soil metagenomes to better understand the evolution of natural product structural diversity: pentangular polyphenols as a case study. *J Am Chem Soc* **136**, 18111-18119, doi:10.1021/ja510606j (2014).

103 Montiel, D., Kang, H. S., Chang, F. Y., Charlop-Powers, Z. & Brady, S. F. Yeast homologous recombination-based promoter engineering for the activation of silent natural product biosynthetic gene clusters. *Proc Natl Acad Sci U S A* **112**, 8953-8958, doi:10.1073/pnas.1507606112 (2015).

104 Campbell, C. D. & Vederas, J. C. Biosynthesis of lovastatin and related metabolites formed by fungal iterative PKS enzymes. *Biopolymers* **93**, 755-763, doi:10.1002/bip.21428 (2010).

105 Lamon-Fava, S. Statins and lipid metabolism: an update. *Curr Opin Lipidol* **24**, 221-226, doi:10.1097/MOL.0b013e3283613b8b (2013).

106 Weng, T. C., Yang, Y. H., Lin, S. J. & Tai, S. H. A systematic review and meta-analysis on the therapeutic equivalence of statins. *J Clin Pharm Ther* **35**, 139-151, doi:10.1111/j.1365-2710.2009.01085.x (2010).

107 Wagschal, K., Yoshizawa, Y., Witter, D. J., Liu, Y. Q. & Vederas, J. C. Biosynthesis of ML-236C and the hypocholesterolemic agents compactin by Penicillium aurantiogriseum and lovastatin by Aspergillus terreus: Determination of the origin of carbon, hydrogen and oxygen atoms by C-13 NMR spectrometry and observation of unusual labelling of acetate-derived oxygens by O-18(2). *J Chem Soc Perk T 1*, 2357-2363, doi:Doi 10.1039/P19960002357 (1996).

108 Berg, V. D. M. A., Hans, M. & Streekstra, H. Method for the production of simvastatin. WO2007147801A1 (2007).

109 Morgan, B. *et al.* Methods for making simvastatin and intermediates. WO2005040107A2 (2005).

110 Xie, X. & Tang, Y. Efficient synthesis of simvastatin by use of whole-cell biocatalysis. *Appl Environ Microbiol* **73**, 2054-2060, doi:10.1128/AEM.02820-06 (2007).

111 Xie, X., Watanabe, K., Wojcicki, W. A., Wang, C. C. & Tang, Y. Biosynthesis of lovastatin analogs with a broadly specific acyltransferase. *Chem Biol* **13**, 1161-1169, doi:10.1016/j.chembiol.2006.09.008 (2006).

112 Xie, X., Wong, W. W. & Tang, Y. Improving simvastatin bioconversion in Escherichia coli by deletion of bioH. *Metab Eng* **9**, 379-386, doi:10.1016/j.ymben.2007.05.006 (2007).

113 Xie, X. *et al.* Rational improvement of simvastatin synthase solubility in Escherichia coli leads to higher whole-cell biocatalytic activity. *Biotechnol Bioeng* **102**, 20-28, doi:10.1002/bit.22028 (2009).

114 Gao, X. *et al.* Directed evolution and structural characterization of a simvastatin synthase. *Chem Biol* **16**, 1064-1074, doi:10.1016/j.chembiol.2009.09.017 (2009).

115 Jimenez-Oses, G. *et al.* The role of distant mutations and allosteric regulation on LovD active site dynamics. *Nat Chem Biol* **10**, 431-436, doi:10.1038/nchembio.1503 (2014).

116 Singh, S. K. & Pandey, A. Emerging approaches in fermentative production of statins. *Appl Biochem Biotechnol* **171**, 927-938, doi:10.1007/s12010-013-0400-2 (2013).

117 Liu, Y. *et al.* Engineered monoculture and co-culture of methylotrophic yeast for de novo production of monacolin J and lovastatin from methanol. *Metab Eng* **45**, 189-199, doi:10.1016/j.ymben.2017.12.009 (2018).

118 Xu, W., Cai, X., Jung, M. E. & Tang, Y. Analysis of intact and dissected fungal polyketide synthase-nonribosomal peptide synthetase in vitro and in Saccharomyces cerevisiae. *J Am Chem Soc* **132**, 13604-13607, doi:10.1021/ja107084d (2010).

119 Brachmann, C. B. *et al.* 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 (1998).

120 Harvey, C. J. B. *et al.* HEx: A heterologous expression platform for the discovery of fungal natural products. *Science advances* **4**, eaar5459, doi:10.1126/sciadv.aar5459 (2018).

121 Tang, M. C. *et al.* Discovery of Unclustered Fungal Indole Diterpene Biosynthetic Pathways through Combinatorial Pathway Reassembly in Engineered Yeast. *J Am Chem Soc* **137**, 13724-13727, doi:10.1021/jacs.5b06108 (2015).

122 Horwitz, A. A. *et al.* Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas. *Cell systems* **1**, 88-96, doi:10.1016/j.cels.2015.02.001 (2015).

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

124 Heigwer, F., Kerr, G. & Boutros, M. E-CRISP: fast CRISPR target site identification. *Nature methods* **11**, 122-123, doi:10.1038/nmeth.2812 (2014).

125 Naito, Y., Hino, K., Bono, H. & Ui-Tei, K. CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites. *Bioinformatics (Oxford, England)* **31**, 1120-1123, doi:10.1093/bioinformatics/btu743 (2015).

126 Engel, S. R. *et al.* The reference genome sequence of Saccharomyces cerevisiae: then and now. *G3 (Bethesda, Md.)* **4**, 389-398, doi:10.1534/g3.113.008995 (2014).

127 Li, M. Z. & Elledge, S. J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nature methods* **4**, 251-256, doi:10.1038/nmeth1010 (2007).

128 Curran, K. A., Karim, A. S., Gupta, A. & Alper, H. S. Use of expression-enhancing terminators in Saccharomyces cerevisiae to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metab Eng* **19**, 88-97, doi:10.1016/j.ymben.2013.07.001 (2013).

129 Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* **2**, 31-34, doi:10.1038/nprot.2007.13 (2007).

130 Looke, M., Kristjuhan, K. & Kristjuhan, A. Extraction of genomic DNA from yeasts for PCR-based applications. *BioTechniques* **50**, 325-328, doi:10.2144/000113672 (2011).

131 Abe, Y. *et al.* Molecular cloning and characterization of an ML-236B (compactin) biosynthetic gene cluster in Penicillium citrinum. *Molecular genetics and genomics : MGG* **267**, 636-646, doi:10.1007/s00438-002-0697-y (2002).

132 Cardenas, J. & Da Silva, N. A. Metabolic engineering of Saccharomyces cerevisiae for the production of triacetic acid lactone. *Metabolic engineering* **25**, 194-203 (2014).

133 Mutka, S. C., Bondi, S. M., Carney, J. R., Da Silva, N. A. & Kealey, J. T. Metabolic pathway engineering for complex polyketide biosynthesis in Saccharomyces cerevisiae. *FEMS yeast research* **6**, 40-47 (2006).

134 Ma, S. M. *et al.* Complete reconstitution of a highly reducing iterative polyketide synthase. *Science* **326**, 589-592 (2009).

135 Lee, K. M. & DaSilva, N. A. Evaluation of the Saccharomyces cerevisiae ADH2 promoter for protein synthesis. *Yeast* **22**, 431-440, doi:10.1002/yea.1221 (2005).

136 Auclair, K., Kennedy, J., Hutchinson, C. R. & Vederas, J. C. Conversion of cyclic nonaketides to lovastatin and compactin by a lovC deficient mutant of Aspergillus terreus. *Bioorg Med Chem Lett* **11**, 1527-1531 (2001).

137 Lee, M. E., DeLoache, W. C., Cervantes, B. & Dueber, J. E. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS synthetic biology* (2015).

138 Da Silva, N. A. & Srikrishnan, S. Introduction and expression of genes for metabolic engineering applications in Saccharomyces cerevisiae. *FEMS yeast research* **12**, 197-214 (2012).

139 Galanie, S. & Smolke, C. D. Optimization of yeast-based production of medicinal protoberberine alkaloids. *Microbial cell factories* **14**, 1 (2015).

140 Jarboe, L. R. Improving the success and impact of the metabolic engineering design, build, test, learn cycle by addressing proteins of unknown function. *Curr Opin Biotechnol* **53**, 93-98, doi:10.1016/j.copbio.2017.12.017 (2018).

141 Renault, H., Bassard, J. E., Hamberger, B. & Werck-Reichhart, D. Cytochrome P450-mediated metabolic engineering: current progress and future challenges. *Current opinion in plant biology* **19**, 27-34, doi:10.1016/j.pbi.2014.03.004 (2014).

142 Trenchard, I. J. & Smolke, C. D. Engineering strategies for the fermentative production of plant alkaloids in yeast. *Metab Eng* **30**, 96-104, doi:10.1016/j.ymben.2015.05.001 (2015).

143 Arendt, P. *et al.* An endoplasmic reticulum-engineered yeast platform for overproduction of triterpenoids. *Metab Eng* **40**, 165-175, doi:10.1016/j.ymben.2017.02.007 (2017).

144 Maciejak, A. *et al.* The effects of statins on the mevalonic acid pathway in recombinant yeast strains expressing human HMG-CoA reductase. *BMC biotechnology* **13**, 68 (2013).

145 Kennedy, J. *et al.* Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science* **284**, 1368-1372 (1999).

146 Hutchinson, C. R. *et al.* Aspects of the biosynthesis of non-aromatic fungal polyketides by iterative polyketide synthases. *Antonie Van Leeuwenhoek* **78**, 287-295 (2000).

147 Xie, X. & Tang, Y. Efficient synthesis of simvastatin by use of whole-cell biocatalysis. *Applied and environmental microbiology* **73**, 2054-2060 (2007).

148 Mikkelsen, M. D. *et al.* Microbial production of indolylglucosinolate through engineering of a multi-gene pathway in a versatile yeast expression platform. *Metab Eng* **14**, 104-111, doi:10.1016/j.ymben.2012.01.006 (2012).

149 Steensels, J. *et al.* Improving industrial yeast strains: exploiting natural and artificial diversity. *FEMS Microbiol Rev* **38**, 947-995, doi:10.1111/1574-6976.12073 (2014).

150 Fang, F. *et al.* A vector set for systematic metabolic engineering in Saccharomyces cerevisiae. *Yeast* **28**, 123-136, doi:10.1002/yea.1824 (2011).

151 Flagfeldt, D. B., Siewers, V., Huang, L. & Nielsen, J. Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae. *Yeast* **26**, 545-551, doi:10.1002/yea.1705 (2009).

152 Jensen, N. B. *et al.* EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae. *FEMS Yeast Res* **14**, 238-248, doi:10.1111/1567-1364.12118 (2014).

153 Lee, M. E., DeLoache, W. C., Cervantes, B. & Dueber, J. E. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth Biol* **4**, 975-986, doi:10.1021/sb500366v (2015).

154 Valli, M. *et al.* Intracellular pH distribution in Saccharomyces cerevisiae cell populations, analyzed by flow cytometry. *Appl Environ Microbiol* **71**, 1515-1521, doi:10.1128/AEM.71.3.1515-1521.2005 (2005).

155 Khattak, W. A. *et al.* Developmental strategies and regulation of cell-free enzyme system for ethanol production: a molecular prospective. *Appl Microbiol Biotechnol* **98**, 9561-9578, doi:10.1007/s00253-014-6154-0 (2014).

156 You, C. & Zhang, Y. H. Cell-free biosystems for biomanufacturing. *Advances in biochemical engineering/biotechnology* **131**, 89-119, doi:10.1007/10_2012_159 (2013).

157 Yang-Boja, E., DeFilippes, F. & Fales, H. M. Electrospray mass spectra of three proprietary detergents. *Anal Biochem* **285**, 205-210, doi:10.1006/abio.2000.4734 (2000).

158 Yamamoto, E., Yamaguchi, S. & Nagamune, T. Synergistic effects of detergents and organic solvents on protein refolding: control of aggregation and folding rates. *J Biosci Bioeng* **111**, 10-15, doi:10.1016/j.jbiosc.2010.08.016 (2011).

159 Boder, E. T. & Wittrup, K. D. Yeast surface display for screening combinatorial polypeptide libraries. *Nature biotechnology* **15**, 553-557, doi:10.1038/nbt0697-553 (1997).

160 Cherf, G. M. & Cochran, J. R. Applications of Yeast Surface Display for Protein Engineering. *Methods in molecular biology* **1319**, 155-175, doi:10.1007/978-1-4939-2748-7_8 (2015).

161 Pepper, L. R., Cho, Y. K., Boder, E. T. & Shusta, E. V. A decade of yeast surface display technology: where are we now? *Combinatorial chemistry & high throughput screening* **11**, 127-134 (2008).

162 Kondo, A. & Ueda, M. Yeast cell-surface display--applications of molecular display. *Appl Microbiol Biotechnol* **64**, 28-40, doi:10.1007/s00253-003-1492-3 (2004).

163 Dranginis, A. M., Rauceo, J. M., Coronado, J. E. & Lipke, P. N. A biochemical guide to yeast adhesins: glycoproteins for social and antisocial occasions. *Microbiology and molecular biology reviews : MMBR* **71**, 282-294, doi:10.1128/MMBR.00037-06 (2007).

164 Chen, I., Dorr, B. M. & Liu, D. R. A general strategy for the evolution of bond-forming enzymes using yeast display. *Proc Natl Acad Sci U S A* **108**, 11399-11404, doi:10.1073/pnas.1101046108 (2011).

165 Ueda, M. & Tanaka, A. Cell surface engineering of yeast: construction of arming yeast with biocatalyst. *J Biosci Bioeng* **90**, 125-136 (2000).

166 Wang, Z., Mathias, A., Stavrou, S. & Neville, D. M., Jr. A new yeast display vector permitting free scFv amino termini can augment ligand binding affinities. *Protein engineering, design & selection : PEDS* **18**, 337-343, doi:10.1093/protein/gzi036 (2005).

167 Rakestraw, J. A., Sazinsky, S. L., Piatesi, A., Antipov, E. & Wittrup, K. D. Directed evolution of a secretory leader for the improved expression of heterologous proteins and full-length antibodies in Saccharomyces cerevisiae. *Biotechnol Bioeng* **103**, 1192-1201, doi:10.1002/bit.22338 (2009).

168 Liu, Z., Tyo, K. E., Martinez, J. L., Petranovic, D. & Nielsen, J. Different expression systems for production of recombinant proteins in Saccharomyces cerevisiae. *Biotechnol Bioeng* **109**, 1259-1268, doi:10.1002/bit.24409 (2012).

169 Eiden-Plach, A. *et al.* Viral preprotoxin signal sequence allows efficient secretion of green fluorescent protein by Candida glabrata, Pichia pastoris, Saccharomyces cerevisiae, and Schizosaccharomyces pombe. *Appl Environ Microbiol* **70**, 961-966 (2004).

170 Riffer, F., Eisfeld, K., Breinig, F. & Schmitt, M. J. Mutational analysis of K28 preprotoxin processing in the yeast Saccharomyces cerevisiae. *Microbiology* **148**, 1317-1328, doi:10.1099/00221287-148-5-1317 (2002).

171 Schmitt, M. J. Cloning and expression of a cDNA copy of the viral K28 killer toxin gene in yeast. *Mol Gen Genet* **246**, 236-246 (1995).

172 Baganz, F., Hayes, A., Marren, D., Gardner, D. C. & Oliver, S. G. Suitability of replacement markers for functional analysis studies in Saccharomyces cerevisiae. *Yeast* **13**, 1563-1573, doi:10.1002/(SICI)1097-0061(199712)13:16<1563::AID-YEA240>3.0.CO;2-6 (1997).

173 Kim, J. M., Vanguri, S., Boeke, J. D., Gabriel, A. & Voytas, D. F. Transposable elements and genome organization: a comprehensive survey of retrotransposons revealed by the complete Saccharomyces cerevisiae genome sequence. *Genome research* **8**, 464-478 (1998).