UNIVERSITY OF CALIFORNIA
RIVERSIDE

Establishment of Phytosterol-Producing Yeast Platforms for Plant Enzyme Characterization and Brassinosteroid Biosynthesis Elucidation

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemical and Environmental Engineering

by

Shanhui Xu

December 2022

Dissertation Committee:
Dr. Yanran Li, Chairperson
Dr. Ian Wheeldon
Dr. Robert Jinkerson

The Dissertation of Shanhui Xu is approved:

Committee Chairperson

University of California, Riverside

Acknowledgement

First and foremost I am deeply grateful to my supervisor, Dr. Yanran Li, for her valuable and inspiring advice, continuous support during my Ph.D. Her advice and support are very important to me, not only for the research, but also for my future career development and personal life. I would also like to thank her for encouraging me when I felt stressful and frustrated during my Ph.D. study. I would also like to thank to my committee members Dr. Ian Wheeldon and Dr. Robert Jinkerson, for taking time to attend my defense and read the thesis. A special thanks to Dr. Takahito Nomura and Dr. Yinsheng Wang, our collaborators, for kindly sharing their knowledge and providing perspective suggestions to my research. I would like to thank all the lab members and collaborates. Their help and support created a harmonious environment in our lab.

I must thank my parents, who are very patient and understanding and support me as always, even when I cannot spend more time with them during these years. Plenty thanks to my friends as well, for sharing laughter and tears with me and make a lot of fun in my life. Without them, it would be harder during Ph.D. study.

Chapter I of this dissertation is published on ACS Synthetic Biology and adapted with permission from Xu, S.; Chen, C.; Li, Y., "Engineering of Phytosterol-Producing Yeast Platforms for Functional Reconstitution of Downstream Biosynthetic Pathways", which can be found on ACS Synth Biol 2020, 9 (11), 3157-3170. https://pubs.acs.org/doi/10.1021/acssynbio.0c00417. Copyright 2020 American Chemical Society.

ABSTRACT OF THE DISSERTATION

Establishment of Phytosterol-Producing Yeast Platforms for Plant Enzyme Characterization and Brassinosteroid Biosynthesis Elucidation

by

Shanhui Xu

Doctor of Philosophy, Graduate Program in Chemical and Environmental Engineering
University of California, Riverside, December 2022
Dr. Yanran Li, Chairperson

Phytosterols that plays important roles in regulating membrane properties and serves as the precursor to multiple specialized metabolites such as the phytohormone brassinosteroids. Brassinosteroids are essential and universal plant hormones, regulating plant growth and development, also involved in biotic and abiotic defense. With the promising potential of agriculture application, the biosynthesis of brassinolide in plants remains elusive although it has been intensively studied in the past decades. The promiscuity of brassinosteroid biosynthetic enzymes and the low abundance of brassinosteroids in nature is likely the reasons. In this project, to decipher the biosynthesis of brassinosteroids and pave the way for other steroid nature products (e.g. withanolides, saponins), we first reconstituted the biosynthetic pathway of phytosterols in a heterologous host *Saccharomyces cerevisiae* and engineered the yeast to produce the free phytosterols and incorporate the sterols in the cell membrane. The resulted phytosterol-producing yeast

strains provided a platform for enhanced reconstitution of plant membrane enzymes. Then, we optimized the growth of the strain by engineering sterol esterification mechanism, tuning upstream FPP pathway, and performing retro genetic engineering. Using the optimized phytosterol-producing yeast platform, a plant scaffold protein MSBP1 was characterized to play an important role in BR biosynthesis, which expanded our understanding of this multifaced scaffold protein. More importantly, it implies the possibility that brassinosteroid biosynthesis takes a place in the metabolon. This finding highlighted the role of non-enzymatic proteins in nature product biosynthesis in plant and broaden our vision to investigate plant secondary metabolism.

Table of Contents

Chapter III: A Multifunctional Scaffold Protein MSBP1 Enhanced the Activities of BR Biosynthetic Enzymes

List of Figures

Chapter I: Engineering of Phytosterol-Producing Yeast Platforms for Functional Reconstitution of Downstream Biosynthetic Pathways

Chapter II: Optimization of Campesterol-Producing Yeast Strains for Plant Membrane-Bound Enzyme Reconstitution

Chapter III: A Multifunctional Scaffold Protein MSBP1 Enhanced the Activities of BR Biosynthetic Enzymes

List of Tables

Chapter I: Engineering of Phytosterol-Producing Yeast Platforms for Functional Reconstitution of Downstream Biosynthetic Pathways

Introduction

Sterols are essential membrane components that regulate membrane fluidity and permeability in eukaryotic organisms1. Animals generally synthesize cholesterol, fungi produce ergosterol, and plants utilize an array of phytosterols to fully functionalize. Phytosterols are cholesterol-like sterols that exhibit distinct side chains from cholesterol. Similar to cholesterol in humans and ergosterol in yeast, phytosterols are essential membrane components and have important physiological functions[2], such as temperature adaption[3], transmembrane signal transduction[4]. The composition of sterols in membrane plays an essential role in various physiological and biochemical progresses[5]. For example, the ratio between two dominant phytosterols, campesterol and β-sitosterol, modulates the growth of plants[6]. The repression of campesterol biosynthesis in plants could result in defected growth with defected cell elongation and reduced fertility[6], due to the decrease production of plant hormone brassinosteroids, which is synthesized from campesterol.

Brassinosteroids (BRs), a group of essential plant hormones derived from campesterol, are widely found in the plant kingdom, and play important roles in growth and development, such as seed germination, cell division, and pollen elongation[7, 8]. Brassinolide (BL) is the most bioactive brassinosteroids in nature, which has a very unique structure that carries a lactone moiety in the B-ring and a pair of vicinal diols in the A-ring[9]. The BR defective mutants usually exhibit dwarf phenotype. In addition, many studies suggest that BRs are required for the resistance of unfavorable abiotic stress, such as high salinity, drought, and insufficient nutrient[10]. Regard to these properties in growth regulation and biotic and abiotic defense, BRs exhibit great potential for application in

agricultural usage and have intrigued a lot of studies on the biofunction of BRs. According to the previous research, applying BRs on trial field not only increased the yield of crops up to 56% at maximum, but also improved the quality of the crops regarding enhanced starch and Vitamin C content[11]. Furthermore, BRs are nontoxic and environmentally friendly for the application in crops yield promotion and plant protection. Besides, accumulating evidence indicate that BRs act as complex multifaceted regulators of plant immunity, whose homeostasis is significant in plant innate immunity system[7]. According to the field trials, BRs treatment exhibits a positive effect on crops resistance to a broad-spectrum of pathogens and can somehow substitute some traditional pesticides[11]. Meanwhile, other evidence suggests that the impact of BRs on plant immunity defense is not always positive. Applying BL can cause rice more susceptible to Pythium graminicola and meloidogyne graminicola[12]. The molecular mechanism of how BRs involve in plant immunity remains elusive but exhibit great significance for the further usage of BRs in agriculture. However, the low abundance of BR in nature (e.g. 10-100 mg/kg in pollen, 10-100 ng/kg in shoots and leaves), which makes them difficult to isolate, detect, and utilize. To fully elucidate the biosynthetic pathway of BRs in plants is a promising way to synthesize them and extent the application of BRs, and more importantly to understand the multifaced role of BR regulation in plants.

According to the pioneering studies[13-18], the biosynthetic pathway of BRs in plants appears a general frame from campesterol (CR) to BRs (Figure 1.1). Many intermediates and BR-related genes have been identified and characterized, but several enzymatic steps remain elusive. Most of these studies employed isotopic-labeling and Arabidopsis thaliana

mutants phenotypic rescue as the major approaches to investigate BR biosynthesis in plants[19, 20]. And in vitro assays are conducted to characterize the enzymes. These methods are limited when it comes to a highly crosslinked synthetic pathway like BR. Plant mutant analysis heavily depends on the phenotype recognition and the analysis of intermediates accumulation which is not very reliable especially within the branched biosynthetic background[15, 16]. The promiscuity of cytochrome P450s also makes it difficult to find their native substrates and connect the enzymes and intermediates to elucidate the pathway. So far, although most of the biosynthetic enzymes and important intermediates in the BR biosynthesis have been identified, there are still components in the biosynthetic pathway missing.

Heterologous hosts are favorable alternatives to characterize the complicated BR biosynthetic pathways. Massive investigations on the plant pathway reconstitution in microbial heterologous hosts have highlighted the potential of using them for pathway elucidation and enzyme characterization in natural product biosynthesis. *Saccharomyces. cerevisiae* has been widely used in plant enzyme characterization and nature product biosynthesis[21-25], such as artemisinic acid, noscapine, Taxol precursor, opioid and gibberellin. Therefore, using *S. cerevisiae* to reconstitute and characterize BR biosynthetic enzymes could be a promising way to understand the connection between the enzymes and eventually elucidate BR biosynthetic pathway.

In this project, *S. cerevisiae* was used as the platform to reconstitute and characterize the enzymes of BR biosynthesis and eventually to elucidate the biosynthetic

pathway of BR. At the first stage, we resolved several bottlenecks and established the free phytosterol biosynthesis in yeast[26]. Campesterol and β-sitosterol, the two major phytosterols in plants, were produced in free form at 7.2 mg/L and 2.0 mg/L, respectively. The establishment of a yeast-based phytosterol biosynthetic platform enabled an efficient reconstitution of the early-stage BR biosynthetic pathway towards the synthesis of 22-hydroxycampest-4-en-3-one. Two BR biosynthetic enzyme CYP90B1 and CYP90A1 were characterized and enzymatic reaction order between them was clarified. However, the engineering of yeast to produce phytosterols instead of ergosterol deteriorated the growth of yeast, likely due to the toxicity of campesterol and the disruption of native sterol metabolism. Therefore, at the second stage, we optimized the growth of the phytosterol-producing strains for further utilization of the platform for plant enzyme characterizations. The platform was optimized by partially restoring sterol esterification mechanism, engineering upstream FPP pathway, and performing retro genetic engineering. The enhanced growth, Li^+ and heat tolerance ensures the applications of this platform to reconstitute plant membrane anchored enzymes. Moreover, the phytosterol-producing "plantinized" yeast also exhibited a clear improvement of the enzymatic activity of plant membrane-bound enzymes, not only to BR biosynthetic enzymes, but also noscapine biosynthetic cytochrome P450. At the third stage, we utilized a scaffold protein membrane steroid-binding protein (MSBP1) to enhance the the production of 22-hydroxycampest-4-en-3-one and extended the understanding of MSBP1. With the help of MSBP1, the bifunction of CYP90C1 was identified for the first time. We also found that the enhancement of MSBP1 brought to BR biosynthetic enzymes did not only rely on the

physical interactions with the synthetic enzymes. The role that MSBP1 played in yeast or in plants remains further studies, but it brought up a possibility that BR biosynthetic enzymes or MSBP1 might formed a complex. It also suggests that the missing components in BR biosynthetic pathways might not have enzymatic functions but be a scaffold protein like MSBP1.

References

1. Dufourc, E. J., Sterols and membrane dynamics. *J Chem Biol* **2008**, 1 (1-4), 63-77.

2. Zauber, H.; Burgos, A.; Garapati, P.; Schulze, W. X., Plasma membrane lipid-protein interactions affect signaling processes in sterol-biosynthesis mutants in Arabidopsis thaliana. *Frontiers in plant science* **2014**, 5, 78.

3. Dufourc, E. J., The role of phytosterols in plant adaptation to temperature. *Plant Signal Behav* **2008**, 3 (2), 133-4.

4. Laloi, M.; Perret, A. M.; Chatre, L.; Melser, S.; Cantrel, C.; Vaultier, M. N.; Zachowski, A.; Bathany, K.; Schmitter, J. M.; Vallet, M.; Lessire, R.; Hartmann, M. A.; Moreau, P., Insights into the role of specific lipids in the formation and delivery of lipid microdomains to the plasma membrane of plant cells. *Plant physiology* **2007**, 143 (1), 461-72.

5. Du, Y.; Fu, X.; Chu, Y.; Wu, P.; Liu, Y.; Ma, L.; Tian, H.; Zhu, B., Biosynthesis and the Roles of Plant Sterols in Development and Stress Responses. *Int J Mol Sci* **2022**, 23 (4).

6. Schaeffer, A.; Bronner, R.; Benveniste, P.; Schaller, H., The ratio of campesterol to sitosterol that modulates growth in Arabidopsis is controlled by STEROL METHYLTRANSFERASE 2;1. *The Plant journal : for cell and molecular biology* **2001**, 25 (6), 605-15.

7. De Bruyne, L.; Hofte, M.; De Vleesschauwer, D., Connecting growth and defense: the emerging roles of brassinosteroids and gibberellins in plant innate immunity. *Molecular plant* **2014**, 7 (6), 943-59.

8. Bajguz, A., Metabolism of brassinosteroids in plants. *Plant physiology and biochemistry* : PPB **2007**, 45 (2), 95-107.

9. Yokota, T., The structure, biosynthesis and function of brassinosteroids. *Trends in plant science* **1997**, 2 (4), 137-143.

10. Eremina, M.; Unterholzner, S. J.; Rathnayake, A. I.; Castellanos, M.; Khan, M.; Kugler, K. G.; May, S. T.; Mayer, K. F.; Rozhon, W.; Poppenberger, B., Brassinosteroids participate in the control of basal and acquired freezing tolerance of plants. *Proceedings of the National Academy of Sciences of the United States of America* **2016**, 113 (40), E5982-E5991.

11. Khripach, V.; Zhabinskii, V.; De Groot, A., Twenty years of brassinosteroids: Steroidal plant hormones warrant better crops for the XXI century. *Ann Bot-London* **2000**, 86 (3), 441-447.

12. Nahar, K.; Kyndt, T.; Hause, B.; Hofte, M.; Gheysen, G., Brassinosteroids Suppress Rice Defense Against Root-Knot Nematodes Through Antagonism With the Jasmonate Pathway. *Mol Plant Microbe In* **2013**, 26 (1), 106-115.

13. Acebedo, S. L.; Alonso, F.; Galagovsky, L. R.; Ramirez, J. A., Synthesis and biological activity of ring-A difluorinated brassinosteroids. *Steroids* **2011**, 76 (10-11), 1016-1020.

14. Fujioka, S.; Yokota, T., Biosynthesis and metabolism of brassinosteroids. *Annual review of plant biology* **2003**, 54, 137-64.

15. Ohnishi, T.; Godza, B.; Watanabe, B.; Fujioka, S.; Hategan, L.; Ide, K.; Shibata, K.; Yokota, T.; Szekeres, M.; Mizutani, M., CYP90A1/CPD, a brassinosteroid biosynthetic cytochrome P450 of Arabidopsis, catalyzes C-3 oxidation. *The Journal of biological chemistry* **2012**, 287 (37), 31551-60.

16. Fujioka, S.; Takatsuto, S.; Yoshida, S., An early C-22 oxidation branch in the brassinosteroid biosynthetic pathway. *Plant physiology* **2002**, 130 (2), 930-939.

17. Fujita, S.; Ohnishi, T.; Watanabe, B.; Yokota, T.; Takatsuto, S.; Fujioka, S.; Yoshida, S.; Sakata, K.; Mizutani, M., Arabidopsis CYP90B1 catalyses the early C-22 hydroxylation of C27, C28 and C29 sterols. *The Plant journal : for cell and molecular biology* **2006**, 45 (5), 765-74.

18. Ohnishi, T.; Szatmari, A. M.; Watanabe, B.; Fujita, S.; Bancos, S.; Koncz, C.; Lafos, M.; Shibata, K.; Yokota, T.; Sakata, K.; Szekeres, M.; Mizutani, M., C-23 hydroxylation by Arabidopsis CYP90C1 and CYP90D1 reveals a novel shortcut in brassinosteroid biosynthesis. *The Plant cell* **2006**, 18 (11), 3275-88.

19. Noguchi, T.; Fujioka, S.; Choe, S.; Takatsuto, S.; Yoshida, S.; Yuan, H.; Feldmann, K. A.; Tax, F. E., Brassinosteroid-insensitive dwarf mutants of Arabidopsis accumulate brassinosteroids. *Plant physiology* **1999**, 121 (3), 743-52.

20. Noguchi, T.; Fujioka, S.; Choe, S.; Takatsuto, S.; Tax, F. E.; Yoshida, S.; Feldmann, K. A., Biosynthetic pathways of brassinolide in Arabidopsis. *Plant physiology* **2000**, 124 (1), 201-9.

21. Ro, D. K.; Paradise, E. M.; Ouellet, M.; Fisher, K. J.; Newman, K. L.; Ndungu, J. M.; Ho, K. A.; Eachus, R. A.; Ham, T. S.; Kirby, J.; Chang, M. C.; Withers, S. T.; Shiba, Y.; Sarpong, R.; Keasling, J. D., Production of the antimalarial drug

precursor artemisinic acid in engineered yeast. *Nature* **2006**, 440 (7086), 940-3.

22. Li, Y.; Smolke, C. D., Engineering biosynthesis of the anticancer alkaloid noscapine in yeast. *Nat Commun* **2016**, 7, 12137.

23. Nett, R. S.; Montanares, M.; Marcassa, A.; Lu, X.; Nagel, R.; Charles, T. C.; Hedden, P.; Rojas, M. C.; Peters, R. J., Elucidation of gibberellin biosynthesis in bacteria reveals convergent evolution. *Nat Chem Biol* **2017**, 13 (1), 69-74.

24. Galanie, S.; Thodey, K.; Trenchard, I. J.; Filsinger Interrante, M.; Smolke, C. D., Complete biosynthesis of opioids in yeast. *Science* **2015**, 349 (6252), 1095-100.

25. Ajikumar, P. K.; Xiao, W. H.; Tyo, K. E.; Wang, Y.; Simeon, F.; Leonard, E.; Mucha, O.; Phon, T. H.; Pfeifer, B.; Stephanopoulos, G., Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. *Science* **2010**, 330 (6000), 70-4.

26. Xu, S.; Chen, C.; Li, Y., Engineering of Phytosterol-Producing Yeast Platforms for Functional Reconstitution of Downstream Biosynthetic Pathways. *ACS Synth Biol* **2020**, 9 (11), 3157-3170.

Chapter I: Engineering of Phytosterol-Producing Yeast Platforms for Functional Reconstitution of Downstream Biosynthetic Pathways

Abstract

As essential structural molecules for plant plasma membranes, phytosterols are key intermediates for the synthesis of many downstream specialized metabolites of pharmaceutical or agricultural significance, such as brassinosteroids and withanolides. *Saccharomyces cerevisiae* has been widely used as an alternative producer for plant secondary metabolites. Establishment of heterologous sterol pathways in yeast, however, has been challenging due to either low efficiency or structural diversity, likely a result of crosstalk between the heterologous phytosterol and the endogenous ergosterol biosynthesis. For example, in this study, we engineered campesterol production in yeast using plant enzymes; although we were able to enhance the titer of campesterol to ~40 mg/L by upregulating the mevalonate pathway, no conversion to downstream products was detected upon the introduction of downstream plant enzymes. Further investigations uncovered two interesting observations about sterol engineering in yeast. First, many heterologous sterols tend to be efficiently and intensively esterified in yeast, which drastically impedes the function of downstream enzymes. Second, yeast can overcome the growth deficiency caused by altered sterol metabolism through repeated culture. By employing metabolic engineering, strain evolution, fermentation engineering, and pathway reconstitution, we were able to establish a set of phytosterol-producing yeast strains: campesterol (~7 mg/L), β-sitosterol (~2 mg/L), 22-hydroxycampesterol (~1mg/L), and 22-hydroxycampest-4-en-3-one (~4 mg/L). This work resolves the technical bottlenecks in phytosterol-derived

pathway reconstitution in the backer's yeast and opens up opportunities for efficient bioproduction and pathway elucidation of this group of phytochemicals.

Introduction

Phytosterols are cholesterol-like sterols produced from plants that exhibit distinct side chains from cholesterol. Similar to cholesterol in humans and ergosterol in yeast, phytosterols are essential membrane components that regulate membrane fluidity and permeability in plants[1, 2]. Campesterol, **1** and β-sitosterol, **2** are the dominant sterols in plants[3], and their biosynthesis has been relatively well elucidated[4-6] (Figure 1.1). In addition to their essential role in membrane maintenance, phytosterols exhibit a broad spectrum of promising bioactivities. Phytosterols are well-documented for their cholesterol-lowering effect[3, 7] and have been listed as a dietary factor that may affect cardiovascular disease risk by the American Heart Association[8]. Phytosterols have also been reported to be associated with reduction in cancer risks[9, 10], anti-oxidant[11, 12] and anti-osteoarthritic properties[3, 13]. In addition, phytosterols serve as the synthetic precursor of downstream phytosteroids, such as the important group of plant hormones brassinosteroids (BRs), and pharmaceutically intriguing withanolides.

BRs are one important group of plant hormones involved in a number of physiological functions and behaviors[14, 15], such as cell division, vascular differentiation, and pollen elongation. BRs exhibit extremely low abundance in nature (e.g. 10-100 mg/kg in pollen, 10-100 ng/kg in shoots and leaves), which makes them difficult to isolate, detect, and utilize. In addition, BRs are of broad pharmaceutical interests with anticancer, antiviral,

and anabolic activities[16]. Brassinolide is the best studied and the most active brassinosteroid[17], and is synthesized from campesterol[18]. Most BRs are synthesized along the brassinolide biosynthetic pathway from campesterol, while a small number of others were proposed to be synthesized from β-sitosterol[19-21]. Enzymes involved in the early and final stage of brassinolide biosynthesis have been identified, while the ones catalyzing the reactions between these enzymes remain unclear (Figure 1.1). In addition, how the enzymes catalyzing early steps are arranged to direct the synthesis towards brassinolide is unknown. One of the hindrances to clearly elucidate the biosynthesis is the lack of intermediates required to examine substrate preference of the enzymes involved and reconstitute the biosynthesis *ex situ*.

Ergosterol, **3**, biosynthesis and phytosterol biosynthesis diverge as early as squalene, and shares the same intermediate episterol, **4**, , which is converted to ergosterol or campesterol in fungi or plants, respectively (Figure 1.1). The bioproduction of campesterol has been achieved in *Saccharomyces cerevisiae*[22] and *Yarrowia lipolytica*[23], through replacing the endogenous C22-sterol desaturase (ERG5) gene with a heterologous 7-dehydrocholesterol reductase (DHCR7) encoding gene. In plants, campesterol is synthesized from **4** by the function of Δ^7-sterol-C5-desaturase (DWF7), 7-dehydrocholesterol reductase (DWF5) and $\Delta^{24(28)}$-sterol reductase (DWF1)[24]. On the other hand, the bioproduction of β-sitosterol has not been established in heterologous microbial hosts, which is likely due to the lack of substrate for the 24-methylenesterol C-methyltransferase SMT2.

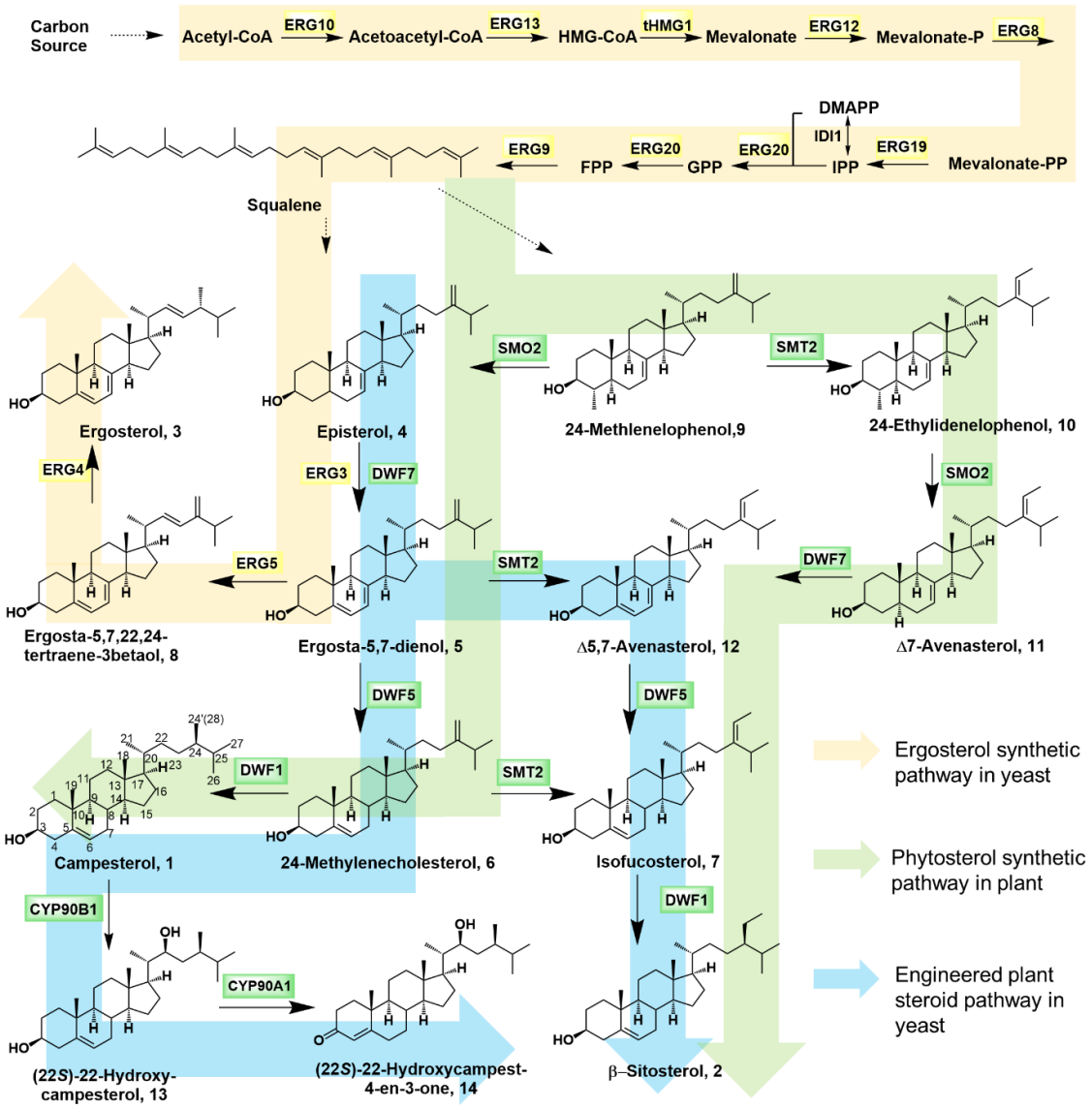


Figure 1.1 Proposed biosynthetic pathway of phytosterol and (22S)-22-Hydroxycampest-4-en-3-one, **14** in plants and yeast. The yellow arrow represents the native ergosterol, **3**, biosynthetic pathway in yeast; the green arrow represents the plant pathway of β-sitosterol, **2**; the green arrow represents the reconstituted biosynthetic pathway from episterol, **4** to **14**. Carbon numbers are marked on the structure of campesterol. SMT2: 24-methylenesterol C-methyltransferase 2; SMO2: methylsterol monooxygenase 2; ERG10: acetyl-CoA C-acetyltransferase; ERG13: 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) synthase; HMG1: HMG-CoA reductase; ERG12: mevalonate kinase; ERG8: phosphomevalonate kinase; ERG19: mevalonate pyrophosphate decarboxylase; ERG20: farnesyl pyrophosphate synthetase; ERG9: squalene synthase; ERG3: Δ^7-sterol-C5(6)-desaturase 1; ERG5: C22-sterol desaturase; ERG4: $\Delta^{24(28)}$-sterol reductase; DWF7: Δ^7-sterol-C5(6)-desaturase 1; DWF5: C7(8)-reductase; DWF1: $\Delta^{24(28)}$-sterol reductase; CYP90B1: C22-hydroxylase; CYP90A1: C3-oxidase; IPP: isopentenyl pyrophosphate; DMAPP: dimethyl allyl pyrophosphate; GPP: geranyl pyrophosphate.

Table 1.1. The overview of phytosteroid synthesized in yeast strains

Products	Genotype		Functions of Genes Introduced	Host	Titer	Ref.
	Genes Deleted	Genes Introduced				
24-methyldesmosterol	Δ*erg4*, Δ*erg5*	*24ISO*	Δ^{24}-isomerase	*S. cerevisiae*	-**	26
campesterol*	Δ*erg5*	*DHCR7*	Δ^{7}-reductase	*Y. lipolytica*	942 mg/L	23
campesterol*	Δ*erg5*	*DHCR7*	Δ^{7}-reductase	*S. cerevisiae*	-**	22
campesterol	Δ*erg4*, Δ*erg5*	*St*DWF5, *Ar*DWF1	Δ^{7}-reductase, $\Delta^{24(28)}$-reductase	*S. cerevisiae*	-**	24
cholesterol	Δ*erg5*, Δ*erg6*	*DHCR7*, *DHCR24*	Δ^{7}-reductase, $\Delta^{24(25)}$-reductase	*S. cerevisiae*	3.84 μg/10^7 cells	22
cholesterol	Δ*erg4*, Δ*erg5*	*St*DWF5, SSR2	Δ^{7}-reductase, $\Delta^{24(25)}$-reductase	*S. cerevisiae*	-**	25
22*R*-hydroxycholesterol	Δ*erg5*, Δ*erg6*	*DHCR7*, *DHCR24*, *CYP90B27*	Δ^{7}-reductase, $\Delta^{24(25)}$-reductase, 22*R*-hydroxylase	*S. cerevisiae*	-**	27
diosgenin	Δ*erg5*, Δ*erg6*	*CYP90G4 or CYP90B50*, *CYP94D108 or CYP82J17*	16*S*,22*S*-dihydroxylase, 27-hydroxylase	*S. cerevisiae*	-**	28

* Likely 24-epicampesterol due to the lack of the plant $\Delta^{24(28)}$-sterol reductase DWF1.
** Titer not reported.

Massive investigations on plant pathway reconstitution in heterologous hosts have highlighted the potential of using microbial hosts (e.g. *S. cerevisiae*) for pathway elucidation and enzyme characterization in natural product biosynthesis (Table 1.1). *S. cerevisiae* has been utilized for enzyme characterization in phytosterol and derivative biosynthesis. For instance, the desmosterol-producing yeast and 24-methylenecholesterol-producing yeast constructed through expressing DWF5 into *erg6* and *erg4*/*erg5* inactivated strains, respectively, have been utilized to characterize sterol $\Delta^{24(25)}$-reductase involved in cholesterol biosynthesis[25], novel sterol $\Delta^{24(28)}$-reductase involved in campesterol

production[24], and Δ^{24}-isomerase putatively used in withanolide biosynthesis[26]. The cholesterol-producing yeast constructed through expressing $\Delta^{24(25)}$-dehydrocholesterol reductase (DHCR24) and DHCR7 in the *erg5*/*erg6* inactivated yeast strain has been utilized to confirm the function of CYP90B27 in synthesizing 22(*R*)-hydroxycholesterol from cholesterol[27]. The same cholesterol-producing yeast strain has also been utilized to reconstruct diosgenin biosynthesis with the expression of *Pp*CYP90G4/*Pp*CYP94D108 or *Tf*CYP90B50/*Tf*CYP82J17[28]. However, no β-sitosterol or campesterol-derived compounds have been successfully synthesized in yeast.

Here, we resolved several bottlenecks in phytosteroid-related pathway reconstitution in *S. cerevisiae* and demonstrated the establishment of a yeast-based biosynthetic platform that enabled an efficient reconstitution of the early-stage brassinolide biosynthetic pathway towards the synthesis of 22-hydroxycampest-4-en-3-one, **14**. The establishment of the **14**-producing yeast strain provides a platform for the furfure reconstitution and elucidation of the biosynthesis of phytosterol-derived natural products such as brassinosteroid and withanolide.

Results

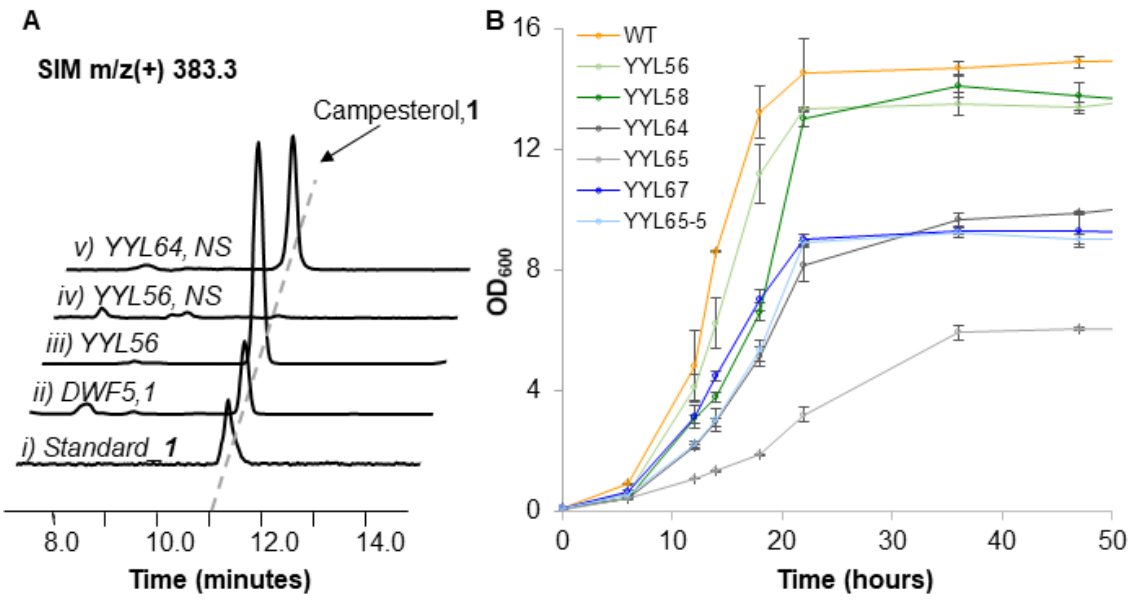
Construction and Optimization of *De Novo* Campesterol Production Using Plant Enzymes in Yeast.

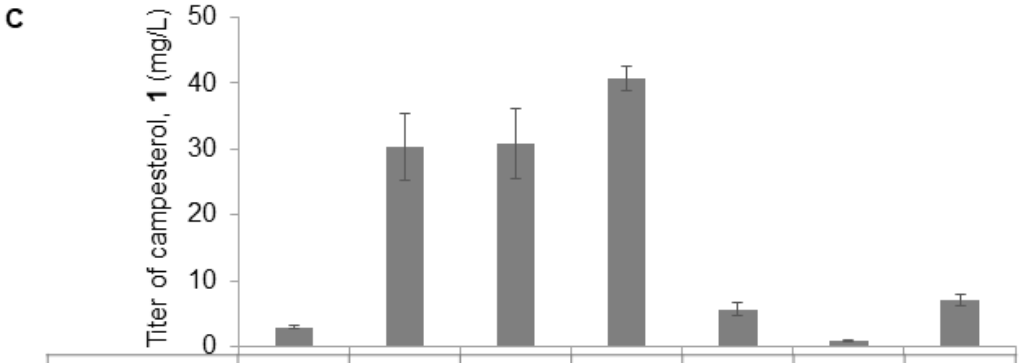
Campesterol, **1** and ergosterol, **3** are both biosynthesized from episterol, **4**, and this enables the redirection of yeast ergosterol synthetic pathway to the synthesis of campesterol (Figure 1.1). Plants use DWF7, DWF5 and DWF1 to convert **4** to campesterol

(Figure 1.1). DWF7, the ortholog of ERG3, catalyzes the conversion of **4** to ergostra-5,7-sterol, **5**. **5** is then converted to 24-methylenecholesterol, **6** through C7-C8 enol reduction by DWF5, followed by C24-C28 enol reduction catalyzed by DWF1 to afford the formation of campesterol. Since DWF7 and ERG3 are functional exchangeable, heterologous expression of DWF5 and DWF1 should lead to the synthesis of campesterol in yeast. The *dwf5* and *dwf1* genes from *Arabidopsis thaliana* were introduced into yeast with low-copy number plasmids and regulated by the yeast constitutive *GPD* promoter. The sterols were extracted from the engineered yeast cells after growing in synthetic defined medium (SDM) for 3 days at 30^{o}C by the standard saponification protocol. The extracts were analyzed with liquid chromatography-mass spectrometry (LC-MS)[2930] and the production of campesterol was confirmed by comparing with the authentic campesterol standard (Figure 1.2A). Further introduction of *dwf7* to the *dwf5*/*dwf1*-expressing strain resulted in significant enhancement in campesterol production at 3.37±0.07 mg/L, 32-fold higher than the *dwf5/dwf1*-expressing strain (Figure S1.1A), and minor effects on yeast growth (Figure 1.2B). As a comparison, introducing an extra copy of *erg3*, the *dwf7* ortholog in yeast, also resulted in the similar boost of campesterol production (Figure S1.1A), which suggests that ERG3 or DWF7 might catalyze a rate-limiting step in campesterol biosynthesis in yeast. Additionally, introducing *dwf7, dwf5* and *dwf1* genes to yeast strain greatly impacted the synthesis of ergosterol, which was barely detected from the engineered campesterol-producing yeast strains (Figure S1.1B). The *dwf7, dwf5 and dwf1* genes were then integrated onto the genome to afford a stable campesterol-producing strain for further engineering (YYL55).

Table 1.2. Yeast strains used in this study.

Strain	Genotype
CENPK2.1D	*MATα, ura3-52, trp1-289, leu2-3,112 his3Δ1, MAL2-8^{C}, SUC2*
YYL55	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, kanmx, P_{GPD}-dwf7-T_{CYC1}*
YYL56	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, P_{GPD}-dwf7-T_{CYC1}* *ybl059wΔ::P_{PGK1}-erg12-T_{PHO5}, P_{TEF1}-erg10-T_{CYC1}, P_{PYK1}-tHmg1-T_{MFA1}, P_{TPI1}-erg13-T_{STE2}*
YYL57	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, P_{GPD}-dwf7-T_{CYC1}* *ybl059wΔ::P_{PGK1}-erg12-T_{PHO5}, P_{TEF1}-erg10-T_{CYC1}, P_{PYK1}-tHmg1-T_{MFA1}, P_{TPI1}-erg13-T_{STE2}* *ybr197cΔ::P_{TPI1}-erg8-T_{STE2}, P_{TEF1}-erg19-T_{CYC1}, P_{GPD}-idi1-T_{CYC1}, P_{PYK1}-tHmg1-T_{MFA1}*
YYL58	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, P_{GPD}-dwf7-T_{CYC1}* *ybl059wΔ::P_{PGK1}-erg12-T_{PHO5}, P_{TEF1}-erg10-T_{CYC1}, P_{PYK1}-tHmg1-T_{MFA1}, P_{TPI1}-erg13-T_{STE2}* *ybr197cΔ::P_{TPI1}-erg8-T_{STE2}, P_{TEF1}-erg19-T_{CYC1} P_{GPD}-idi1-T_{CYC1}, P_{PYK1}-thmg1-T_{MFA1}* *ymr206wΔ::P_{TEF1}-erg20-T_{CYC1}, P_{GPD1}-upc2 -T_{ADH1}, P_{PYK1}-tHmg1-T_{MFA1}, P_{PGK1}-erg7-T_{PHO5}*
YYL60	*erg5Δ::dhcr7*
YYL63	*are1Δ, are2Δ*
YYL64	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, P_{GPD}-dwf7-T_{CYC1}* *ybl059wΔ::P_{PGK1}-erg12-T_{PHO5}, P_{TEF1}-erg10-T_{CYC1}, P_{PYK1}-tHmg1-T_{MFA1}, P_{TPI1}-erg13-T_{STE2}* *are1Δ, are2Δ*
YYL65	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, P_{GPD}-dwf7-T_{CYC1}* *ybl059wΔ::P_{PGK1}-erg12-T_{PHO5}, P_{TEF1}-erg10-T_{CYC1}, P_{PYK1}-tHmg1-T_{MFA1}, P_{TPI1}-erg13-T_{STE2}* *are1Δ, are2Δ* *erg4Δ::ura3*
YYL66	*are1Δ, are2Δ* *erg4Δ::ura3*
YYL67	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, P_{GPD}-dwf7-T_{CYC1}* *ybl059wΔ::P_{PGK1}-erg12-T_{PHO5}, P_{TEF1}-erg10-T_{CYC1}, P_{PYK1}-tHmg1-T_{MFA1}, P_{TPI1}-erg13-T_{STE2}* *are1Δ, are2Δ* *erg4Δ*
YYL69	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, P_{GPD}-dwf7-T_{CYC1}* *ybl059wΔ::P_{PGK1}-erg12-T_{PHO5}, P_{TEF1}-erg10-T_{CYC1}, P_{PYK1}-tHmg1-T_{MFA1}, P_{TPI1}-erg13-T_{STE2}* *ybr197cΔ::P_{TPI1}-CYP90A1-T_{PHO5}, P_{GPD}-CYP90B1-T_{CYC1}* *are1Δ, are2Δ* *erg4Δ*





Strain	YYL55	YYL56	YYL57	YYL58	YYL64	YYL65	YYL67
mg/L	3.02	30.37	30.79	40.74	5.72	1.00	7.13
OD_{600}	13.51 ±0.49	13.87 ±0.11	13.68 ±0.18	13.62 ±0.23	10.52 ±0.26	6.12 ±0.30	9.48 ±0.18
DWF1/5/7	+	+	+	+	+	+	+
cassette 1		+	+	+	+	+	+
cassette 2			+	+			
cassette 3				+			
Δare1/ Δare2					+	+	+
Δerg4						+	+
Evolved							+

Figure 1.2 Biosynthesis of campesterol in yeast. **(A)** Selected ion monitoring (SIM) extracted ion chromatogram (EIC) using campesterol's characteristic m/z^+ signal (MW=400.69Da, $[C_{28}H_{49}O\text{-}H_2O]^+=[C_{28}H_{47}]^+$=383.4) of i) campesterol standard, ii) yeast strain harboring DWF5 and DWF1, iv) YYL56, v) YYL56 without KOH saponification, vi) YYL64 (YYL56 with *are1*Δ*are2*Δ*)* without KOH saponification. NS: non-saponification extraction method. Yeast strains were extracted with saponification if not specified. **(B)** Growth of strains engineered for campesterol production. The yeast strains were cultured in synthetic complete SDM medium supplemented with 2% (w/v) glucose at 30°C for 72 hours. WT represents wildtype yeast strain. Bars represent mean values of at least three biological replicates, and the error bars represent the standard deviation of the replicates. **(C)** Campesterol production from yeast strains harboring DWF7/DWF5/DWF1 with enhanced expression level of mevalonate (MVA) pathway, *are1*/*are2* inactivation, and *erg4* inactivation. Genes overexpressed in each MVA-enhancing cassette: MVA1: *erg12*, *erg13*, *erg8*, *erg19*; MVA2: *erg8*, *erg19*, *idi1*, *tHmg1*; MVA3: *erg20*, *upc2*, *tHmg1*, *erg7*. The campesterol-producing yeast strains were cultured in synthetic complete SDM medium supplemented with 2% (w/v) glucose at 30°C for 72 hours. OD_{600} values of the culture at the collecting time were included. All traces are representative of at least three biological replicates for each engineered yeast strain. Bars represent mean values of three biological replicates, and the error bars represent the standard deviation of the replicates.

To increase the campesterol production to a level sufficient for the efficient reconstitution of downstream phytosteroid pathways, the endogenous mevalonate (MVA) pathway was upregulated in yeast to increase the synthesis of squalene, the precursor of sterols (Figure 1.1). Based on a previously established pathway[31]: *erg7*, *erg10*, *erg12*, *erg13*, *erg19*, *erg20* and *idi1* were overexpressed by strong, constitutive promoters in the yeast genome. Earlier reports also indicate that overexpressing the truncated 3-hydroxy-3-methylglutaryl-coenzyme A reductase (tHMG1) and the sterol regulatory element binding protein UPC2 could increase the production of squalene[31, 32]. Thus, three squalene-upregulating modules (module I: *thmg1*, *erg10*, *erg12*, and *erg13*; module II: *erg8*, *erg19*, *idi1*, and *thmg1*; module III: *erg7*, *erg20*, *upc2*, and *thmg1*) were integrated sequentially into *ybl059w*, *ybr197c*, and *ymr206w* locus of YYL55 to afford YYL56, YYL57, and YYL58, respectively (Figure 1.1, Table 1.2). Consistent with previous studies and as expected, the overexpression of genes on MVA pathway improved campesterol production. Out of all the strains, YYL58 exhibited the highest production of campesterol at 40.72 mg/L, a 13-fold increase from the original campesterol-producing YYL55. However, upregulating squalene synthesis also led to a deleterious effect on the growth of yeast: compared with wild-type strain, YYL56, YYL57 and YYL58 all exhibited slightly longer lag phases and slower growth rates (Figure 1.2B).

Construction and Characterization of β-Sitosterol Biosynthesis in Yeast

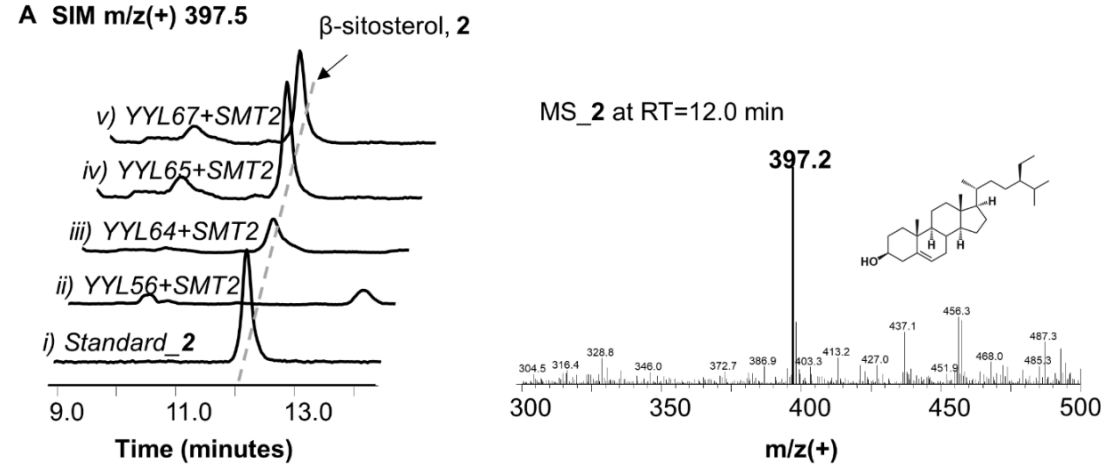
Episterol, **4** is an important precursor of ergosterol and campesterol in fungi and plants, respectively. In plants, **4** is proposed to be synthesized from 24-methylenelophenol, **9**, by the function of the ERG25 ortholog sterol 4α-methyl oxidase SMO2[33]. **9** is proposed

to be converted to 24-ethylidenelophenol, **10**, with the function of the 24-methylenesterol C-methyltransferase SMT2[34], and SMO2, DWF7, DWF5, and DWF1 then function to convert **10** to β-sitosterol, **2** (Figure 1.1). Due to the structural similarity between **9** and **4**, and the fact that sterol pathway in plants is highly cross-linked, it is likely that SMT2 could function on 24-methylenesterols, such as **4**, **5** or **6**. Therefore, SMT2 was cloned from *A. thaliana* cDNA library and expressed under a strong constitutive *GPD* promoter in YYL58. However, no β-sitosterol or any 24-methylated products were detected (Figure 1.3). LC-MS analysis of the SMT2-expressing YYL58 indicated that **6** was accumulated in saponified extracts of the cell pellets, while **4** and **5** were barely detected, which is likely due to the efficient conversion of **4** and **5** to Ergosta-5,7,22,24-tertraene-3betaol, **8**, catalyzed by ERG3 and ERG5 enzymes. Normally, surplus or foreign sterols are esterified and stored in lipid droplets in yeast to reduce the toxicity[35]. This mechanism of regulating steroid synthesis by sterol esterification is also found in plants. Previous investigations found that changing the expression level of a putative acyl transferase BAT1 (BR-related acyltransferase 1) may influence BR levels in plants: overexpression of BAT1 in *A. thaliana* leads to BR-deficient phenotypes, which can be rescued by BR feeding[36]. Thus, we hypothesize that **6** may be esterified in yeast and became inaccessible to SMT2. To confirm this possibility, YYL58 cell pellets were extracted without saponification. Interestingly, compared with the saponified samples, **6** and campesterol were barely detected in the unsaponified extracts of YYL58, indicating highly efficient esterification of these two phytosterols in yeast (Figure S1.2). Complete esterification of campesterol and **6** was found in all the other campesterol producing strains as well.

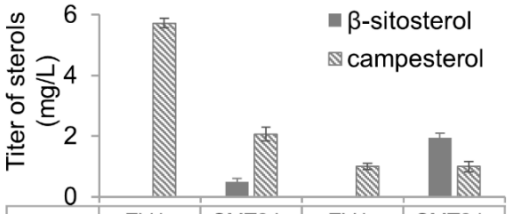
To enhance the availability of free phytosterols in yeast, two acyltransferases responsible for sterol esterification were selected to be deleted from YYL58. Previous investigations indicate that two acyl-CoA sterol acyltransferases ARE1 and ARE2 are responsible for the esterification of sterols in *S. cerevisiae*[37, 38]. Inactivation of *are1* and *are2* demolished the synthesis of sterol ester and enhanced free sterols levels[37, 38]. However, inactivation of *are1* and *are2* was lethal to YYL58, likely due to the accumulation of free phytosterols, isopentenyl pyrophosphate (IPP), or dimethylallyl pyrophosphate (DMAPP)[39]. YYL56, the campesterol-producing strain exhibited more robust growth than YYL58 (Figure 1.2B) but with a similar level of campesterol production, was selected for *are1* and *are2* inactivation to generate YYL64. YYL64 synthesized campesterol and **6** as free sterols, with a decreased campesterol titer at 5.72 mg/L (Figure 1.2A, 2C).

Introducing SMT2 to YYL64 resulted in β-sitosterol production although at only ~0.5 mg/L and with a cost of campesterol production dropped around 60%, which was synthesized by the same set of enzymes except for SMT2 (Figure 1.1, 1.3A, 1.3C). The reduced production of campesterol in YYL64 expressing SMT2 indicates that the synthesis of β-sitosterol likely consumed compound **6** due to the promiscuity of SMT2. In SMT2-expressing YYL64, isofucosterol, **7** was detected while the methylated products of **4** and **5** were not, which indicates that **6** is likely a substrate of SMT2 in YYL64 (Figure 1.1, 1.3B, 1.3D). The inactivation of *are1* and *are2* in YYL64 also led to growth defects (Figure 1.2B), altered the sterol composition, and diminished total sterol production in YYL64 (Figure S1.1). Since the inactivation of *are1* and *are2* in wild type strain (YYL63, Table 1.2) had a minimal effect on yeast viability (Figure 1.2B), the growth defect in YYL56 is likely due

to the accumulation of foreign sterols or squalene. The weakened growth of YYL64 in comparison with YYL56 also implies the importance of esterification in reducing the toxicity of free foreign sterols in yeast.



C



Strain	EV in YYL64	SMT2 in YYL64	EV in YYL65	SMT2 in YYL65
OD_{600}	10.28 ±0.11	10.18 ±0.39	5.97 ±0.06	6.35 ±0.17

D

sterol	4	5	6	4'	12	7	2
YYL64	-	-	-	-	-	+	+
YYL65	+	-	-	-	-	+	+
YYL60	-	+	-	-	+	+	-

*4' is C24'-methylated 4.

Figure 1.3 Biosynthesis of β-sitosterol in yeast. **(A)** Selected ion monitoring SIM (SIM) extracted ion chromatogram (EIC) using β-sitosterol's characteristic m/z^+ signal (MW=414.71Da, $[C_{29}H_{51}O\text{-}H_2O]^+=[C_{29}H_{49}]^+$=397.4) of i) β-sitosterol standard, ii) YYL56 harboring SMT2, iii) YYL64 (YYL56 with *are1*Δ*are2*Δ) harboring SMT2, iv) YYL65 (YYL56 with *are1*Δ*are2*Δ*erg4*Δ) harboring SMT2, v) YYL67 (adaptive evolution strain from YYL65) harboring SMT2. **(B)** EIC SIM using isofucosterol, **7**, characteristic m/z^+ signal (MW=412.70Da, $[C_{29}H_{49}O\text{-}H_2O]^+=[C_{29}H_{47}]^+$ =395.3) of i) isofucosterol standard, ii) YYL64 expressing SMT2, iii) YYL65 expressing SMT2, iv) YYL67 expressing SMT2. **(C)** Campesterol and β-sitosterol production in YYL64 and YYL65 expressing SMT2. "EV" represents "empty vector" as negative control. **(D)** Sterol composition of YYL64, YYL65 and YYL60 (*erg5*Δ*::dhcr7*) in the presence of SMT2. YYL64 and YYL65 were extracted without saponification; YYL60 was extracted with saponification method. All traces are representative of at least three biological replicates for each engineered yeast strain. Bars represent mean values of three biological replicates, and the error bars represent the standard deviation of the replicates.

Enhancing the Production of β-Sitosterol by Inactivating *erg4*

ERG4 and DWF1 are both Δ^{24}-sterol reductases that function on the C24 double carbon bond. When functioning on **6**, DWF1 catalyzes the formation of both 24(*S)*- and 24(*R)*-campesterol; while only 24(*R*)-β-sitosterol is synthesized in plants[24, 40]. On the other side, ERG4 has only been reported to synthesize 24(*S*)-methyl sterol, ergosterol, but not 24-ethyl sterols such as β-sitosterol[41], which indicates that ERG4 is likely a 24(*S*)-reductase functioning on sterols with 24-methylene group but not with 24-ethylene group. Consistent with previous results that campesterol could be synthesized from DHCR7-expressing yeast[22, 23], expressing DWF5 in wild type yeast led to the synthesis of campesterol (Figure S1.3); while no campesterol was detected when DWF5 was introduced to the *erg4*Δ mutant strain(Figure S1.3). Moreover, the expression of SMT2 with DWF7 and DWF5 in YYL63 (*are1*Δ*are2*Δ) resulted in no detection of β-sitosterol (Figure S1.4). These facts imply that while DWF1 can function on both **6** and **7**, ERG4 may only function on **6**. Thus, ERG4 likely competes with SMT2 in the conversion of **6** in SMT2-expressing YYL64. To improve the synthesis of β-sitosterol, *erg4* was then deleted from YYL64 to generate YYL65. Inactivation of *erg4* resulted in growth deficiency, such as the hypersensitivity to certain drugs and impaired mating process[41]. Here, inactivation of *erg4* resulted in dramatically reduced campesterol synthesis (~1 mg/L) (Figure 1.2) and a significant growth burden on YYL65, which exhibits an extended lag phase and largely decreased cell density at the stationary phase (Figure 1.4, Figure 1.2B). However, the expression of SMT2 in YYL65 resulted in an enhanced production of β-sitosterol at 2 mg/L, an approximately 4-fold increase compared to YYL64 (Figure 1.3A, 1.3C). Different from what we observed

in YYL64, introducing SMT2 to YYL65 did not affect campesterol production (Figure 1.3B); although **6** was not detected, **7** was observed in the SMT2-expressing YYL65 (Figure 1.3D). The decreased cell density with a higher β-sitosterol production, compared to YYL64, agreed with the possibility that ERG4 may compete with SMT2 in β-sitosterol synthesis, and **6** is a major substrate of SMT2 in YYL64 and YYL65.

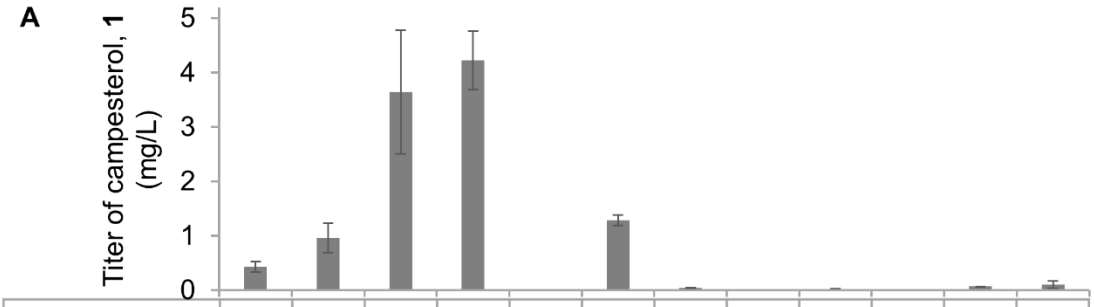
In addition, to identify the possible substrates of SMT2 in addition to **6** in yeast, the methylated products of **4** and **5** were examined in **4** and **5**-accumulating strains. Methylated episterol, **4'**, was not detected in YYL65, which accumulated **4**. It suggested that **4** cannot be converted by SMT2 in yeast (Figure 1.3D). We also introduced SMT2 to the **5**-accumulating strain YYL60 (Table 1.2), and observed the synthesis of a compound with m/z^+=393.4 that agrees with **12** ($[C_{29}H_{47}O\text{-}H_2O]^+=[C_{29}H_{45}]^+$=393.4, Figure S1.5), indicating that SMT2 can also function on **5**. Different from **6** and campesterol, we were able to detect free **5** in YYL60 (Figure S1.2C). Taken together, these results suggest that **5** and **6** maybe the substrates of SMT2 in YYL64 and YYL65 that route the synthesis towards β-sitosterol upon the introduction of SMT2 (Figure 1.1).

Optimization of Phytosterol Production in Yeast

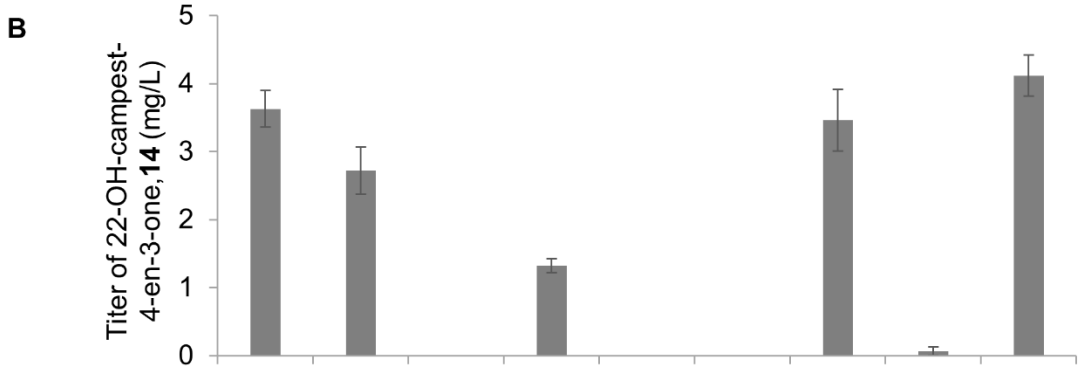
The successful construction of β-sitosterol production enabled us to identify the obstacles in metabolic engineering of phytosterol synthesis in yeast. The cross talk between the endogenous sterol metabolism and the heterologous sterol pathway often make the substrates not readily accessible for the downstream heterologous enzymes. Thus, disruption of sterol metabolism may be necessary for the successful reconstitution of

phytosterol biosynthesis, but this also likely introduce growth deficiencies, such as in the case of YYL65. Due to the growth deficiency introduced by the *are1/are2* and *erg4* inactivation and the relatively low titer of campesterol, functional reconstitution of more heterologous enzymes in YYL65 became unlikely.

To improve phytosterol production and alleviate the growth stress, we optimized the medium for YYL65. Different carbon sources were used to improve the growth condition and enhance campesterol production (Figure 1.4A). We also examined the effects of methyl-β-cyclodextrin considering that it may alleviate the sterol stress in YYL65 by transporting deleterious sterols out of yeast[42]. The effect of supplementing 2% (w/v) sucrose with different amounts of ethanol (2%, 5%, and 10%, v/v) was examined as well (Figure S1.6). Interestingly, supplementing 10% (v/v) ethanol to SDM with 2% (w/v) sucrose or glucose led to up to 10-fold increase in campesterol production but also extended the lag phase of yeast growth. The enhanced production of campesterol is consistent with previous reports of yeast production of amorphadiene, which is likely due to the fact that ethanol enhanced the supply of acetyl-CoA[32]. In addition, a high concentration of ethanol was reported to trigger an increased ergosterol synthesis in order to protect damaged plasma membrane[43]. Thus, the addition of ethanol may upregulate sterol synthesis and result in enhanced supply of **4** and **5**, substrates of campesterol. However, despite the elevated campesterol titer in the presence of ethanol, we did not observe obvious alleviation of growth stress in the medium conditions tested.



YP/SDM	SDM	SDM	SDM	SDM	SDM	SDM	SDM	SDM	SDM	SDM	YP	YP
carbon source (2%)	dex	suc	dex	suc	suc	suc	gal	treh	mal	gly	dex	suc
10% ethanol			+	+		+						
5mM cyclodextrin					+	+						
OD_{600}	7.54 ±1.08	6.01 ±0.45	4.18 ±1.46	5.96 ±0.98	-	3.00 ±1.93	3.15 ±0.31	-	1.51 ±0.32	-	8.99 ±0.28	12.58 ±0.37



YP/SDM	SDM	SDM	SDM	YP	YP	YP	YP	YP	SDM
carbon source	2%dex	2%suc	2%tre	2%dex	5%dex	2%suc	2%dex +2%gly	2%tre	2%dex +10% ethanol
OD_{600}	13.47 ±0.58	8.49 ±0.70	-	15.14 ±1.25	-	-	15.90 ±1.59	21.10 ±0.95	11.48± 0.54

Figure 1.4. Medium optimization of **(A)** YYL65 and **(B)** YYL69. SDM represents as synthetic defined minimal medium with 1.7 g/L yeast nitrogen bases, corresponding dropout and 5 g/L ammonium sulfate, pH 5.8. YP represents complex medium with 10 g/L yeast extract, 20 g/L peptone and 80 mg/L adenine. 2% (w/v) of various carbon sources were used: dex (D-glucose), suc (sucrose), gal (galactose), treh (trehalose), mal (malose) and gly (glycerol). Bars represent mean values of three biological replicates, and the error bars represent the standard deviation of the replicates.

Meanwhile, we noticed that re-streaking YYL65 on agar plates resulted in two different morphologies of yeast colonies – some exhibited a much larger size than the remaining majority, indicating that some yeast cells may have mutated to circumvent the

growth burden caused by the altered sterol composition. Based on this interesting observation, we conducted adaptive evolution on YYL65 to screen for derivatives with enhanced growth. For inoculum preparation, YYL65 was cultivated at 250 rpm and 30°C in 2mL SDM till it reached the stationary phase (OD_{600} ~6.0). An appropriate volume of this culture was used to inoculate 2mL fresh SDM with an initial OD_{600} of 0.1. Once the culture reached early stationary phase, serial dilutions were repeated. In this procedure, the time interval that YYL65 took to reach the stationary phase became shorter after each cycle, indicating enhanced growth. The culture after 10 cycles were plated on agar plate and 10 individual colonies were further analyzed by culturing in SDM with and without 10% (v/v) ethanol. Campesterol titer was then quantified by LC-MS. Excitingly, all 10 colonies grew more robustly than the original YYL65 and the production of campesterol was all boosted. Seven colonies synthesized a higher amount of campesterol in the presence of 10% (v/v) ethanol and one colony didn't survive in the high concentration of ethanol (Figure 1.5A). Among them, YYL65-5 had the highest production of campesterol in the presence of ethanol, while YYL65-1 and YYL65-2 exhibited comparable level of campesterol titer without the presence of ethanol (Figure 1.5A). YYL65-1, with both enhanced growth and higher campesterol production in normal medium condition (campesterol titer at ~ 7.13 mg/L, Figure 1.2B, Figure 1.5B), was annotated as YYL67 and selected for the further pathway reconstitution.

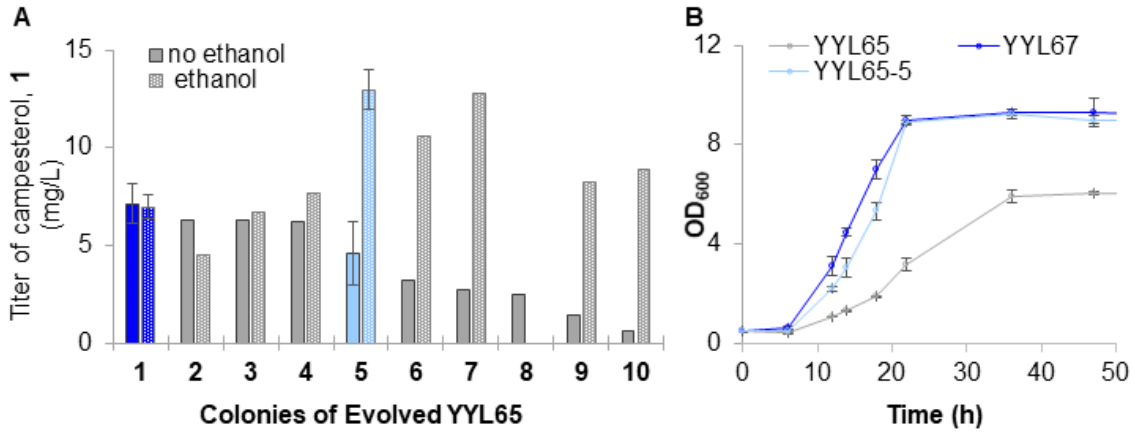


Figure 1.5 Adaptive evolution of YYL65. **(A)** Campesterol production of YYL65 and the ten mutants picked from the 10th generation of adaptive evolution in synthetic complete SDM supplemented with 2% (w/v) sucrose, with and without 10% (v/v) ethanol. **(B)** Growth curve of YYL65, YYL65-7, and YYL67. The yeast strains were cultured in synthetic complete SDM medium at 30°C for 72 hours.

Synthesis of 22-Hydroxycampest-4-en-3-one in Yeast

With the phytosterol-producing strain engineered to overcome the growth deficiency and produce phytosterols accessible for downstream enzymes, YYL67 can be utilized as a phytosterol-producing platform to reconstitute and elucidate the downstream biosynthetic pathway. Since the production of YYL67 in glucose and sucrose was similar (Figure S1.6), we switched to normal SDM supplemented with 2% (w/v) glucose for the subsequent YYL67-related engineering and culturing. To explore this possibility, we attempted to produce BRs using YYL67. BRs are a group of universal hormones derived from campesterol and play essential roles in higher plants[14, 15]. They are involved in plant growth regulation, developmental processes, and responses to abiotic or biotic stress[14, 15, 46, 47]. For example, exogenous application of BR on potatoes have led to enhanced productivity and starch content[14]. The ability to boost crop yield makes BRs important bio-stimulants in agriculture. In addition, BRs also show promising antiviral, anticancer, and

antitumor activities[20]. However, the low abundance of BRs in nature and the structural complexity make these molecules very difficult and expensive to obtain through direction isolation or chemical synthesis, and the high price and low accessibility of natural BRs of high purity and the lack of pathway intermediates are largely impeding the investigation and development of BRs for agricultural and pharmaceutical applications. Reconstituting BR pathways in microbial organism is a promising alternative approach to produce various BRs in high purity and economically. However, the biosynthetic pathways of BRs are not fully elucidated, with at least two enzymes missing and the sequence of the biosynthetic enzymes elusive. The construction of YYL67 enabled us to establish BR biosynthetic pathways in yeast, which may serve as a suitable platform for future BR production. CYP90B1 was demonstrated as the first and rate-limiting step in BR biosynthesis[48]. Expression of CYP90B1 under the regulation of *GPD* promoter in YYL67, which contains NADPH–cytochrome P450 oxidoreductase ATR1, led to an efficient conversion of campesterol to (22*S*)-22-hydroxy-campesterol, **13** (m/z^+ = 399.3, Figure 1.6A, 1.6B), which is consistent with the previous *in vitro* biochemical characterization of CYP90B1[48]. The LC-MS/MS analysis further confirmed that the hydroxyl group was add at the C22 position (Figure 1.6B, S1.7). We also introduced CYP90B1 to YYL58 and YYL64, but the synthesis of **13** was either undetected or inconclusive (Figure 1.6A). We were unable to express CYP90B1 in YYL65 due to significant growth deficiency. The failed functional reconstitution of CYP90B1 in YYL58, YYL64, and YYL65 again highlights the importance of inactivating sterol acyltransferase genes and *erg4* from YYL56, and the

directed evolution of YYL65 for the efficient construction of plant metabolite biosynthesis downstream of simple phytosterols (e.g. **1**, **2**, **6**) in *S. cerevisiae*.

The conversion efficiency of campesterol to **13** is around 86% (Figure S1.8). In addition, the activities of an array of C22-sterol-hydroxylases were tested in YYL67, including CYP90B1 variants and cholesterol C22-hydroxylases (Table S1.4). Among all the enzymes, only the analogues of CYP90B1 in *Solanum lycopersicum*, *sl*CYP724B2 and *sl*CYP90B3, can function on campesterol in yeast (Figure S1.8), as reported before[49]. The membrane bound signaling peptide (~30 amino acids) at the N-terminus was truncated to afford t30CYP90B1, which can convert campesterol to **13**, but with a lower efficiency compared with the full-length protein (Figure S1.8). This agrees with the previous investigation that the activity of plant cytochrome P450s is retained upon the N-terminus truncation, which can be utilized as an expression strategy in the membrane infrastructure-lacking *E. coli*[50]. Compared with *sl*CYP724B2, *sl*CYP90B3 and t30CYP90B1, CYP90B1 exhibits the highest efficiency towards the synthesis of **13** (Figure S1.8).

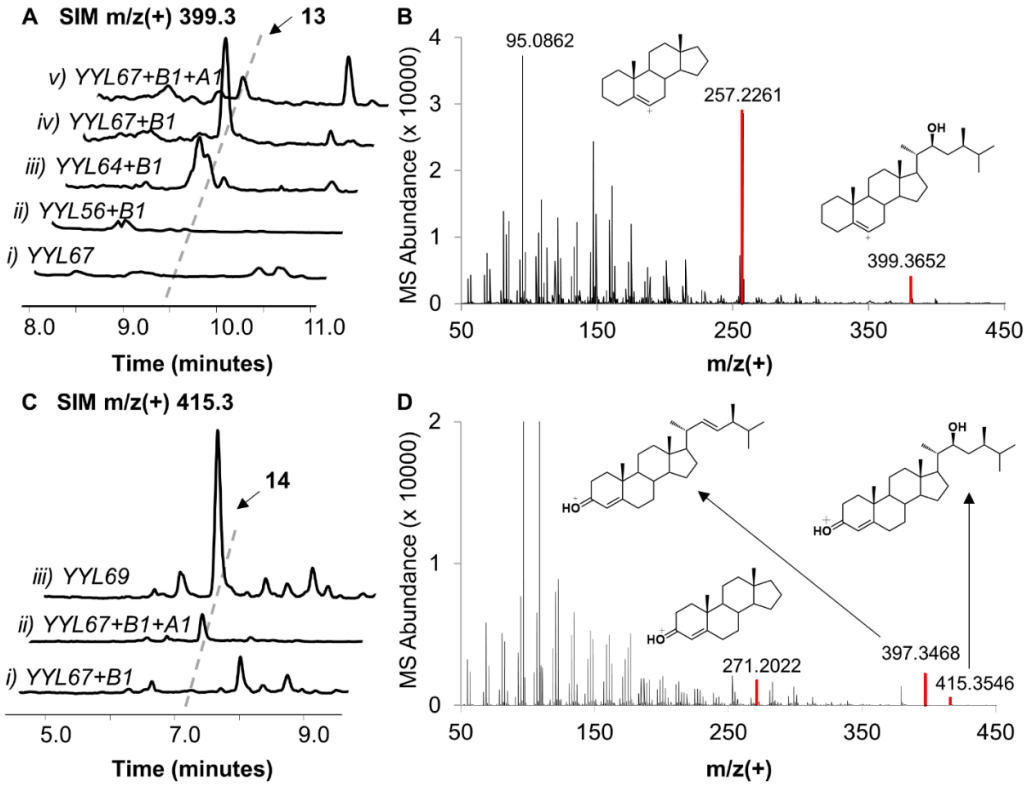
CYP90A1 was first proposed as a C23-hydroxylase[51], yet subsequent *in vitro* characterization argues that CYP90A1 is more likely a C3-oxidase on BR intermediates with 22-hydroxylated and 22,23-dihydroxylated side chains[52]. Expression of CYP90A1 in YYL67 didn't lead to the synthesis of either 3-oxidized or 23-hydroxylated product, which agrees with the previous study that campesterol is not a substrate of CYP90A1[52]. Co-expression of CYP90A1 with CYP90B1 in YYL67 led to the synthesis of a new peak of m/z^+=415.3, which agrees with the putative product 22-hydroxycampest-4-en-3-one, **14**.

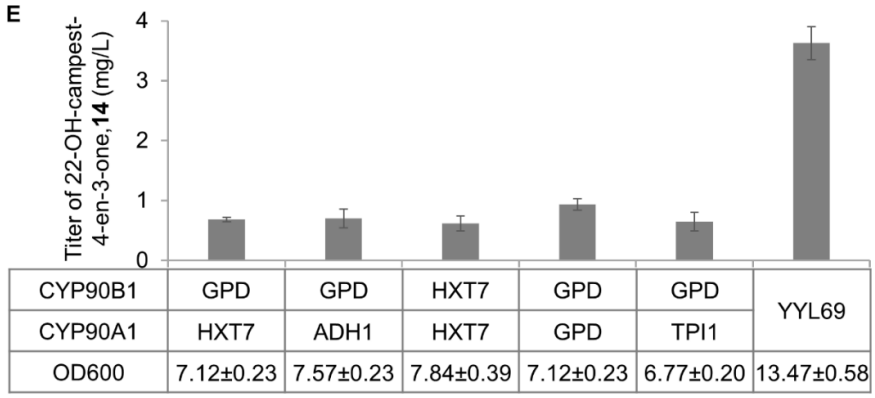
The LC-MS/MS analysis further confirmed the desaturation of the hydroxyl group on the A-ring (Figure 1.6D). This functional reconstitution of CYP90A1 and CYP90B1 confirmed the reaction order between these two enzymes. The conversion of **13** to **14** was very efficient, with very little of **13** detected in YYL67 expressing CYP90A1 and CYP90B1 (Figure 1.6A, 1.6C), which implies that CYP90B1 may be the rate-limiting step and could be a key to regulate BR biosynthesis in yeast.

We also examined the effect of different promoters on the activities of CYP90A1 and CYP90B1 towards the synthesis of **14**. When CYP90A1 and CYP90B1 are both regulated by strong, constitutive promoters, a slightly higher production of **14** was detected (Figure 1.6E). We then integrated *CYP90A1* and *CYP90B1* into YYL67 to generate YYL69 (Table 1.1). Remarkably, the production of **14** was substantially enhanced in comparison to when these two genes were expressed from low-copy number plasmids (Table S1.5), from 0.94 mg/L to 3.63 mg/L, together with an enhanced growth (Figure 1.6F). This result is surprising and, simultaneously highlights the significance of stable genome integration in pathway reconstruction in *S. cerevisiae*. The establishment of YYL69, the **14**-producing strain, confirms the reaction sequence of CYP90A1 and CYP90B1 in yeast, which is consistent with previous *in planta* and *in vitro* biochemical characterizations. Our work provides a yeast-based platform for the elucidation and engineering of downstream BR biosynthesis.

To further optimize the production of **14** in YYL69, INO2 was overexpressed under the strong, constitutive *GPD* promoter. INO2 was recently found to be able to enlarge the

endoplasmic reticulum membrane, where the steroids were synthesized[53]. However, although overexpressing INO2 was very successful in enhancing squalene production, it did not help in the context of sterol production; instead, the titer of **14** dropped ~40% upon the overexpression of INO2 (Figure S1.19A). Since changing the medium conditions has made substantial differences to the production of campesterol in YYL65, we again explored the effects of various base media and carbon sources on **14** production from YYL69 (Figure 1.4B). Specifically, YPD resulted in better growth but less production of **14**, in comparison to when cultured in SDM. Across all of the carbon sources examined, glucose exhibited the most promising production of **14**. Further optimization was done through titration between YP and SDM supplemented with 2% (w/v) glucose to locate the optimal balance between **14**-production and cell growth. However, the titer of **14** is proportional to the ratio of SDM (Figure S1.19B), and the highest production of **14** was still found in SDM with 2% (w/v) glucose. In addition, supplementing 10% (v/v) ethanol in SDM with 2% (w/v) glucose showed ~13% increase in the titer of **14**, with a similar cell density as without ethanol.





CYP90B1	GPD	GPD	HXT7	GPD	GPD	YYL69
CYP90A1	HXT7	ADH1	HXT7	GPD	TPI1	
OD600	7.12±0.23	7.57±0.23	7.84±0.39	7.12±0.23	6.77±0.20	13.47±0.58

Figure 1.6 Biosynthesis of (22S)-22-Hydroxycampest-4-en-3-one, **14**. **(A)** SIM EIC using (22S)-22-Hydroxy-campesterol, **13**'s characteristic m/z^+ signal (MW=416.69Da, $[C_{28}H_{49}O_2\text{-}H_2O]^+$ $[C_{28}H_{47}O]^+$=399.3) of i) YYL67 (adaptive evolution strain from YYL56 with *are1*Δ*are2*Δ*erg4*Δ), ii) YYL56 expressing CYP90B1, iii) YYL64 (YYL56 with *are1*Δ*are2*Δ) expressing CYP90B1, iv) YYL67 expressing CYP90B1, v) YYL67 expressing CYP90B1 and CYP90A1. **(B)** High resolution mass spectrum of **13**. **(C)** SIM EIC using (22S)-22-Hydroxycampest-4-en-3-one, **14**'s characteristic m/z^+ signal (MW=414.67Da, $[C_{28}H_{47}O_2]^+$=415.3) of, i) YYL67 expressing CYP90B1, ii) YYL67 expressing CYP90B1 and CYP90A1, iii) YYL69. **(D)** High resolution mass spectrum of **14**. (E) Quantification of **14** with different promoter combination of CYP90B1 and CYP90A1 expressed in YYL67. All traces and spectrums are representative of at least three biological replicates for each engineered yeast strain and the error bars represent the standard deviation of the replicates.

Discussion

Although campesterol production has been previously established in *S. cerevisiae* and *Y. lipolytica* by the function of an animal DHCR7 and the promiscuity of ERG4[22, 23], such campesterol-producing strain has not been utilized for the synthesis of downstream metabolites, such as BR. In this study, we first established campesterol production using the plant enzymes DWF7, DWF5, and DWF1. Upregulating the mevalonate pathway, a classic strategy to enhance upstream terpene supply, led to more than 10-fold enhancement in campesterol titer to around 40 mg/L. However, such high titer of campesterol and its precursors cannot be converted to **13** and the corresponding 24-methylated sterols upon the introduction of CYP90B1 and SMT2, respectively. Analysis of the campesterol-producing strain YYL56 implies that heterologous sterols in this strain (**6**, campesterol) are only detected in the saponified sample, which indicates the vigorous esterification of these plant specified sterols in yeast (Figure 1.2A). The inactivation of the sterol acyltransferase genes *are1* and *are2* resulted in the synthesis of **7** and **13**, but in low conversions from their precursors (Figure 1.1, Figure 1.3A, 1.6A). Further inactivation of *erg4*, the paralog of *dwf1*, greatly enhanced the conversion towards the 24'-methylated products such as **7** and β-sitosterol by SMT2, but with a significant growth deficiency. Interestingly, the engineered yeast strains were able to circumvent the growth deficiency through adaptive evolution, with enhanced phytosterol production as well. Although the campesterol titer of the best mutant was only ~7 mg/L, less than 1/4 of the highest campesterol titer achieved in this study, the synthesized campesterol is much more accessible by the downstream enzymes, and the C24 stereochemistry is not affected by the presence of ERG4 enzyme.

Using this strain, we were able to reconstitute and optimize the early steps of BR biosynthesis, from campesterol to **14**, in yeast with titer of **14** up to ~4 mg/L.

Although ERG4 and DWF1 can both function on **6** to afford the synthesis of campesterol, the stereochemistry at C24 can be very different. It seems that introduction of DHCR7 only likely leads to predominant 24(*S*)-campesterol synthesis. Previous investigations implied that DWF1 from different plants exhibit different stereochemistry, some exclusively catalyze the formation of 24(*R*)-campesterol, while others result in racemic mixtures[24]. In this study, we used DWF1 from *A. thaliana*. According to previous investigations of *At*DWF1[24] and the high efficiency of ERG4 on **6**, YYL67 likely synthesizes a racemic mixture, while YYL64 synthesizes a mixture with a higher ratio of 24(*S*)-campesterol. Thus, it is critical to use plant enzymes (DWF1, DWF5, DWF7) to synthesize plant specialized campesterol in yeast, if further downstream pathway is to be constructed.

Through engineering for altered sterol composition, yeast strains were more or less stressed in this study. However, we were surprised to learn the superior stress resistance of *S. cerevisiae*. Apart from the capability to adapt themselves for both enhanced growth and sterol production through simple repeated culture, we also noted that yeast can alter their genetic complement to release the stress. For example, when we re-streaked YYL65 on YPD plate, two phenotypes were detected: there was one colony which grew much faster than remaining majority (Figure S1.10), which could not produce campesterol and is marked as YYL65-BC. Colony PCR was reperformed on this strain and the previously

integrated *dwf5* was no longer detected. In addition, the campesterol production can be rescued by expressing DWF5 from a plasmid in YYL65-BC (Figure S1.10), which further confirms the loss of function of *dwf5* but not the other integrated genes. On the other hand, we noticed that YYL65 harboring SMT2 showed slightly better growth than YYL65 alone. The enhanced growth in the *dwf5* missing YYL65-BC and SMT2-expressing YYL65 may be attributed to the loss or a decreased level of 24-methylenecholesterol, respectively. We thus hypothesized that 24-methylenecholesterol is toxic to yeast and maybe the major cause of the growth deficiency of YYL67. Additionally, in this study, the ergosterol synthesis was much affected in all the engineered strains, which also implies the flexibility of *S. cerevisiae* to altered sterol composition. Surprisingly, expressing SMT2 in YYL67 led to a slightly decreased titers of β-sitosterol and **7**, compared to YYL65 (Figure 1.3A). The possible mechanism for the unexpected lower titer of β-sitosterol in YYL67 is not clear, but implies the complexity of sterol metabolism regulation in *S. cerevisiae*.

Similar to many previously reported metabolic engineering efforts, different medium conditions can cause distinct effects on the target compound production. YYL65 exhibits a significant growth deficiency (Figure 1.5B) that cannot be used to express cytochrome P450s, and was not able to grow in a number of medium conditions. YYL65 grows slightly better in glucose, yet exhibits a higher campesterol titer in sucrose. This agrees with previous investigations that ergosterol production tends to be higher when using dihexoses (e.g. sucrose and maltose) as carbon sources in yeast, compared to monosaccharides (e.g. glucose)[54]. However, such a trend was diminished in the adapted strain YYL67 (Figure S1.6B). Upon introducing plant cytochrome P450s, however,

glucose became a more favorable carbon source in respect of both sterol synthesis and growth (Figure 1.4B). Unlike previous studies that addition of glycerol or trehalose helped the noscapine production in *S. cereivsiae*[55], addition of glycerol did not help the production of **14** and YYL69 cannot grow in the presence of trehalose (Figure 1.4B). Addition of ethanol did enhance the titer of **14** but the effect was subtle, which might be due to the positive effect of ethanol on acetyl-coA supply and sterol biosynthesis as discussed previously. The distinct preference in carbon sources between YYL65, YYL67, and YYL69 may shed light on potential engineering strategies for phytosterol-production and plant cytochrome P450 reconstitution in yeast.

Through reconstituting the biosynthesis of β-sitosterol and **14**, we demonstrated the critical bottlenecks in converting phytosterol to downstream products in yeast – crosstalk between heterologous and yeast endogenous sterol metabolism, and the growth burden caused by altered sterol composition in yeast. We believe that these challenges are generally true in the reconstitutions of other phytosterol-derived pathways in yeast. Thus, this work has broad implications on engineering strategies to establish phytosteroid-producing yeast strains in general. In addition, this work highlights the potential of using baker's yeast to elucidate and engineer the biosynthesis of various steroids, such as steroidal alkaloids and withanolides. The establishment of a **14**-producing strain will provide a platform to elucidate the biosynthesis of BRs. Thus, *S. cerevisiae* is an optimal host for phytosterol pathway elucidation and reconstitution. In addition, considering the growth deficiency in YYL67 and YYL69, non-conventional yeast strains with more robust growth or higher lipid production, such as *Y. lipolytica*, should also be considered for the

manufacturing of phytosteroids once the biosynthesis is fully elucidated and reconstituted in *S. cerevisiae*.

Materials and Methods

Materials and Culture Conditions

Chemicals: campesterol (~98%) and ergosterol (⩾75%) are obtained from Sigma-Aldrich. β-Sitosterol (~65%) and stigmasterol (~95%) are purchased from Fisher Scientific. All engineered yeast strains in this work are listed in Table 1.1 and constructed in a haploid CENPK2.1D background (*MATα; his3D;1 leu2-3_112; ura3-52; trp1-289; MAL2-8c; SUC2*). Yeast strains were cultured at 30 ^{o}C in complex yeast extract peptone dextrose (YPD, all components from BD Diagnostics) medium or synthetic defined medium (SDM) containing yeast nitrogen base (YNB) (BD Diagnostics), ammonium sulfate (Fisher Scientific), 2% (w/v) glucose unless specified and the appropriate dropout (Takara Bio) solution for selection. 200 mg/L G418 sulfate (Calbiochem) or 200 mg/L Hygromycin B (Life Technologies) were used in YPD medium for selection.

General Technique for DNA Manipulation

PCR reactions were performed with Expand high Fidelity system (Sigma-Aldrich), Phusion DNA Polymerase (NEB), Q5 High-Fidelity DNA Polymerase (New England Biolabs) and Taq Polymerase (NEB) according to manufacturer's protocols. PCR products were purified by Zymoclean Gel DNA Recovery Kit (Zymo Research). Plasmids were prepared with Econospin columns (Epoch Life Science) according to manufacturer's protocols. All DNA constructs were confirmed through DNA sequencing by Source

Bioscience Inc and Retrogen sequencing Inc. BP Clonase II Enzyme Mix, Gateway pDONR221 Vector and LR Clonase II Enzyme Mix (Life Technologies) and the *S. cerevisiae* Advanced Gateway Destination Vector Kit[56] (Addgene) were used to perform Gateway Cloning. Gibson one-pot, isothermal DNA assembly[57] was conducted at the scale of 10 μL by incubating T5 exonuclease (NEB), Phusion polymerase (NEB), Taq ligase (NEB) and 50 ng of each DNA fragment at 50 ^{o}C for 1 h to assemble multiple DNA fragments. Yeast strains are constructed through homologous recombination and DNA assembly[58]. *ATR1*, *DWF1*, *DWF5*, *DWF7*, *SMT2* were amplified from *A. thaliana* cDNA (kindly provided by Prof. Patricia Springer). *DHCR7* from Xenopus laevis, *CYP90A1*, *CYP90B1*, *CYP90D1* and *DET2* from *A. tahliana* were *S.cerevisiae* codon optimized and synthesized from TWIST Bioscience Inc.

Plasmids used in this study are in Table S1.4. Sequences of genes used in this work are listed in Table S1.5.

Culture and Fermentation Conditions

For all the functional assays reported in this work,, yeast strains were grown in 500 μL media in 96-well plates (BD falcon) covered with AeraSeal film (Excel Scientific), shaking at 250 r.p.m.. For all the engineered yeast metabolites analysis, yeast strains were first cultured overnight in 500 μL SDM with 2% (w/v) glucose. Appropriate volume of the overnight seed culture was inoculated in fresh 500 μL SDM with carbon sources to make the OD_{600} around 0.1 and incubated at 30 ^{o}C for 72 h before metabolite analysis of the yeast pellets.

For growth monitoring, yeast strains were grown in 2 mL SDM with 2% (w/v) sucrose in test tubes. The yeast strains were first cultured overnight in 500 µL SDM with 2% (w/v) glucose. Appropriate volume of the overnight seed culture was inoculated in fresh 2 mL SDM with 2% (w/v) glucose to make the OD_{600} around 0.1 incubated at 30 ^{o}C for 72 h, at 350 r.p.m..

Analysis of Phytosterol and Intermediates in BR Pathway

All yeast metabolites of interests in this study were all collected from yeast pellets. For the strains producing steryl esters, 100 µL 60% (w/v) KOH and 100 µL ethanol was added to each sample, then incubated at 86 °C for 1 to 2 hours for saponification. The mixture was cooled down to room temperature and then 400 µL petroleum ether was added and vortexed thoroughly for 15 min. Organic phases were removed to new tubes and dried in vacufuge (Eppendorf). For Δ*are1*Δ*are2* double mutants, 100 µL of dimethylformamide (DMF) was added to each sample. Then 400 µL of petroleum ether was add and vortex to mix. The petroleum ether phase was taken to new tubes and dried in the vacufuge. For the intermediates in BR synthetic pathway, yeast pellets from 1 ml culture were collected and mixed with 100 µL of DMF and 900 µL acetone. The mixture was vortexed for 15 min and spun at 13,3000 rpm for 10 min. The liquid phases were taken to new tubes and dried in the vacufuge. All dried samples were eluted in ethanol and run through LC-MS.

LC-MS used here is Shimadzu serial 2020. The column used in analysis is from Agilent HPLC column Poroshell 120, EC-C18, 3.0 x 100 mm, 2.7 µm. On Shimadzu LC-MS 2020, phytosterols were separated on isometric acetonitrile: methanol (80: 20, v/v, 0.1%

formic acid) over 15 min with 0.6 ml/min flow rate at 25 °C. The intermediate **13**, **14**, **15** and **16** on BR pathway were separated under the linear gradient started from 80% (v/v) methanol in water (1% formic acid) to 100% (v/v) methanol over 8 min and then stay at 100% (v/v) methanol for 12 min at the flow rate of 0.5 ml/min. The mass spectrometer was operated in positive ionization mode. Data was acquired by the electrospray ionization mode. Characteristic m/z+ of the compounds was set in SIM mode as described in figure captions.

Quantification of Campesterol, Sitosterol and Intermediates on BR Synthetic Pathway

For the quantification of campesterol, calibration curve of mass abundance to concentration was made with the authentic campesterol standard solution (Figure S1.11). The R^2=0.9913 shows the reliability of this standard curve, and the peak areas of tested samples were in the range of the standard curve. Mass abundance was represented by the peak area at SIM m/z^+=383.3 mode, which is automatically calculated by Labsolution of Shimadzu LC-MS 2020. For the intermediates **13**, **14**, **15** and **16** on BR pathway, the authentic standards are not available. The concentrations of them were reported as equivalents of campesterol using the standard curve of campesterol and the relative abundance of characteristic ion signals of the corresponding compounds: SIM m/z^+=399.3, 415.3, 439.3, 455.4, respectively.

Adaptive Evolution

Seed culture of yeast strain YYL65 was started with a single colony from an agar plate and grown to stationary phase. Then the culture was back-diluted in 2mL synthetic complex SDM with 2% (w/v) sucrose, starting at OD_{600} ~ 0.1 to early stationary phase (OD_{600} ~6.0). Then appropriate amount of the culture was used to inoculate 2mL fresh synthetic complete SDM with 2% (w/v) sucrose, with the starting OD_{600} ~ 0.1. This procedure was repeated for more than 10 rounds. Then 10 μL culture from the last batch of culture was plated on synthetic complex SDM agar plate. 10 colonies were cultured in 1mL SDM liquid medium at 30 °C for 3 days, which are subjected to LC-MS analysis for campesterol production.

Genome Sequencing

Genomic DNA was extracted with VWR Life Science Yeast Genomic DNA Purification Kit, following the standard protocol from the vendor. The sequencing was done by Novogen Corporation. The chromatogram files generated by NGS platforms (like Illumina HiSeq TM 2000, MiSeq) are transformed by CASAVA Base Calling into sequencing reads, which are called Raw data or Raw reads. Both the sequenced reads and quality score information would be contained in FASTQ files. Raw data is filtered off the reads containing adapter and low-quality reads to obtain clean data for subsequent analysis. The resequencing analysis is based on reads mapping to a common reference sequence by BWA software. SAMTOOLS is used to detect SNP and InDel in functional genomics and get the mutation statistics.

LC-MS/MS Analysis of Intermediates on BR Pathway

Structure of intermediates **13**, **14**, **15** and **16** on BR pathway were identified by Waters Synapt G2-Si Q-TOF system. Yeast sample preparation and LC chromatograph were the same as described in "Analysis of phytosterol and intermediates in BR pathway" session. The mass spectrometer was operated in positive ionization mode. MS/MS fragmentation data was obtained using Collision Induced Dissociation (CID) fragmentation, with collision voltage at 25 V.

Yeast Growth Monitoring

Yeast growth was monitored with OD_{600} measurement. Yeast strains were grown in 2 mL SDM medium in test tubes as described in "Culture and fermentation conditions" session. The yeast culture was collected into microplate (from Coring) at certain time intervals after induction, and diluted with blank SDM medium if necessary. The data was obtained by Synergy HTX Multi-Mode Microplate reader from Bio Tek. The reading from microplate reader was calibrated with spectrophotometer and then converted to OD_{600}.

Abbreviation

BR: brassinosteroids; MVA: mevalonate; IPP: isopentenyl pyrophosphate; DMAPP: dimethylallyl pyrophosphate; SIM: Selected ion monitoring; EIC: extracted ion chromatogram; CYP: cytochrome P450 monooxygenase; InDel: Insert and deletion; SNP: single-nucleotide polymorphism; YPD: yeast extract peptone dextrose medium; SDM: synthetic defined medium; SMT2: 24-methylenesterol C-methyltransferase 2; SMO2: methylsterol monooxygenase 2; ERG10: acetyl-CoA C-acetyltransferase; ERG13: 3-

hydroxy-3-methylglutaryl-CoA (HMG-CoA) synthase; HMG1: HMG-CoA reductase; ERG12: mevalonate kinase; ERG8: phosphomevalonate kinase; ERG19: mevalonate pyrophosphate decarboxylase; ERG20: farnesyl pyrophosphate synthetase; ERG9: squalene synthase; ERG3: Δ^7-sterol-C5(6)-desaturase 1; ERG5: C22-sterol desaturase; ERG4: $\Delta^{24(28)}$-sterol reductase; DWF7: Δ^7-sterol-C5(6)-desaturase 1; DWF5: C7(8)-reductase; DWF1: $\Delta^{24(28)}$-sterol reductase.

Copy Right and Citation

References

1. Dufourc, E. J., Sterols and membrane dynamics. *J Chem Biol* **2008,** *1* (1-4), 63-77.

2. Hartmann, M. A., Plant sterols and the membrane environment. *Trends Plant Sci* **1998,** *3* (5), 170-175.

3. Piironen, V.; Lindsay, D. G.; Miettinen, T. A.; Toivo, J.; Lampi, A. M., Plant sterols: biosynthesis, biological function and their importance to human nutrition. *J Sci Food Agr* **2000,** *80* (7), 939-966.

4. Sonawane, P. D.; Pollier, J.; Panda, S.; Szymanski, J.; Massalha, H.; Yona, M.; Unger, T.; Malitsky, S.; Arendt, P.; Pauwels, L.; Almekias-Siegl, E.; Rogachev, I.; Meir, S.; Cardenas, P. D.; Masri, A.; Petrikov, M.; Schaller, H.; Schaffer, A. A.; Kamble, A.; Giri, A. P.; Goossens, A.; Aharoni, A., Plant cholesterol biosynthetic pathway overlaps with phytosterol metabolism. *Nature plants* **2016,** *3*, 16205.

5. Miettinen, K.; Inigo, S.; Kreft, L.; Pollier, J.; De Bo, C.; Botzki, A.; Coppens, F.; Bak, S.; Goossens, A., The TriForC database: a comprehensive up-to-date resource of plant triterpene biosynthesis. *Nucleic Acids Res* **2018,** *46* (D1), D586-D594.

6. Ohyama, K.; Suzuki, M.; Kikuchi, J.; Saito, K.; Muranaka, T., Dual biosynthetic pathways to phytosterol via cycloartenol and lanosterol in Arabidopsis. *Proc Natl Acad Sci U S A* **2009,** *106* (3), 725-30.

7. Ostlund, R. E., Jr., Phytosterols and cholesterol metabolism. *Curr Opin Lipidol* **2004,** *15* (1), 37-41.

8. American Heart Association Nutrition, C.; Lichtenstein, A. H.; Appel, L. J.; Brands, M.; Carnethon, M.; Daniels, S.; Franch, H. A.; Franklin, B.; Kris-Etherton, P.; Harris, W. S.; Howard, B.; Karanja, N.; Lefevre, M.; Rudel, L.; Sacks, F.; Van Horn, L.; Winston, M.; Wylie-Rosett, J., Diet and lifestyle recommendations revision 2006: a scientific statement from the American Heart Association Nutrition Committee. *Circulation* **2006,** *114* (1), 82-96.

9. Bradford, P. G.; Awad, A. B., Phytosterols as anticancer compounds. *Mol Nutr Food Res* **2007,** *51* (2), 161-170.

10. Woyengo, T. A.; Ramprasath, V. R.; Jones, P. J. H., Anticancer effects of phytosterols. *Eur J Clin Nutr* **2009,** *63* (7), 813-820.

11. Yoshida, Y.; Niki, E., Antioxidant effects of phytosterol and its components. *J Nutr Sci Vitaminol* **2003,** *49* (4), 277-280.

12. Wang, T.; Hicks, K. B.; Moreau, R., Antioxidant activity of phytosterols, oryzanol, and other phytosterol conjugates. *J Am Oil Chem Soc* **2002,** *79* (12), 1201-1206.

13. Gabay, O.; Sanchez, C.; Salvat, C.; Chevy, F.; Breton, M.; Nourissat, G.; Wolf, C.; Jacques, C.; Berenbaum, F., Stigmasterol: a phytosterol with potential anti-osteoarthritic properties. *Osteoarthr Cartilage* **2010,** *18* (1), 106-116.

14. Khripach, V.; Zhabinskii, V.; De Groot, A., Twenty years of brassinosteroids: Steroidal plant hormones warrant better crops for the XXI century. *Ann Bot-London* **2000,** *86* (3), 441-447.

15. Peres, A. L. G. L.; Soares, J. S.; Tavares, R. G.; Righetto, G.; Zullo, M. A. T.; Mandava, N. B.; Menossi, M., Brassinosteroids, the Sixth Class of Phytohormones: A Molecular View from the Discovery to Hormonal Interactions in Plant Development and Stress Adaptation. *Int J Mol Sci* **2019,** *20* (2).

16. Zhabinskii, V. N.; Khripach, N. B.; Khripach, V. A., Steroid plant hormones: effects outside plant kingdom. *Steroids* **2015,** *97*, 87-97.

17. Noguchi, T.; Fujioka, S.; Choe, S.; Takatsuto, S.; Tax, F. E.; Yoshida, S.; Feldmann, K. A., Biosynthetic pathways of brassinolide in Arabidopsis. *Plant Physiol* **2000,** *124* (1), 201-209.

18. Jetter, R., Phytochemicals-Biosynthesis, Function and Application. Springer: 2014.

19. Fujioka, S.; Yokota, T., Biosynthesis and metabolism of brassinosteroids. *Annual review of plant biology* **2003,** *54*, 137-64.

20. Oklestkova, J.; Rarova, L.; Kvasnica, M.; Strnad, M., Brassinosteroids: synthesis and biological activities. *Phytochem Rev* **2015,** *14* (6), 1053-1072.

21. Yokota, T., The structure, biosynthesis and function of brassinosteroids. *Trends in plant science* **1997,** *2* (4), 137-143.

22. Souza, C. M.; Schwabe, T. M.; Pichler, H.; Ploier, B.; Leitner, E.; Guan, X. L.; Wenk, M. R.; Riezman, I.; Riezman, H., A stable yeast strain efficiently producing cholesterol instead of ergosterol is functional for tryptophan uptake, but not weak organic acid resistance. *Metab Eng* **2011,** *13* (5), 555-69.

23. Du, H. X.; Xiao, W. H.; Wang, Y.; Zhou, X.; Zhang, Y.; Liu, D.; Yuan, Y. J., Engineering Yarrowia lipolytica for Campesterol Overproduction. *PloS one* **2016,**

11 (1), e0146773.

24. Tsukagoshi, Y.; Suzuki, H.; Seki, H.; Muranaka, T.; Ohyama, K.; Fujimoto, Y., Ajuga Delta24-Sterol Reductase Catalyzes the Direct Reductive Conversion of 24-Methylenecholesterol to Campesterol. *J Biol Chem* **2016,** *291* (15), 8189-98.

25. Sawai, S.; Ohyama, K.; Yasumoto, S.; Seki, H.; Sakuma, T.; Yamamoto, T.; Takebayashi, Y.; Kojima, M.; Sakakibara, H.; Aoki, T.; Muranaka, T.; Saito, K.; Umemoto, N., Sterol side chain reductase 2 is a key enzyme in the biosynthesis of cholesterol, the common precursor of toxic steroidal glycoalkaloids in potato. *Plant Cell* **2014,** *26* (9), 3763-74.

26. Knoch, E.; Sugawara, S.; Mori, T.; Poulsen, C.; Fukushima, A.; Harholt, J.; Fujimoto, Y.; Umemoto, N.; Saito, K., Third DWF1 paralog in Solanaceae, sterol Delta(24)-isomerase, branches withanolide biosynthesis from the general phytosterol pathway. *Proc Natl Acad Sci U S A* **2018,** *115* (34), E8096-E8103.

27. Yin, Y.; Gao, L.; Zhang, X.; Gao, W., A cytochrome P450 monooxygenase responsible for the C-22 hydroxylation step in the Paris polyphylla steroidal saponin biosynthesis pathway. *Phytochemistry* **2018,** *156*, 116-123.

28. Christ, B.; Xu, C.; Xu, M.; Li, F. S.; Wada, N.; Mitchell, A. J.; Han, X. L.; Wen, M. L.; Fujita, M.; Weng, J. K., Repeated evolution of cytochrome P450-mediated spiroketal steroid biosynthesis in plants. *Nat Commun* **2019,** *10* (1), 3206.

29. Zinser, E.; Paltauf, F.; Daum, G., Sterol composition of yeast organelle membranes and subcellular distribution of enzymes involved in sterol metabolism. *J Bacteriol* **1993,** *175* (10), 2853-8.

30. Quail, M. A.; Kelly, S. L., The extraction and analysis of sterols from yeast. *Methods Mol Biol* **1996,** *53*, 123-31.

31. Ro, D. K.; Paradise, E. M.; Ouellet, M.; Fisher, K. J.; Newman, K. L.; Ndungu, J. M.; Ho, K. A.; Eachus, R. A.; Ham, T. S.; Kirby, J.; Chang, M. C.; Withers, S. T.; Shiba, Y.; Sarpong, R.; Keasling, J. D., Production of the antimalarial drug precursor artemisinic acid in engineered yeast. *Nature* **2006,** *440* (7086), 940-3.

32. Westfall, P. J.; Pitera, D. J.; Lenihan, J. R.; Eng, D.; Woolard, F. X.; Regentin, R.; Horning, T.; Tsuruta, H.; Melis, D. J.; Owens, A.; Fickes, S.; Diola, D.; Benjamin, K. R.; Keasling, J. D.; Leavell, M. D.; McPhee, D. J.; Renninger, N. S.; Newman, J. D.; Paddon, C. J., Production of amorphadiene in yeast, and its conversion to dihydroartemisinic acid, precursor to the antimalarial agent artemisinin. *Proc Natl Acad Sci U S A* **2012,** *109* (3), E111-8.

33. Darnet, S.; Rahier, A., Plant sterol biosynthesis: identification of two distinct families of sterol 4alpha-methyl oxidases. *Biochem J* **2004,** *378* (Pt 3), 889-98.

34. Carland, F.; Fujioka, S.; Nelson, T., The sterol methyltransferases SMT1, SMT2, and SMT3 influence Arabidopsis development through nonbrassinosteroid products. *Plant Physiol* **2010,** *153* (2), 741-56.

35. Jacquier, N.; Schneiter, R., Mechanisms of sterol uptake and transport in yeast. *The Journal of steroid biochemistry and molecular biology* **2012,** *129* (1-2), 70-8.

36. Choi, S.; Cho, Y. H.; Kim, K.; Matsui, M.; Son, S. H.; Kim, S. K.; Fujioka, S.; Hwang, I., BAT1, a putative acyltransferase, modulates brassinosteroid levels in Arabidopsis. *Plant J* **2013,** *73* (3), 380-91.

37. Yang, H.; Bard, M.; Bruner, D. A.; Gleeson, A.; Deckelbaum, R. J.; Aljinovic, G.; Pohl, T. M.; Rothstein, R.; Sturley, S. L., Sterol esterification in yeast: a two-gene process. *Science* **1996,** *272* (5266), 1353-6.

38. Zweytick, D.; Leitner, E.; Kohlwein, S. D.; Yu, C.; Rothblatt, J.; Daum, G., Contribution of Are1p and Are2p to steryl ester synthesis in the yeast Saccharomyces cerevisiae. *Eur J Biochem* **2000,** *267* (4), 1075-82.

39. Wang, Q.; Quan, S.; Xiao, H., Towards efficient terpenoid biosynthesis: manipulating IPP and DMAPP supply. *Bioresources and Bioprocessing* **2019,** *6* (1), 6.

40. Morikawa, T.; Mizutani, M.; Aoki, N.; Watanabe, B.; Saga, H.; Saito, S.; Oikawa, A.; Suzuki, H.; Sakurai, N.; Shibata, D.; Wadano, A.; Sakata, K.; Ohta, D., Cytochrome P450 CYP710A encodes the sterol C-22 desaturase in Arabidopsis and tomato. *Plant Cell* **2006,** *18* (4), 1008-22.

41. Zweytick, D.; Hrastnik, C.; Kohlwein, S. D.; Daum, G., Biochemical characterization and subcellular localization of the sterol C-24(28) reductase, erg4p, from the yeast saccharomyces cerevisiae. *FEBS Lett* **2000,** *470* (1), 83-7.

42. Zidovetzki, R.; Levitan, I., Use of cyclodextrins to manipulate plasma membrane cholesterol content: evidence, misconceptions and control strategies. *Biochim Biophys Acta* **2007,** *1768* (6), 1311-24.

43. Aguilera, F.; Peinado, R. A.; Millan, C.; Ortega, J. M.; Mauricio, J. C., Relationship between ethanol tolerance, H+ -ATPase activity and the lipid composition of the plasma membrane in different wine yeast strains. *Int J Food Microbiol* **2006,** *110* (1), 34-42.

44. Soares, E. V., Flocculation in Saccharomyces cerevisiae: a review. *J Appl Microbiol* **2011,** *110* (1), 1-18.

45. Cheng, C.; Zhang, M.; Xue, C.; Bai, F.; Zhao, X., Development of stress tolerant Saccharomyces cerevisiae strains by metabolic engineering: New aspects from cell flocculation and zinc supplementation. *J Biosci Bioeng* **2017,** *123* (2), 141-146.

46. Vriet, C.; Russinova, E.; Reuzeau, C., Boosting Crop Yields with Plant Steroids. *Plant Cell* **2012,** *24* (3), 842-857.

47. Eremina, M.; Unterholzner, S. J.; Rathnayake, A. I.; Castellanos, M.; Khan, M.; Kugler, K. G.; May, S. T.; Mayer, K. F. X.; Rozhon, W.; Poppenberger, B., Brassinosteroids participate in the control of basal and acquired freezing tolerance of plants. *P Natl Acad Sci USA* **2016,** *113* (40), E5982-E5991.

48. Fujita, S.; Ohnishi, T.; Watanabe, B.; Yokota, T.; Takatsuto, S.; Fujioka, S.; Yoshida, S.; Sakata, K.; Mizutani, M., Arabidopsis CYP90B1 catalyses the early C-22 hydroxylation of C27, C28 and C29 sterols. *Plant J* **2006,** *45* (5), 765-74.

49. Ohnishi, T.; Watanabe, B.; Sakata, K.; Mizutani, M., CYP724B2 and CYP90B3 function in the early C-22 hydroxylation steps of brassinosteroid biosynthetic pathway in tomato. *Biosci Biotechnol Biochem* **2006,** *70* (9), 2071-80.

50. Biggs, B. W.; Lim, C. G.; Sagliani, K.; Shankar, S.; Stephanopoulos, G.; De Mey, M.; Ajikumar, P. K., Overcoming heterologous protein interdependency to optimize P450-mediated Taxol precursor synthesis in Escherichia coli. *Proc Natl Acad Sci U S A* **2016,** *113* (12), 3209-14.

51. Szekeres, M.; Nemeth, K.; KonczKalman, Z.; Mathur, J.; Kauschmann, A.; Altmann, T.; Redei, G. P.; Nagy, F.; Schell, J.; Koncz, C., Brassinosteroids rescue the deficiency of CYP90, a cytochrome P450, controlling cell elongation and de-etiolation in arabidopsis. *Cell* **1996,** *85* (2), 171-182.

52. Ohnishi, T.; Godza, B.; Watanabe, B.; Fujioka, S.; Hategan, L.; Ide, K.; Shibata, K.; Yokota, T.; Szekeres, M.; Mizutani, M., CYP90A1/CPD, a brassinosteroid biosynthetic cytochrome P450 of Arabidopsis, catalyzes C-3 oxidation. *J Biol Chem* **2012,** *287* (37), 31551-60.

53. Kim, J. E.; Jang, I. S.; Son, S. H.; Ko, Y. J.; Cho, B. K.; Kim, S. C.; Lee, J. Y., Tailoring the Saccharomyces cerevisiae endoplasmic reticulum for functional assembly of terpene synthesis pathway. *Metab Eng* **2019,** *56*, 50-59.

54. Charles E. Bills, O. N. M., Paul S. Prickett, Factors Determining the Ergosterol Content of Yeast. *Journal of Biological Chemistry* **1930,** *87*, 259-264.

55. Li, Y.; Li, S.; Thodey, K.; Trenchard, I.; Cravens, A.; Smolke, C. D., Complete biosynthesis of noscapine and halogenated alkaloids in yeast. *Proc Natl Acad Sci U S A* **2018,** *115* (17), E3922-E3931.

56. Alberti, S.; Gitler, A. D.; Lindquist, S., A suite of Gateway cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae. *Yeast* **2007,** *24* (10), 913-9.

57. Gibson, D. G.; Young, L.; Chuang, R. Y.; Venter, J. C.; Hutchison, C. A., 3rd; Smith, H. O., Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **2009,** *6* (5), 343-5.

58. Shao, Z.; Zhao, H.; Zhao, H., DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res* **2009,** *37* (2), e16.

Chapter II: Optimization of Campesterol-Producing Yeast Strains for Plant Membrane-Bound Enzyme Reconstitution

Abstract

Campesterol is a major phytosterol that plays important roles in regulating membrane properties and serves as the precursor to multiple specialized metabolites such as the phytohormone brassinosteroids. Recently, we established a campesterol-producing yeast strain and extended the bioproduction to 22-hydroxycampesterol and 22-hydroxycampest-4-en-3-one, precursors to brassinolide. However, there is a the tradeoff in growth due to disrupted sterol metabolism. In this study, we enhanced the growth of the campesterol-producing yeast by partially restoring the activity of the sterol acyl transferase and engineering upstream FPP supply. Furthermore, genome sequencing analysis also revealed a pool of genes possibly associated with the altered sterol metabolism. Retro engineering implies an essential role of ASG1, especially the C-terminal asparagine-rich domain of ASG1, in the sterol metabolism of yeast especially under stress. The performance of the campesterol-producing yeast strain was enhanced with the titer of campesterol to 18.4 mg/L, and the stationary OD_{600} was improved by ~ 33% comparing to the unoptimized strain. In addition, we examined the activity of a plant cytochrome P450 in the engineered strain, which exhibits more than 9-fold higher activity than when expressed in the wild-type yeast strain. Therefore, the engineered campesterol-producing yeast strain also serves as a robust host for the functional expression of plant membrane protein.

Introduction

Sterols are essential membrane components in mammals, plants, fungi, and bacteria, regulating membrane fluidity and permeability[1], and involved in protein trafficking and signal transduction[2]. Different organisms produce different types of sterols. Mammals produce cholesterol that serves as the precursor to steroid hormones, bile acids, and vitamin D; fungi produce ergosterol, the precursor of vitamin D2; plants produce a mixture with campesterol and β-sitosterol as the dominant components[1]. Phytosterols function as precursors to several plant specialized metabolites, such as the plant hormone brassinosteroids (BRs) and the pharmaceutically intriguing withanolides. Similar to sterol hormones in animals, BRs play important roles in regulating plant growth and development, as well as resistance to biotic and abiotic stress[3-5].

Cholesterol has been reported to modulates the function of various types of membrane proteins[6], such as receptors, ion channels, transporters, and peptides, either by directly binding to the proteins or indirectly affecting membrane properties. The cholesterol-producing yeast has been believed to provide a more preferred membrane microenvironment for the functional reconstitution of mammalian membrane proteins, and has been validated by enhanced stability and activity of a Na,K-ATPase α3β1 isoform in a "humanized" cholesterol-producing *Pichia pastoris* strain[7]. Similarly, the composition of the plasma membrane has been shown to be critical to the activity of membrane proteins in plants[8]. Phytosterols differ from cholesterol and ergosterol primarily in the side chains, e.g., campesterol and sitosterol are equipped with a C24-methyl or C24-ethyl group, respectively. Plants exhibit distinct membrane properties from fungi and metazoans, such

as lower membrane ordering than cholesterol or ergosterol composed membrane[1]. Plant membrane proteins play important roles in plant perception, immunity, and metabolism[9-12]. The microenvironment created by phytosterols is believed to be crucial for optimal enzymatic activities[13]. Thus, a phytosterol-producing yeast strain might also function like a "plantinized" yeast more adapted to the functional reconstitution of plant membrane proteins. However, whether phytosterol-producing yeast exhibits similar characteristics remains to be explored.

In our previous study, we engineered *Saccharomyces cerevisiae* to produce campesterol instead of ergosterol and reconstituted the early-stage pathway towards the synthesis of BRs[14]. The biosynthesis of campesterol was achieved using plant enzymes DWF1, DWF5, and DWF7 in *S. cerevisiae*. Further upregulation of upstream MVA pathway enzymes enhanced the production of campesterol yet not accessible to downstream biosynthetic enzymes, and further investigation implies that it is mainly because phytosterols are almost completely esterified in the engineered strain. Thus, acyltransferases ARE1, ARE2, and sterol C-24(28) reductase ERG4 were inactivated to yield yeast strain YYL65 to produce free campesterol that is accessible to the downstream enzyme such as CYP90B1. However, YYL65 exhibits significant growth burden due to these genetic modifications. The growth rate at the exponential phase of YYL65 decreased to 1/7 of the wild-type. The OD_{600} at the stationary phase of YYL65 was reduced to less than 1/2 of that of the wildtype yeast strain under the same cultivation condition. Further adaptive evolution of YYL65 generated YYL67 with both enhanced growth (stationary OD_{600} only ~ 2/3 of wild type) and campesterol production (7 mg/L), yet YYL67 still grows

much weaker than the wild-type yeast. Additionally, YYL67 has lower tolerance to Li^+ and heat, compared to the wild type.

In this study, we extended the engineering efforts to further enhance the performance of the engineered phytosterol-producing yeast platform. In addition to more robust growth, the engineered strain also exhibits higher resistance to Li^+ and heat. We also examined whether the change of the sterol composition could plantinize the engineered yeast strain through testing the activity of a cytochrome P450 CYP82Y1 in the optimized campesterol-producing yeast strains. Surprisingly, the activity of CYP82Y1 in producing 1-hydroxy-N-methylcandadine was enhanced by more than 9-fold in the engineered campesterol-producing strains than that in the wildtype yeast strain. The engineered yeast strain not only serves as a more robust platform to produce campesterol, further reconstitute and investigate BR biosynthesis, but also as a potential workhorse for the functional reconstitution of plant membrane-bound proteins.

Results and Discussion

Partially Restoring the Function of Acyltransferases Can Rescue Yeast Growth

Previously, the inactivation of the two acyl-CoA sterol acyltransferases ARE1 and ARE2 in YYL67 (Table S2.1) resulted in the synthesis of free campesterol as well as an evident growth deficiency[14]. The growth deficiency is likely caused by the defected sterol biosynthesis as well as the potential toxicity of free phytosterols, specifically campesterol in yeast. Prior investigation showed that ARE1 and ARE2 exhibit distinct substrate

specificity: ARE2 prefers ergosterol as the substrate while ARE1 has higher promiscuity and prefers sterol precursors than the end product ergosterol[15].

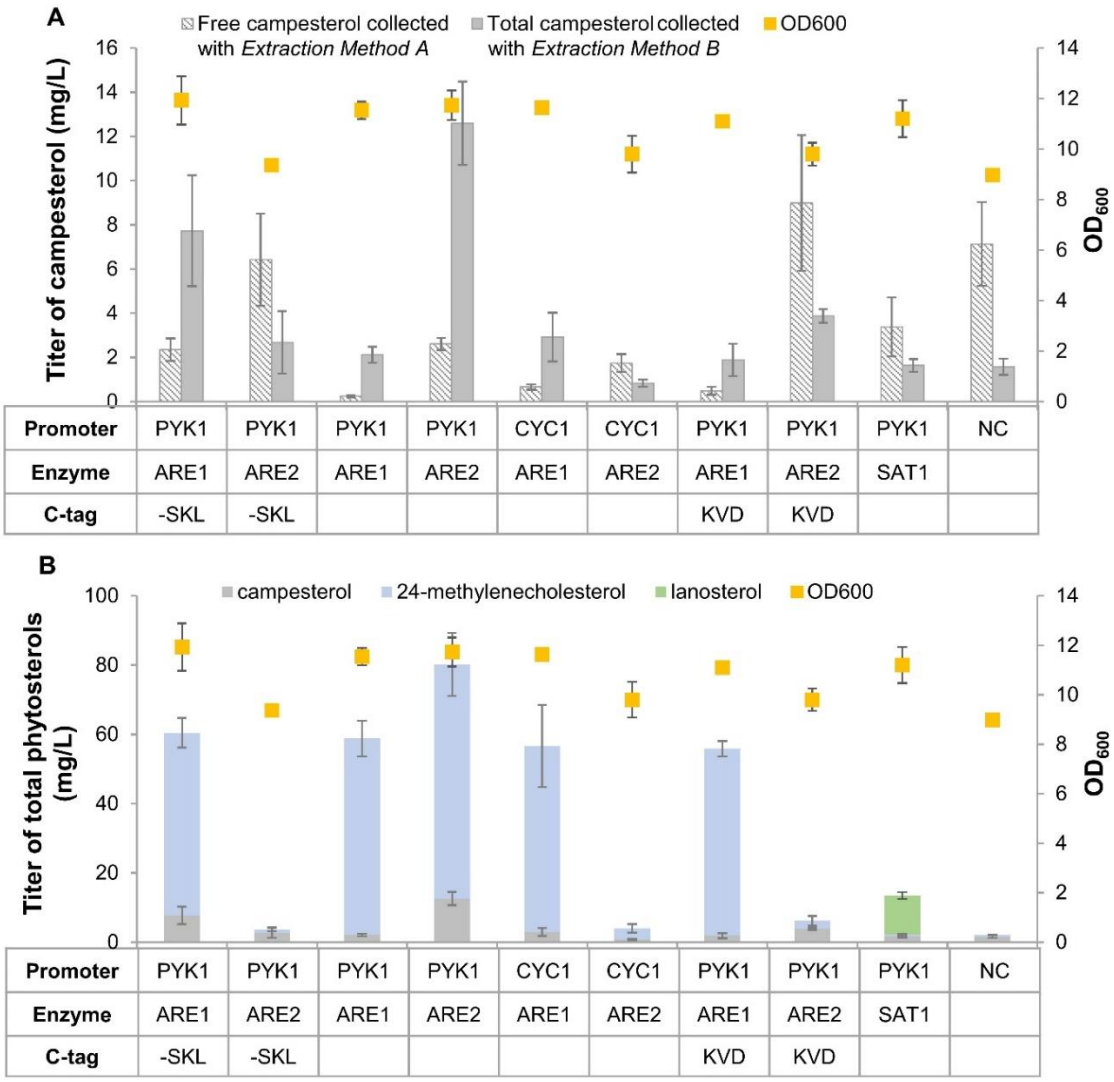


Figure 2.1. The effects of ARE1 and ARE2 to sterol production in YYL67. **(A)** Campesterol production of YYL67 expressing ARE1 or ARE2 using different promoters and/or with different C-terminal localization tags. **(B)** Production of total phytosterols in YYL67 expressing ARE1 or ARE2 using different promoters and/or with different C-terminal localization tags. NC represents negative control which means the strain harboring an empty vector. Free campesterol was extracted without saponification using *Extraction Method A*. Total campesterol was extracted by saponification using KOH, as described in *Extraction Method B*. *Extraction Method B* exhibits more significant sample loss than *Extraction Method A* due to the harsh treatment of the sample during saponification. Error bars represent standard deviation of three biological replicates.

Thus, ARE1 and ARE2 were each introduced into YYL67 with different expression conditions (different promoter strengths and intracellular localization) to find a solution that can restore the growth of YYL67 with a minimal effect on campesterol production. ARE1 and ARE2 were expressed downstream of either a strong constitutive *PYK1* promoter (P_{PYK1}) or a weak *CYC1* promoter (P_{CYC1}). In addition, ARE1 and ARE2 were fused to C-terminal signaling peptides to redirect the acyltransferases to localize on organelles besides the ER membrane, where ARE1 and ARE2 are natively localized and where sterols are synthesized and esterified [16]. The effects of two signal peptides were examined: the C-terminal tag -SKL that was reported to localize enzymes to peroxisomes [17], and the C-terminal tag -KVD which was a Golgi retention C-terminal motif [18]. The localization of ARE1, ARE2 with C-terminus tags was confirmed in YYL63 (Δ*are1*Δ*are2*; Table S2.1) using confocal fluorescence microscopy (Figure S2.1). The C-terminus tag -SKL led ARE2 to peroxisomes which were marked by a native peroxisomal membrane protein PEX3-YFP[19]; the C-terminus tag -KVD led ARE2 to Golgi which was marked by a native Golgi membrane protein VRG4-Dsred[19] (Figure S2.1). The OD_{600} of each strain at stationary phase was measured as an indicator of yeast growth, and the productions of free and total (including free and esterified form) campesterol and the precursor 24-methylenecholesterol (the two major phytosterols synthesized in YYL67) were estimated (Figure 2.1). Free sterols were extracted using *Extraction Method A* (Materials and Methods) without saponification, while "total sterols" (both free sterols and sterol esters) were extracted using *Extraction Method B* (Materials and Methods). Due to the harsh saponification procedure involving high concentration treatment of KOH, using *Extraction*

Method B exhibits a lower efficiency compared to *Extraction Method A* in recovering the corresponding sterols.

As expected, introducing ARE1 or ARE2 into YYL67 generally resulted in enhanced stationary OD_{600}, and the production of free campesterol was decreased but at different levels. Notably, ARE1 and ARE2 exhibit distinct esterification efficiency towards campesterol and 24-methylenecholesterol. When expressed downstream of the constitutive P_{PYK1} in YYL67, ARE2 enhanced the level of total campesterol by ~7 folds with ~60% decrease of free campesterol, while expressing ARE1 downstream of P_{PYK1} did not change the level of total campesterol much in YYL67 but substantially decreased the level of free campesterol (Figure 2.1A). This set of data indicates that ARE2 more efficiently esterifies campesterol compared to ARE1. Meanwhile, the production of 24-methylenecholesterol was significantly enhanced when ARE1 or ARE2 was introduced into YYL67 and regulated by P_{PYK1} (Figure 2.1B). Similar phenotypes (enhanced growth, enhanced production of 24-methylenecholsterol in ester form) were observed when ARE1 was expressed downstream of weak P_{CYC1} or redirected to other organelles than ER (Figure 2.1B). However, under the same expression conditions (downstream of P_{CYC1} or redirected to the other organelles than ER), ARE2 exhibits a similar phenotype to YYL67. This indicates that ARE1 is relatively more efficient in esterifying 24-methylenecholesterol than ARE2. Although ARE1 and ARE2 can both esterify campesterol and 24-methylenecholesterol when expressed at high level at ER membranes, they do exhibit distinct substrate specificity. The observation that ARE1 prefers 24-methylenecholesterol (unsaturated C24-28) is consistent with previous reports that ARE1 prefers yeast sterol

intermediates with unsaturated C24-28[15]. ARE2, on the other hand, prefers campesterol (saturated C24-28) and ergosterol (saturated C24-28) as reported[15].

SAT1 is a plant sterol O-acyltransferase and was characterized to esterify mainly phytosterol precursors, such as cycloartenol and lanosterol, while campesterol is not its favorable substrate[20]. Thus, we also introduced SAT1 into YYL67, downstream of a strong P_{PYK1} to enhance the growth and minimize the effects on the production of free campesterol. Although SAT1 barely esterified 24-methylenecholesterol or campesterol, expressing SAT1 in YYL67 did enhance the stationary OD_{600} (Figure 2.1). Since SAT1 has been reported to mainly function at the upstream part of the sterol biosynthesis, the level of lanosterol was measured (Figure 2.1B). Consistent with previous investigations[20], a significant amount of lanosterol in the ester form was detected only when SAT1 was expressed in YYL67, which also exhibited a higher stationary OD_{600} than YYL67 alone. These results indicate that although sterol esters do not directly constitute membrane, they play a role in yeast growth under stress condition. In general, the growth of yeast strain seems to be positively correlated with the total levels of sterols (mostly in ester form), either lanosterol or phytosterols (Figure 2.1B). ARE2 expressed downstream of P_{PYK1} was selected for the subsequent study due to the relatively robust growth and higher free campesterol production among strains with higher stationary OD_{600} than YYL67.

Increasing Production of FPP or GGPP Can Improve the Growth of YYL67

ERG20 is a farnesyl pyrophosphate synthetase (FPS) that catalyzes the conversion of isopentenyl pyrophosphate (IPP) to geranyl pyrophosphate (GPP) and GPP to farnesyl pyrophosphate (FPP). FPP is the precursor to steroid synthesis and has often been

engineered to enhance yeast growth or isoprenoid precursor supply[21-24]. ERG20^{F96C} has been identified to be a dual function enzyme that is not only able to synthesize FPP but also exhibits the function of BTS1, the native geranylgeranyl diphosphate synthase (GGPPS) in yeast, to convert FPP to geranylgeranyl diphosphate (GGPP). Overexpressing ERG20^{F96C} into yeast was reported to improve the production of diterpenes, and the improvement was higher than overexpressing BTS1[23]. In addition to ERG20^{F96C}, ERG20^{A99G} was reported to improve the growth and increase the total isoprene production[24]. ERG20^{F96C}, ERG20^{A99G}, wild type ERG20, and BTS1 were expressed downstream of a constitutive promoter P_{TEF1} in YYL67, all of which resulted in enhanced growth but decreased production of campesterol by ~22 folds, ~1.3 folds, 0.8 folds and ~4 folds, respectively (Figure 2.2A, 2.2B). No accumulation of 24-methylenecholesterol or lanosterol was detected. The enhanced growth when expressing ERG20^{F96C} and BTS1 indicates that YYL67 is likely deficient in GGPP supply, while the decrease in campesterol production is likely due to the enhanced consumption of FPP towards GGPP by the function of ERG20^{F96C} and BTS1 as GGPPS.

ERG20^{A99G} exhibited the most enhanced growth as well as the least decreased campesterol production in comparison to YYL67 and was introduced into YYL67 along with *ARE2* through genome integration to generate YYL102 (Table S2.1) for downstream optimization. The optimized strain YYL102 has higher growth rate and a minor decrease of campesterol production compared to YYL67 (Figure 2.2C, 2.2D). The tolerance of YYL102 to Li^+ and heat was increased substantially as well (Figure 2.2E).

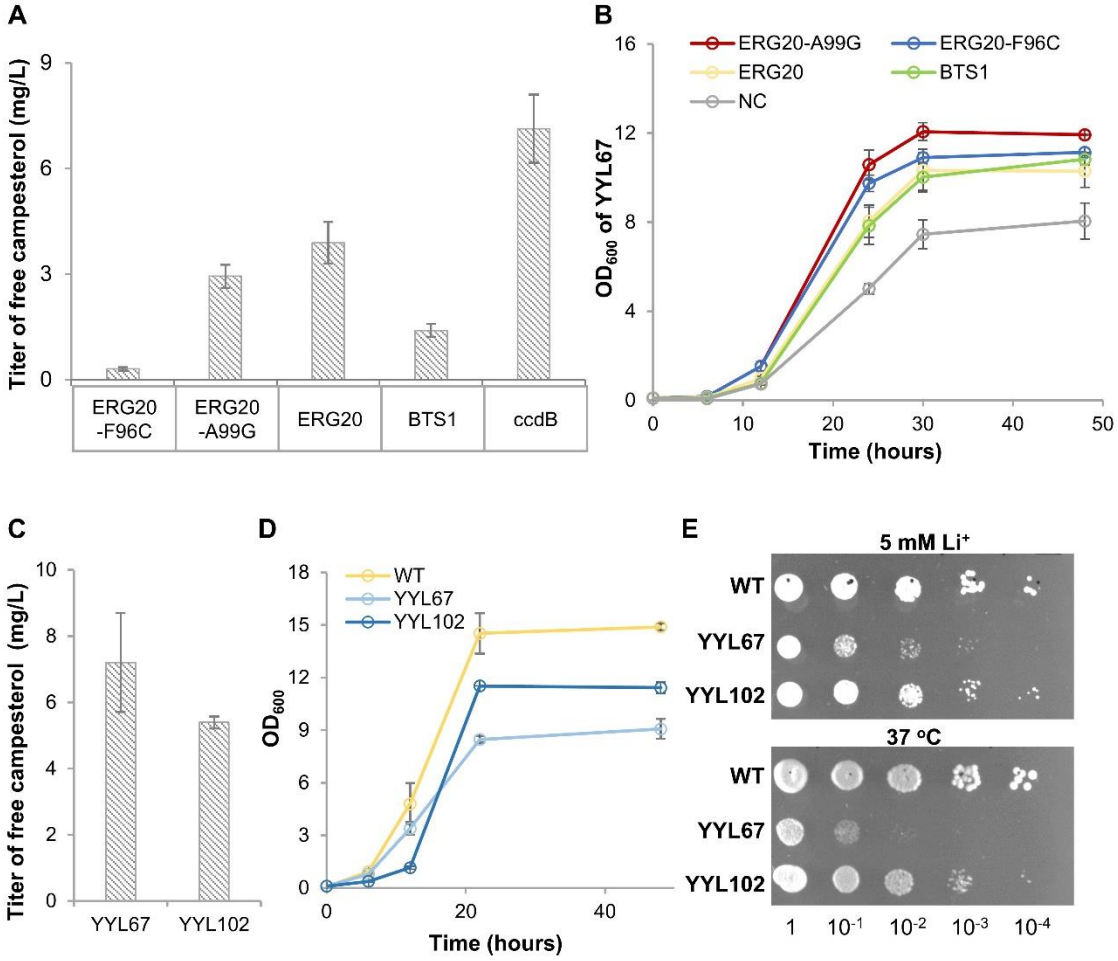


Figure 2.2. Effects of ERG20 mutants on campesterol production in YYL67. (A) Production of campesterol in YYL67 expressing ERG20, ERG20^{F96C}, ERG20^{A99G}, and BTS1. NC represents negative control which means the strain harboring an empty vector. (B) Growth of YYL67 expressing ERG20, ERG20^{F96C}, ERG20^{A99G}, and BTS1 when cultured in SDM with auxotrophic selections at 30 ℃. (C) Production of campesterol in the optimized strain YYL102 (YYL67 expressing ERG20^{A99G} and ARE2 through genomic integration, Table S2.1), YYL67. (D) Growth of wild-type (WT) yeast, YYL67, and YYL102 when cultured in SDM at 30 ℃. (E) Li+ tolerance and heat tolerance of YYL67 and YYL102. Error bars represent standard deviation of three biological replicates.

Genome Sequencing of YYL67 Provides Candidates to Engineer for Further Optimization of Growth and Campesterol Production

YYL67 is an adaptive-evolved strain exhibiting higher production of campesterol and more robust growth in comparison to its parental strain, YYL65[14]. The mutations in YYL67 may play a role in enhancing the growth and campesterol production, and are also promising candidates to engineer for enhanced performance of the strain. Therefore, we sequenced the genomes of YYL65, YYL67, and another evolved strain YYL65-5[14]. Paired-end libraries (150-bp) were generated by NGS platforms and sequenced to an average depth of over 50× for each strain. The high coverage enabled reliable genome-wide analysis of insertion and deletion (InDel) and single-nucleotide polymorphisms (SNPs), which were widely found through the genomes of YYL67 and YYL65-5 (Figure S2.2, Table S2.1, S2.3, S2.4, S2.5). Compared to the parental strain YYL65, YYL67 and YYL65-5 carry 27 and 35 InDels in gene-coding regions (Table S2.4), respectively. 18 of these InDels are present in both YYL67 and YYL65-5 (Table S2.4); many of these genes are related to transcriptional regulation and growth stress response (Table S2.3B). On the other hand, approximately 300 SNPs occur within open reading frames (ORFs) in YYL67 and YYL65-5. The SNPs affect 19 genes in YYL67 and 22 genes in YYL65-5 at amino acid level, with 14 genes found in both YYL67 and YYL65-5 (Table S2.3C and S2.5). The genes affected by InDels or SNPs were considered promising candidates to engineer for enhanced growth or campesterol production in yeast.

To examine the influence of these genes on the growth and sterol production in YYL67, the 41 mutated genes (27 from InDel analysis, 19 from SNPs, with 3 redundant

genes with both InDels and SNPs, 2 genes cannot be inactivated due to technique problems, Table S2.4-S2.6) were inactivated individually in YYL67 as well as the wildtype CEN.PK2-1D using CRISPR/CAS9 system. Although the inactivation of most of the genes did not result in any change of phenotypes compared to the negative control (inactivation of the corresponding gene in CEN.PK2-1D, Figure S2.3), inactivation of ASG1 was lethal to YYL67 and greatly deteriorated the growth of wild-type yeast. Thus, ASG1 was considered an important gene contributing to YYL67′s enhanced performance in comparison to the parental strain YYL65 (Figure S2.3). ASG1 is an activator of stress response genes and the null mutants of ASG1 have a respiratory deficiency. We quantified the mRNA level of ASG1 in YYL67, the parental strain YYL65, and the growth optimized strain YYL102 by reverse transcription polymerase chain reaction (RT-PCR). The transcription level of ASG1 was the highest in YYL65, the one with the lowest growth rate[14] among the three strains (Figure S2.4). Meanwhile YYL102 had the most robust growth among the three strains exhibited the lowest level of *ASG1* mRNA (Figure S2.4). This result suggests that ASG1 is likely to be overexpressed when the strain is under stress.

ASG1 was then upregulated in YYL67 by introducing an extra copy of the genes for further analysis. *ASG1* gene was cloned from genomic DNA of YYL67 (marked as $ASG1^{m}$) and YYL65 (marked as $ASG1^{ori}$). The cloned genes were sequenced, and $ASG1^{m}$ has one less asparagine (N) at the C-terminal N-rich region, compared to $ASG1^{ori}$ (Table S2.2, S2.7). Both versions of ASG1 were expressed downstream of the strong, constitutive *GPD* promoter (P_{GPD}) from a low-copy number plasmid in YYL67 and YYL102 (Table S2.2). Surprisingly, the overexpression of ASG1 in YYL67 enhanced the campesterol

production by ~ 1 fold but did not improve the yeast growth (Figure 2.3A). However, the overexpression of ASG1s in the more robust strain YYL102 did not lead to significant change in either the stationary OD_{600}, the free campesterol production (Figure 2.3A), or the saponified total campesterol level (Figure S2.5). ASG1 likely plays a role in sterol metabolism in yeast under stress condition. In addition, overexpression of $ASG1^{ori}$ in YYL67 enhanced campesterol production to a slightly higher extent than overexpression of $ASG1^{m}$, which hints at the importance of the N-rich region to the function of ASG1. $ASG1^{\Delta N}$ without N-rich region and $ASG1^{½N}$ with half of the N-rich region were constructed and overexpressed downstream of P_{GPD} in YYL67 and YYL102 (Table S2.2, S2.7, Figure 2.3A). In YYL67, $ASG1^{\Delta N}$ resulted in the highest production of campesterol (~2-fold higher) among all the mutants with a similar stationary OD_{600} in comparison to YYL67, and $ASG1^{½N}$ showed similar phenotype as $ASG1^{m}$. No significant differences in campesterol production, either in free or esterified form, were detected when overexpressing ASG1 and mutants in YYL102 (Figure 2.3A, S2.5).

To investigate how ASG1 responds to the growth stress, YYL67 expressing different ASG1 variants or mutants were cultured in two different medium conditions: synthetic defined medium (SDM) supplemented with 5% ethanol and 2×SDM, representing culture conditions under stress[25] and with extra nutrients, respectively. When cultured in 2× SDM, YYL67 expressing ASG1 and mutants exhibit more robust growth and a similar level of campesterol production, compared to when cultured in the standard medium (Figure 2.3B). On the other hand, in the presence of 5% ethanol, the expression of $ASG1^{ori}$ or $ASG1^{m}$ did not result in obvious changes in yeast growth or campesterol

production in comparison to YYL67 harboring an empty vector (negative control, NC); while the introduction of $ASG1^{\Delta N}$ or $ASG1^{½N}$ into YYL67 decreased both the growth and the campesterol production comparing with NC (Figure 2.3C). Interestingly, the stationary OD_{600} was positively correlated with the length of the N-rich region of ASG1 when cultured in the presence of 5% ethanol: YYL67 expressing $ASG1^{\Delta N}$ showed the lowest OD_{600} ~5.0, while $ASG1^{ori}$ did not affect the growth of YYL67 much (Figure 2.3C). The different influences of expressing $ASG1^{\Delta N}$ and $ASG1^{ori}$ under stress (5% ethanol) suggests that the N-rich region of ASG1 plays an important role in the stress response in yeast, which is related to yeast growth and the sterol synthesis. As the N-rich region is usually related to protein aggregation[26], the morphology of $ASG1^{\Delta N}$, $ASG1^{½N}$, and $ASG1^{ori}$ was examined by confocal fluorescence microscopy (Figure S2.6). However, the deletion of N-rich region did not affect the aggregation morphology of ASG1 in yeast. This agrees with the previous study on another yeast transcription factor AZF1 that the deletion of N-rich region does not affect the localization or prion formation[27]. Further mechanistic investigations are needed to understand the precise role of ASG1 and the function of its N-rich region in yeast metabolism and stress response.

Among the other genes that were inactivated in YYL67, *IMD2* also significantly deteriorated the growth of YYL67 but was not lethal, yet did not affect the growth of CEN.PK2-1D (Figure S2.3). IMD2 is inosine monophosphate dehydrogenase, catalyzing the rate-limiting step in GTP biosynthesis. IMD2 was then cloned from the genomic DNA of YYL67 (marked as $IMD2^{m}$) and YYL65 (marked as $IMD2^{ori}$). $IMD2^{m}$ has 7 mutated sites at amino acid level than $IMD2^{ori}$ (Table S2.7). $IMD2^{m}$ and $IMD2^{ori}$ were individually

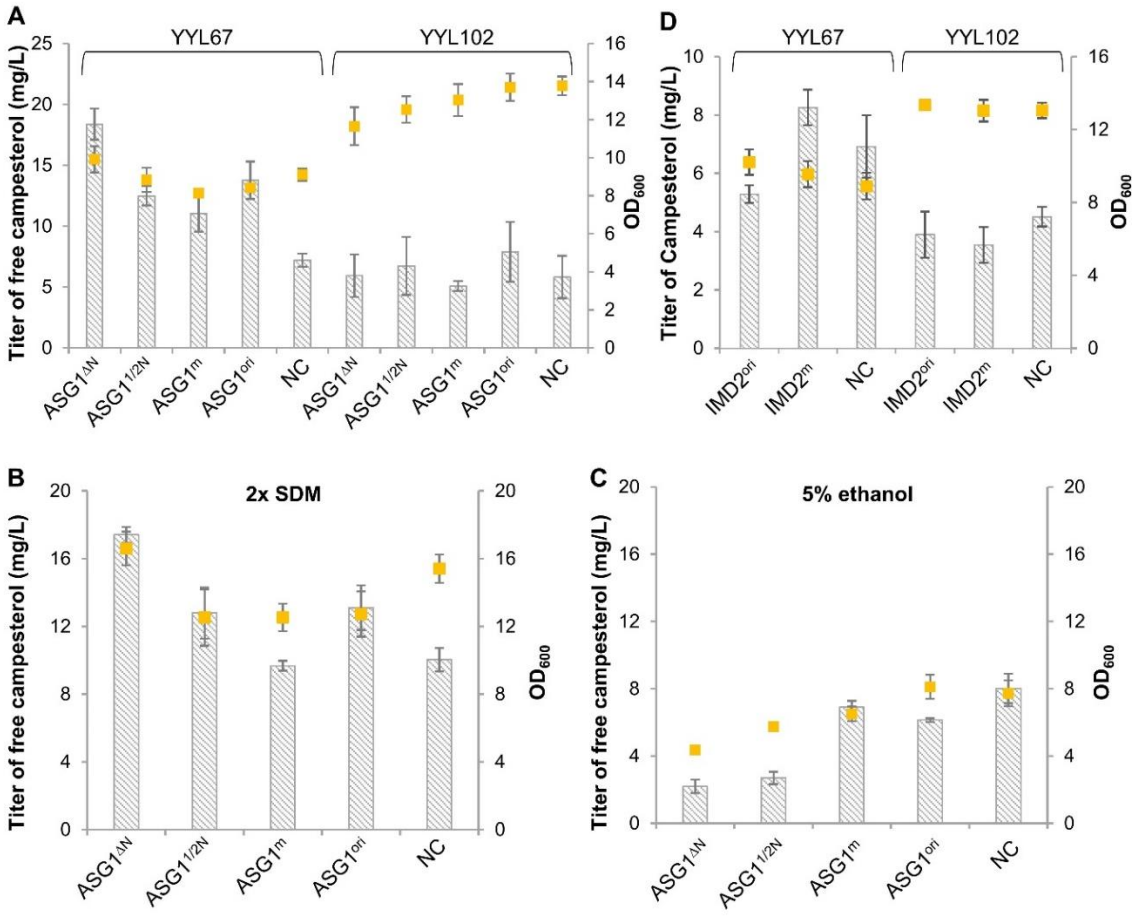


Figure 2.3. Effects of ASG1 mutants on sterol production and yeast growth in YYL67 and YYL102. (A) Production of campesterol and stationary OD_{600} of YYL67 or YYL102 expressing different ASG1 variants ($ASG1^{\Delta N}$, $ASG1^{1/2N}$, $ASG1^{m}$, and $ASG1^{ori}$) cultured in SDM. (B) Production of campesterol and stationary OD_{600} of the YYL67 expressing different ASG1 variants ($ASG1^{\Delta N}$, $ASG1^{1/2N}$, $ASG1^{m}$, and $ASG1^{ori}$) when cultured in the 2× SDM. (C) Production of campesterol and stationary OD_{600} of YYL67 expressing different ASG1 variants ($ASG1^{\Delta N}$, $ASG1^{1/2N}$, $ASG1^{m}$, and $ASG1^{ori}$) when cultured in SDM supplement with 5% ethanol. (D) Production of campesterol and stationary OD_{600} of YYL67 or YYL102 expressing $IMD2^{m}$ or $IMD2^{ori}$. NC represents negative control which means the strain harboring an empty vector. Error bars represent standard deviation of three biological replicates.

overexpressed in YYL67 downstream of the P_{GPD} from a low-copy number plasmid, but no significant effects on the growth or campesterol production were observed (Figure 2.3D). This result suggests that IMD2, though nonessential, plays a role in yeast growth, but not specifically involved in sterol metabolism or stress.

However, we also realized that the InDels and SNPs proposed from the genome analysis did not describe the genetic divergences between YYL67 (built from CEN.PK2-1D) and its parental strain YYL65 precisely enough. This is likely due to the fact that we only sequenced one strain each and used the genome data of *S. cerevisiae* S288C due to the lack of open-source genome information of CEN.PK2-1D, the basal strain of YYL65 and YYL67. Genome-wide association studies (GWAS) analysis with the genome sequencing data of more biological replicates of the related strains (YYL65, YYL67) should be conducted to thoroughly understand the mechanism of yeast to adapt the growth and metabolic burden caused by altered genetic material and sterol metabolism.

The Presence of Free Campesterol Enhanced the Activity of Plant Membrane Enzyme in Yeast

Membrane composition and morphology are essential to the activities of plant membrane proteins, which are heavily involved in plant metabolism and defense[9, 10]. The altered sterol composition in the yeast strains may provide a different microenvironment for the functional reconstitution of plant membrane proteins. We examined the activity of a cytochrome P450 CYP82Y1, which converts N-methylcanadine to 1-hydroxy-N-methylcanadine, in the set of phytosterol-producing strains we constructed in this study: YYL56 (producing campesterol ester)[14], YYL67 (producing free campesterol), and

YYL102 (producing free campesterol with enhanced growth) (Table S2.1). CYP82Y1 was expressed downstream of a constitutive promoter *TPI1* (P_{TPI1}) from a low copy number plasmid (Table S2.2). TNMT was co-expressed with CYP82Y1 to afford the synthesis of the precursor of CYP82Y1, N-methylcanadine (Figure 2.4). Yeast strains were cultured in SDM medium fed with 12.5 μM canadine, the substrate of TNMT. The activity of CYP82Y1 was estimated using the production of 1-hydroxy-N-methylcanadine, the product of CYP82Y1. CYP82Y1 exhibited similar level of activities in YYL56 (Table S2.2), the strain producing campesterol mainly in the easter form, and the wildtype yeast strain (CEN-PK2.1D). Surprisingly, in comparison to CEN-PK2.1D, the production of 1-hydroxy-N-methylcanadine was up to 3-fold higher in YYL67, and ~ 4-fold higher in YYL102 (Figure 2.4). On the other hand, the enzymatic activity of TNMT1, likely a cytosolic enzyme, exhibited the same efficiency towards the synthesis of N-methylcanadine across different campesterol-producing yeast strains (Figure S2.7). Our results indicate that the production of free campesterol in yeast does "plantinize" the yeast membranes (both plasma membranes and intracellular membranes) and enhance the enzymatic efficiency of plant membrane-bound proteins.

Because it was found that overexpressing $ASG1^{\Delta N}$ in YYL67 enhanced the production of free campesterol, we wished to investigate the effect of free campesterol level on the activity of CYP82Y1 in yeast, so $ASG1^{\Delta N}$ was co-expressed with TNMT1 and CYP82Y1 in YYL67. Interestingly, co-expressing $ASG1^{\Delta N}$ did not enhance the enzymatic efficiency of CYP82Y1 in YYL67 although the titer of free campesterol was ~2-fold higher than that of YYL67 (Figure 2.4, S2.7B). On the other hand, co-expressing $ASG1^{\Delta N}$

enhanced the enzymatic efficiency of CYP82Y1 in YYL102 although the titer of free campesterol is the same in comparison with YYL102 alone (Figure 2.4, S2.7B). Although free campesterol is an essential component for the "platinization" of the yeast membrane, the activity of plant membrane enzyme is not positively correlated with the amount of free sterol.

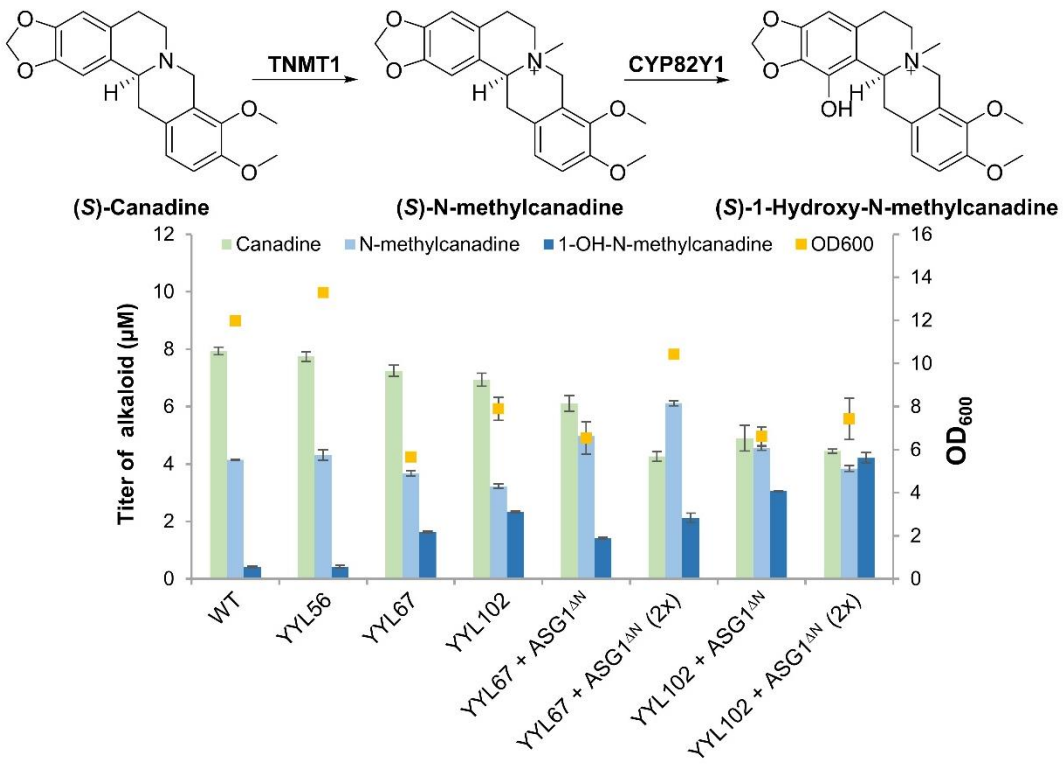


Figure 2.4. Activity of the plant cytochrome P450 CYP82Y1 towards the synthesis of 1-hydroxy-N-methylcanadine in different yeast. The biosynthetic pathway from canadine to 1-hydroxy-N-methylcanadine catalyzed by TNMT and CYP82Y1 is illustrated on the top. Production of 1-hydroxy-N-methylcanadine of different yeast strains expressing TNMT1 and CYP82Y1: wile type (WT) yeast strain, YYL56, YYL67, YYL102, YYL67 expressing $ASG1^{\Delta N}$, YYL102 expressing $ASG1^{\Delta N}$, cultured in either standard SDM or 2× SDM (marked as ″2×″). Error bars represent standard deviation of three biological replicates.

We also observed enhancement in the activity of CYP82Y1 in both YYL67 and YYL102 expressing $ASG1^{\Delta N}$ when cultured in the 2× SDM (in comparison to the standard SDM; Figure 2.4). Since culturing in 2× SDM did not enhance the level of campesterol in either YYL67 or YYL102 (Figure S2.7B) expressing $ASG1^{\Delta N}$ in comparison to standard SDM, this further implies that the activity of plant membrane-bound enzymes is not positively correlated with the level of free phytosterol in yeast. However, how $ASG1^{\Delta N}$ improves the functional reconstitution of CYP82Y1 in YYL102 but not YYL67 remains unclear.

In this study, we further optimize the "plantinized" yeast platform by partially restoring sterol esterification mechanism, engineering upstream FPP pathway, and performing retro genetic engineering. The enhanced growth, Li^+ and heat tolerance ensures the applications of this platform to reconstitute plant membrane anchored enzymes. The discovery of ASG1's roles in yeast growth and sterol production under stress as well as plant membrane-bound enzyme reconstitution highlights ASG1 as an important target for further investigation and engineering for sterol-related metabolic engineering in yeast.

Materials and Methods

Materials

Campesterol (~98%) was obtained from Sigma-Aldrich. Yeast strains were cultured at 30 ^{o}C in complex yeast extract peptone dextrose (YPD, all components from BD Diagnostics) medium or synthetic defined medium (SDM) containing yeast nitrogen

base (YNB) (BD Diagnostics), ammonium sulfate (Fisher Scientific), 2% (w/v) glucose unless specified and the appropriate dropout (Takara Bio) solution for selection.

General Technique for DNA Manipulation

PCR reactions were performed with Expand high Fidelity system (Sigma-Aldrich), Phusion DNA Polymerase (NEB), Q5 High-Fidelity DNA Polymerase (New England Biolabs) or Taq Polymerase (NEB) according to manufacturer's protocols. PCR products were purified by Zymoclean Gel DNA Recovery Kit (Zymo Research). Plasmids were prepared with Econospin columns (Epoch Life Science) according to manufacturer's protocols. All DNA constructs were confirmed through DNA sequencing by Azenta life science Inc. BP Clonase II Enzyme Mix, Gateway pDONR221 Vector, LR Clonase II Enzyme Mix (Life Technologies), and the *S. cerevisiae* Advanced Gateway Destination Vector Kit [28] (Addgene) were used to perform the Gateway Cloning. The Gibson one-pot, isothermal DNA assembly[57] was conducted at the scale of 10 μL by incubating T5 exonuclease (NEB), Phusion polymerase (NEB), Taq ligase (NEB) and 50 ng of each DNA fragment at 50 ^{o}C for 1 hour to assemble multiple DNA fragments. Yeast strains are constructed through homologous recombination and DNA assembly[58]. Yeast strains, plasmids, and sequences of genes used in this work are listed in Table S2.1, Table S2.2, and Table S2.7, respectively.

Culture Conditions

For metabolite analysis, yeast strains were first cultured overnight in 500 μL SDM with 2% (w/v) glucose in 96-well plates (BD faulcon) at 30 ^{o}C overnight, at 250 r.p.m..

Appropriate volume of the overnight seed culture was inoculated in fresh 500 µL SDM with glucose to make OD_{600} around 0.1 and incubated at 30 ^{o}C for 72 hours before metabolite analysis of the yeast pellets. To monitor the growth, yeast strains were grown in 2 mL SDM with 2% (w/v) glucose in test tubes. The yeast strains were first cultured overnight in 500 µL SDM with 2% (w/v) glucose. Appropriate volume of the overnight culture was inoculated in fresh 2 mL SDM with 2% (w/v) glucose to make OD_{600} around 0.1 and incubated at 30 ^{o}C for 72 hour, at 250 r.p.m.. Yeast growth was monitored through measuring OD_{600}: appropriate amount of yeast culture was collected into the cuvette and diluted with water as needed to make the reading within the linear range of the instrument (usually from OD_{600} 0.1 to 0.8). OD_{600} of the culture was measured by Nanodrop (Molecular Device^{TM}).

Analysis and Quantification of Phytosterols in Yeast

Phytosterols were collected and analyzed from yeast pellets. *Extraction Method A*: To extract free sterols from yeast, 100 µL acetone was added to yeast cell pellets (from 500 µL yeast culture), followed by the addition of 400 µL methanol with vigorous vortex to mix. The supernatant was transferred to new tubes. Repeat the extraction of the pellets by adding an additional 400 µL ethyl acetate. The supernatants were combined and dried in the vacufuge (Eppendorf). *Extraction Method B*: To extract total sterols (including sterol esters and free sterols) from yeast, 100 µL 60% (w/v) KOH and 100 µL ethanol was added to the cell pellets, then incubated at 86 °C for 1 hour for saponification. The mixture was then cooled down to room temperature and 400 µL petroleum ether was added followed by vortexing for 5 minutes. Organic phases were collected to new tubes. 400 µL ethyl

acetate was added to the remaining aquatic phase and repeat the vortex step. Organic phases were combined and dried in vacufuge. The dried samples were dissolved in 40 μL methanol and analyzed using LC-MS. Samples extracted by *Extraction Method A* has higher recover ratio than that of *Extraction Method B.*

The samples were then analyzed by reverse phase LC-MS on a Shimadzu 2020 single quadrupole LC-MS (Poroshell 120, EC-C18, 3.0 × 100 mm, 2.7 μm) using positive ionization. To analyze sterols, the metabolites were separated under the linear gradient started from 80% methanol (v/v in water, 0.1% formic acid) to 100% methanol (v/v in water, 0.1% formic acid) for 8 min followed by 100% methanol (v/v in water, 0.1% formic acid) for 12 min at the flow rate of 0.5 ml/min. Data was acquired by the electrospray ionization mode. The sterols were analyzed using selected ion monitoring (SIM) with either the characteristic or possible mass-to-charge ratio (m/z^+) according to the structures and molecular weights. The quantities of campesterol were estimated by comparing to the standard curves using authentic campesterol standard solutions.

Genome Sequencing

Genomic DNA was extracted with VWR Life Science Yeast Genomic DNA Purification Kit following the standard protocol from the vendor. The sequencing was done by Novogen Corporation. The chromatogram files generated by NGS platforms (like Illumina HiSeq TM 2000, MiSeq) are transformed by CASAVA Base Calling into sequencing reads, which are called Raw data or Raw reads. Both the sequenced reads and quality score information would be contained in FASTQ files. Raw data is filtered off the

reads containing adapter and low-quality reads to obtain clean data for subsequent analysis. The resequencing analysis is based on reads mapping to a common reference sequence (*S. cerevisiae* S288C) by BWA software. SAMTOOLS is used to detect SNP and InDel in functional genomics and get the mutation statistics.

Quantification of RNA Expression Level by Quantitative RT-PCR

2 mL of 48-hour culture of yeast strains was cultivated as described in "Culture Conditions" section. Cell pellet was collected, and RNA was purified by Zymo YeaStar™ RNA kit. RNA concentration was quantified by Nanodrop (Molecular Device™). 200 ng of RNA of each strain was converted to cDNA by reverse transcription PCR kit from NEB. 1 μL of converted cDNA was used as the template and amplified by PCR using Taq polymerase (NEB). The constitutively expressed gene TDH3 was used as a control to normalize the cDNA concentration. PCR products were run electrophoresis and analyzed by ChemiDoc MP Imager (BIO RAD). The band intensity was quantified by the analysis tool of ChemiDoc MP Imager (BIO RAD). The intensity of target genes was normalized by the intensity of TDH3. Three independent replicates were examined in this experiment.

Application of CRISPR-Cas9 to Inactive Genes in Yeast

The plasmids used for CRISPR-based gene inactivation in yeast (Fig. S2.8) were constructed through Gibson assembly of the backbone vector and DNA fragment encoding sgRNA and repair DNA (synthesized from IDT). The repair DNA contains 100 bp of homologous sequence of target cDNA incorporating an in-frame stop codon within ± 30 bp from the cutting site recognized by sgRNA. The sgRNA used to inactivate each gene

were listed in Table S2.6. The sgRNA not targeting to any genes in yeast was used as a negative control (pYL1228, Table S2.2).

Analysis and Quantification of N-Methylcanadine, 1-Hydroxy-N-Methylcanadine in Yeast

Canadine was fed to the yeast culture at inoculation with a final concentration of 12.5 μM. Culture was centrifuged to separate the liquid medium and the cell pellets, and the liquid medium was analyzed by LC-MS. The metabolites were separated under the linear gradient started from 20% methanol to 60% methanol (v/v in water, 0.1% formic acid) for 7 min at the flow rate of 0.5 ml/min. The quantities of canadine were estimated by comparing to the standard curves using authentic canadine standard solutions. The quantities of N-methylcanadine and 1-hydroxy-N-methylcanadine were estimated with mass intensity area multiplied by the ionization coefficient which calculated by using the principle of conservation of mass.

Reference

1. Dufourc, E. J., Sterols and membrane dynamics. *J Chem Biol* **2008**, *1* (1-4), 63-77.

2. Tong, J.; Manik, M. K.; Im, Y. J., Structural basis of sterol recognition and nonvesicular transport by lipid transfer proteins anchored at membrane contact sites. *Proceedings of the National Academy of Sciences of the United States of America* **2018**, *115* (5), E856-E865.

3. De Bruyne, L.; Hofte, M.; De Vleesschauwer, D., Connecting growth and defense: the emerging roles of brassinosteroids and gibberellins in plant innate immunity. *Molecular plant* **2014**, *7* (6), 943-59.

4. Peres, A. L. G. L.; Soares, J. S.; Tavares, R. G.; Righetto, G.; Zullo, M. A. T.; Mandava, N. B.; Menossi, M., Brassinosteroids, the Sixth Class of Phytohormones: A Molecular View from the Discovery to Hormonal Interactions in Plant Development and Stress Adaptation. *Int J Mol Sci* **2019**, *20* (2).

5. Khripach, V.; Zhabinskii, V.; De Groot, A., Twenty years of brassinosteroids: Steroidal plant hormones warrant better crops for the XXI century. Ann Bot-London **2000**, 86 (3), 441-447.

6. Grouleff, J.; Irudayam, S. J.; Skeby, K. K.; Schiott, B., The influence of cholesterol on membrane protein structure, function, and dynamics studied by molecular dynamics simulations. *Biochim Biophys Acta* **2015**, *1848* (9), 1783-95.

7. Hirz, M.; Richter, G.; Leitner, E.; Wriessnegger, T.; Pichler, H., A novel cholesterol-producing Pichia pastoris strain is an ideal host for functional expression of human Na,K-ATPase alpha 3 beta 1 isoform. *Appl Microbiol Biot* **2013**, *97* (21), 9465-9478.

8. Zauber, H.; Burgos, A.; Garapati, P.; Schulze, W. X., Plasma membrane lipid-protein interactions affect signaling processes in sterol-biosynthesis mutants in Arabidopsis thaliana. *Frontiers in plant science* **2014**, *5*, 78.

9. Luschnig, C.; Vert, G., The dynamics of plant plasma membrane proteins: PINs and beyond. *Development* **2014**, *141* (15), 2924-38.

10. Hrycay, E. G.; Bandiera, S. M., The monooxygenase, peroxidase, and peroxygenase properties of cytochrome P450. *Arch Biochem Biophys* **2012**, *522* (2), 71-89.

11. Simons, K.; Sampaio, J. L., Membrane organization and lipid rafts. *Cold Spring Harb Perspect Biol* **2011**, *3* (10), a004697.

12. Du, Y.; Fu, X.; Chu, Y.; Wu, P.; Liu, Y.; Ma, L.; Tian, H.; Zhu, B., Biosynthesis and the Roles of Plant Sterols in Development and Stress Responses. *Int J Mol Sci* 2022, *23* (4).

13. Schaller, H., The role of sterols in plant growth and development. *Progress in lipid research* **2003**, *42* (3), 163-75.

14. Xu, S.; Chen, C.; Li, Y., Engineering of Phytosterol-Producing Yeast Platforms for Functional Reconstitution of Downstream Biosynthetic Pathways. *ACS Synth Biol* **2020**, 9 (11), 3157-3170.

15. Zweytick, D.; Leitner, E.; Kohlwein, S. D.; Yu, C.; Rothblatt, J.; Daum, G., Contribution of Are1p and Are2p to steryl ester synthesis in the yeast Saccharomyces cerevisiae. *Eur J Biochem* **2000**, *267* (4), 1075-82.

16. Jacquier, N.; Schneiter, R., Mechanisms of sterol uptake and transport in yeast. *The Journal of steroid biochemistry and molecular biology* **2012**, *129* (1-2), 70-8.

17. DeLoache, W. C.; Russ, Z. N.; Dueber, J. E., Towards repurposing the yeast peroxisome for compartmentalizing heterologous metabolic pathways. *Nat Commun* **2016**, *7*, 11152.

18. Gao, C.; Cai, Y.; Wang, Y.; Kang, B. H.; Aniento, F.; Robinson, D. G.; Jiang, L., Retention mechanisms for ER and Golgi membrane proteins. *Trends in plant science* **2014**, *19* (8), 508-15.

19. Zhu, J.; Zhang, Z. T.; Tang, S. W.; Zhao, B. S.; Li, H.; Song, J. Z.; Li, D.; Xie, Z., A Validated Set of Fluorescent-Protein-Based Markers for Major Organelles in Yeast (Saccharomyces cerevisiae). *mBio* **2019**, *10* (5).

20. Chen, Q.; Steinhauer, L.; Hammerlindl, J.; Keller, W.; Zou, J., Biosynthesis of phytosterol esters: identification of a sterol o-acyltransferase in Arabidopsis. *Plant physiology* **2007**, *145* (3), 974-84.

21. Fischer, M. J.; Meyer, S.; Claudel, P.; Bergdoll, M.; Karst, F., Metabolic engineering of monoterpene synthesis in yeast. *Biotechnology and bioengineering* **2011**, *108* (8), 1883-92.

22. Ignea, C.; Pontini, M.; Maffei, M. E.; Makris, A. M.; Kampranis, S. C., Engineering monoterpene production in yeast using a synthetic dominant negative geranyl

diphosphate synthase. *ACS Synth Biol* **2014**, *3* (5), 298-306.

23. Ignea, C.; Trikka, F. A.; Nikolaidis, A. K.; Georgantea, P.; Ioannou, E.; Loupassaki, S.; Kefalas, P.; Kanellis, A. K.; Roussis, V.; Makris, A. M.; Kampranis, S. C., Efficient diterpene production in yeast by engineering Erg20p into a geranylgeranyl diphosphate synthase. *Metab Eng* **2015**, *27*, 65-75.

24. Rubat, S.; Varas, I.; Sepulveda, R.; Almonacid, D.; Gonzalez-Nilo, F.; Agosin, E., Increasing the intracellular isoprenoid pool in Saccharomyces cerevisiae by structural fine-tuning of a bifunctional farnesyl diphosphate synthase. *Fems Yeast Res* **2017**, *17* (4).
25. Ding, J.; Huang, X.; Zhang, L.; Zhao, N.; Yang, D.; Zhang, K., Tolerance and stress response to ethanol in the yeast Saccharomyces cerevisiae. *Appl Microbiol Biotechnol* **2009**, *85* (2), 253-63.

26. Michelitsch, M. D.; Weissman, J. S., A census of glutamine/asparagine-rich regions: implications for their conserved function and the prediction of novel prions. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97* (22), 11910-5.

27. Stewart, T.; Wolfe, B. E.; Fuchs, S. M., Defining the role of the polyasparagine repeat domain of the S. cerevisiae transcription factor Azf1p. *PloS one* **2021**, *16* (5), e0247285.

28. Alberti, S.; Gitler, A. D.; Lindquist, S., A suite of Gateway cloning vectors for high-throughput genetic analysis in Saccharomyces cerevisiae. *Yeast* **2007**, *24* (10), 913-9.

Chapter III: A Multifunctional Scaffold Protein MSBP1 Enhanced the Activities of BR Biosynthetic Enzymes

Abstract

Brassinosteroids (BRs) are a group of plant hormones derived from phytosterols, involved in growth and development regulation, and biotic and abiotic response. Brassinolide (BL) is the most bioactive BR. The biosynthesis of BL from campesterol requires the function of at least four cytochrome P450s and one reductase. In the previous chapters, the functions of the first two cytochrome P450s, CYP90A1 and CYP90B1, have been reconstituted in the phytosterol-producing yeast platforms. Further investigation and characterization of the remaining BR biosynthetic enzymes was impeded likely due to the low enzymatic efficiency of CYP90B1 and insufficient BR intermediate production. Here, we discovered that MSBP1, a scaffold protein previously reported to support the biosynthesis of lignin, can substantially enhance the production of BR intermediate, 22-hydroxycampest-4-en-3-one. Moreover, MSBP1 also improved the activity of CYP90C1 towards the synthesis of C23-hydroxylation efficiency. The new enzymatic function of CYP90C1 to catalyze C2-hydroxylation was discovered for the first time when co-expressed with MSBP1 in yeast. The physical interactions between MSBP1 and BR biosynthetic enzymes were characterized using Yeast-Two-Hybrid (Y2H) system, suggesting that MSBP1 is likely to play an important role in recruiting BR synthetic enzymes to form a metabolon in yeast and plants.

Introduction

Brassinosteroids (BRs) are essential plant hormones that are widely found in plant kingdoms, regulating growth and development processes, including seed germination, root growth, stem elongation and flowering[1-3]. They elicit a broad-spectrum resistance to pathogens and also response to abiotic stress, such as salinity and drought[4-6]. The BR defective mutants exhibit dwarf phenotypes. Brassinosteroids is a big group that around 70 naturally occurring brassinosteroids have been reported since it was found in 1979, among which brassinolide (BL) is the most bioactive BR[7]. The biosynthesis of brassinolide have been intensively studied for decades and many BR intermediates and related enzymes have been identified and characterized, but several enzymatic steps remain elusive (Figure 3.1). The biosynthesis of BL from campesterol involves at least 5 enzymes (CYP90A1, CYP90B1, CYP90C1, CYP85A2, DET2) (Table S3.3). Based on our previous study of BR biosynthetic enzyme reconstitution in yeast (Chapter I), campesterol is first converted to 22-hydroxy-campesterol, **13**, by CYP90B1, followed by CYP90A1 to oxidize **13** to 22-hydroxy-campest-4-en-3-one, **14**. This C22-hydroxylation pathway is also believed to be the first and rate-limiting step of BR biosynthetic pathway in *Arabidopsis. thaliana*. **14** is converted to 22-Hydroxy-5α-campest-3-one, **17**, by 5α-reductase DET2 in *A. thaliana*. (Figure 3.1). CYP90C1 and CYP90D1, have the same function to catalyze hydroxylation at C23 on the C22-hydroxylated sterols (e.g., 22-hydroxy-5α-campest-3-one)3-epi-6-deoxocathasterone,,)[8]. CYP85A2 is reported to catalyze C6 oxidation and Baeyer-Villiger Oxidation of castasterone to brassinolide[9] (Figure 3.1). CYP85A1 has a putative function in the C6 oxidation as CYP85A2[10]. However, the enzymes responsible for the reduction of

C3 keto and C2 hydroxylation remains unknown. The biosynthetic pathways of BR remains unclear, likely due to the high promiscuity of BR biosynthetic enzymes making the biosynthetic pathway highly cross-linked, and the low abundance of BRs in nature and the instability of certain intermediates making them hard to identified.

In Chapter II, we optimized the phytosterol-producing yeast YYL102 with enhanced growth and higher tolerance to heat and Li^+. The activity of CYP90A1 and CYP90B1 in YYL102 (Table S3.1) towards the synthesis of **14** is slightly lower than in YYL67 (Figure S3.1). Further reconstitution of enzymes putatively functions downstream of CYP90A1, CYP90C1 or CYP90D1, did not convert **14** towards downstream products; while the reconstitution of DET2 vigorously consumed **14** but a clear products was not detected using LC-MS analysis (Figure S3.2), which could be due to the relatively low production and low ionization efficiency of the product of DET2. Thus, we attempted to further enhance the production of **14** in yeastmatabolon. Recently, a membrane steroid binding protein, MSBP1, has been found to play an important role in forming a lignin biosynthetic metabolon with monolignol P450 monooxygenases[11]. MSBP1 acted as essential structural components in ER membrane and stabilize the P450 monooxygenase complex in plants[11]. It turns out that MSBP1 not only involved in lignin biosynthesis but also in BR biosynthesis. In addition to forming a lignin biosynthetic metabolon, MSBP1 has also been found to play a number of different but important physiological roles in plant, including but not limited to, interacting with BAK1 and negatively regulating BR signaling[13]. This led to the hypothesis that MSBP1 may play a similar role in the phytosterol biosynthesis; and BR biosynthesis may also require physical interactions

among the biosynthetic enzymes to form a kinetically more efficient metabolon as lignin biosynthesis.

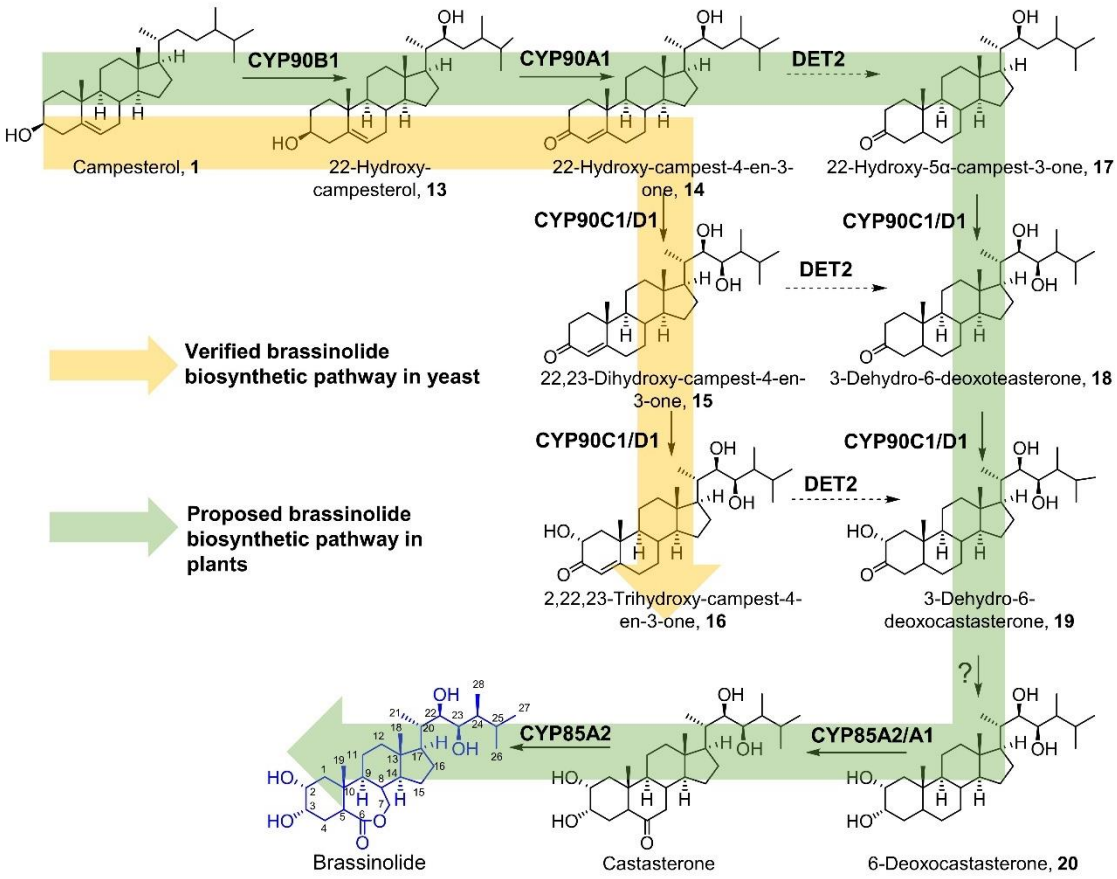


Figure 3.1 The proposed biosynthetic pathway of brassinosteroids. The question marker represents that the enzymes catalyzing the corresponding conversion is unknown. Conversions that remain to be identified or confirmed are represented using dashed arrows.

In this study, we found that expression of MSBP1 enhanced the production of **14** by 2 folds in YYL69 (Table S3.1). Co-expression with MSBP1 in YYL69 also enhanced the enzymatic efficiency of CYP90C1 and CYP90D1. CYP90C1 and CYP90D1 have been believed to be functional redundant enzymes that catalyze C23-hydroxylation on substrates such as **14**. The C23 hydroxylation of **14** was confirmed in YYL69 expressing

CYP90C1/CYP90D1 with MSBP1. Additionally, our results also indicate that CYP90C1 or CYP90D1 might catalyze the C2 hydroxylation, which has not been reported before. The fact that MSBP1 enhanced the efficiency of BR biosynthetic enzymes implies that these BR biosynthetic enzymes might function as a complex. To validate this hypothesis, ubiquitin-based yeast-two-hybrid system was used to confirm the physical interactions between MSBP1 to BR biosynthetic enzymes. MSBP1 shows interaction with ATR1, CYPP90A1, CYP90B1, and CYP85A2, but not with CYP90C1, CYP90D1 or DET2. Moreover, we also observed interactions between different pairs of BR biosynthetic enzymes. Our results confirmed the physical contact/interactions between BR biosynthetic enzymes and non-enzymetic steroid bining proteins (MSBP1 or MSBP2), and strongly implies that BR biosynthesis takes place within an enzyme complex, i.e., efficient BR biosynthesis requires the formation of a metabolon in plant. For long, the C2 keto-reductase and C2 hydroxylase are missing from BR biosynthesis. Our results on the surprising role MSBP1 in BR biosynthesis, improving the putative C2-hydroxylation activity of CYP90C1 or CYP90D1, and the failure in the detection of CS and BL in yeast, suggests that the key missing components for reconstituting BL biosynthesis in yeast might be non-catalytic proteins rather than enzymes.

Results

Expression of MSBP1 Can Enhance the Production of BR Intermediates in Yeast

MSBP1 is a membrane steroid-binding protein which was proved to bind to several plant steroids in *in vitro* experiments, such as progesterone, 5-dihydrotestosterone, 24-epi-

brassinolide (24-eBL), and stigmasterol, with different affinities[12]. MSBP1 can negatively regulate the BL signaling by suppressing CYP90A1 and CYP90C1 expression in plants upon BL exogenous treatment[13]. Lately, MSBP1 was reported to play a role as a scaffold protein to recruit three lignin biosynthetic cytochrome P450s together and increase the stability of the complex, thus enhancing the enzymatic efficiency[11]. The multiple roles of MSBP1 in physiological progress and its correlation to BL signaling and steroid binding intrigued us to investigate its role in BR biosynthesis. MSBP1 and its homologue MSBP2 were expressed downstream of a constitutive promoter P_{GPD} through a low-copy plasmid together or individually to YYL69 (Table S3.1, S3.2), a strain that expressed CYP90A1 and CYP90B1 and produce a BR intermediate, 22-hydroxycampest-4-en-3-one, **14** (Figure 3.1). The production of **14** was quantified to estimate the activity of CYP90A1 and CYP90B1. The expression of MSBP1 and MSBP2 enhanced the production of **14**, by ~1.8 and ~1.0 folds, respectively (Figure 3.2A). Co-expression of MSBP1 and MSBP2 in YYL69 did not further enhance the production of **14**. MSBP1 shows slightly higher enhancement on **14** production than MSBP2 and thus was used for further investigation. MSBP1 was predicted to have four domains or regions, including an ER lumen domain (domain A), a transmembrane domain (domain B), a steroid-binding domain (domain C) and a disorder domain (domain D). Different combinations of these domains were examined by introducing them to YYL69 and quantifying the production of **14** (Figure 3.2A). Expression of MSBP1-BCD in YYL69 lead to an enhanced titer of **14** than when wildtype MSBP1 was expressed in YYL69, which indicated that domain A of MSBP1 is not essential to the function of MSBP1. The titer of **14** in YYL69 expressing MSBP1-ABC

or MSBP1-BC were at the same level and slightly lower than that of MSBP1, but significantly higher than MSBP1-ABD (p value ~ 0.2%) (Figure 3.2A). It suggests that C domain is important for the full function of MSBP1. However, only expressing C domain in YYL69 did not improve the production of **14**. MSBP1-BC and MSBP1-CD both exhibited higher production of **14** than C domain only (Figure 3.2A). These results indicates that B and D domains were also essential for MSBP1 function.

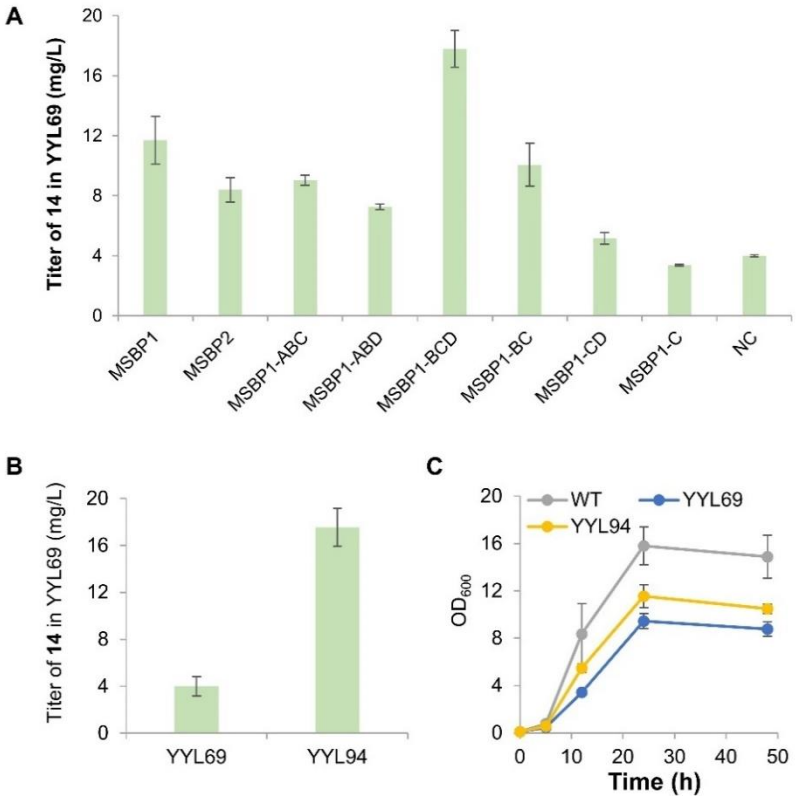


Figure 3.2 Effects of MSBP1 on 22-hydroxycampest-4-en-3-one, **14**, production in YYL69. (A) Production of 22-hydroxycampest-4-en-3-one, **14**, of YYL69 expressing MSBP1 and MSBP2 and different MSBP1 domain combinations (MSBP1-ABC, ABD, BCD, BC, CD, C) and the empty vector (NC) cultured in SDM. (B) Production of campesterol and stationary OD_{600} of the YYL67 expressing different ASG1 variants ($ASG1^{\Delta N}$, $ASG1^{1/2N}$, $ASG1^{m}$, and $ASG1^{ori}$) when cultured in the 2× SDM. (B) Production of 22-hydroxycampest-4-en-3-one, **14**, in strain YYL69 and the optimized strain YYL94 (YYL69 expressing $ERG20^{A99G}$, ARE2 and MSBP1 through genomic integration). (C) Growth of wild-type (WT) yeast, YYL69, and YYL94 when cultured in SDM at 30 ℃. Error bars represent standard deviation of three biological replicates.

Characterization of BR Biosynthetic Enzymes in the Presence of MSBP1

The optimized strain YYL94 allowed us to reconstitute more BR biosynthetic enzymes in this platform. DET2, the 5α-reductase to reduce **14**, was first introduced downstream of a constitutive promoter P_{GPD} in YYL94, but the products were not detected even though the substrate **14** was largely consumed (Table S3.1, S3.2). To avoid this bottleneck, we tested CYP90C1 and CYP90D1 in YYL94, which were reported to have the same putative catalytic activities to hydroxylate C23 on C22 hydroxylated BR intermediates[8] (Figure 3.1). CYP90C1 and CYP90D1 were expressed individually downstream of a constitutive promoter P_{GPD} in YYL94 (integrated with *MSBP1*) (Table S3.1, S3.2) and 22, 23-dihydroxyl-campest-4-en-3-one, **15**, the C23 hydroxylated product of compound **14** was detected, which agrees with the previous report (Figure 3.3A). The function of CYP90C1 or CYP90D1 was re-examined in YYL67 and YYL69 in the presence of MSBP1.The production of **15** was ~10-fold higher in YYL94 (integrated with *MSBP1*) than in YYL69 (Figure 3.3 A), which was barely detected in the absence of MSBP1. Interestingly, another new peak with a characteristic m/z^+ 447.3 and derivative m/z^+ 429.3, consistent with that of 2, 22, 23-trihydroxyl-campest-4-en-3-one, **16**,was detected only when CYP90C1 or CYP90D1 was expressed in YYL94 (Figure 3.3). The putative production of **16** suggests that CYP90C1 and CYP90D1 could be bifunctional hydroxylases. However, the enzymatic efficiency of CYP90C1 or CYP90D1 in catalyzing C2 hydroxylation is very low. It is hard to draw a conclusion that this C2 hydroxylation only happened when co-express with MSBP1 or C2 hydroxylated product is below the detection limit without MSBP1. However, the function of MSBP1 in enhancing the

enzymatic efficiency of CYP90A1, CYP90B1, CYP90C1, and CYP90D1 was confirmed, although the mechanism further investigation.

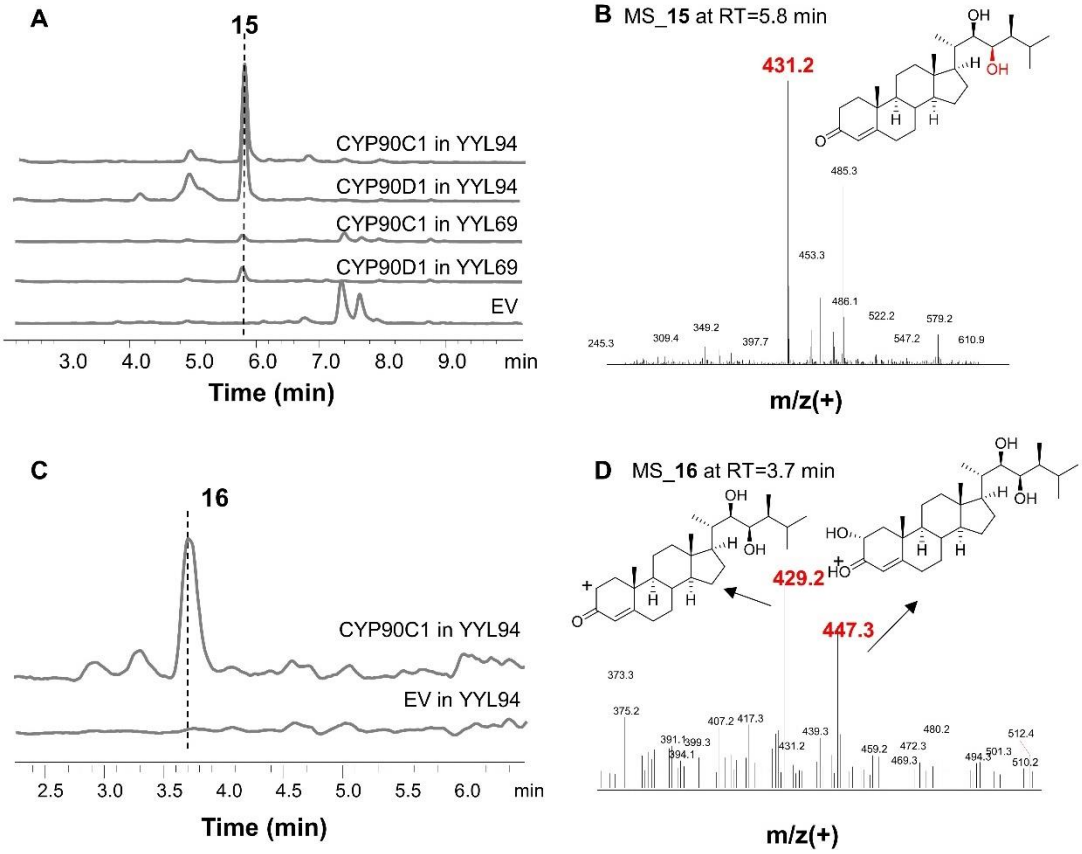


Figure 3.3 Characterization of 22, 23-dihydroxycampest-4-en-3-one, **15**. (A) SIM EIC using 22, 23-dihydroxy-campesterol, **15**'s characteristic m/z^+ signal (MW=430.67 Da, $[C_{28}H_{47}O_3]^+$=431.3) of YYL69 expressing CYP90C1 and CYP90D1, with or without co-expressing with MSBP1. (B) Mass spectrum of **15**. (C) SIM EIC using 22, 23-dihydroxy-campesterol, **16**'s characteristic m/z^+ signal (MW=446.67 Da, $[C_{28}H_{47}O_4]^+$=447.3) of YYL69 expressing CYP90C1 and CYP90D1, with or without co-expressing with MSBP1. (D) Mass spectrum of **16**.

Interaction Between MSBP1 and BR Biosynthetic Enzymes

To explore whether the role of MSBP1 in BR biosynthesis is related to enzyme complex formation as lignin biosynthesis, we examined the protein-protein interaction between MSBP1 and the BR biosynthetic enzymes using a mating-based split-ubiquitin yeast two-hybrid (mbSUS-Y2H) system[14]. In the mbSUS-Y2H system, the physical interactions of the bait and prey proteins can activate the reunion of split ubiquitin and enable the yeast strain harboring the prey and bait to grow on a 6-drop-out selective medium. The bait protein is expressed downstream of promoter Met (P_{Met}), the expression level of which can be suppressed and adjusted by methionine[14]. MSBP1 was used as the prey and BR biosynthetic enzymes (CYP90A1, CYP90B1, CYP90C1, CYP90D1, CYP85A1, CYP85A2) as the baits. BAK1 was confirmed to interact with MSBP1 in the previous research and was used as the positive control[13]. The empty vector of the bait construct was used as the negative control. The mated strain harboring BAK1 and MSBP1 grew well on the selective plate, while the negative control did not (Figure 3.4A). The strains harboring CYP90A1, CYP90B1 as baits, mating with MSBP1 harboring strain grew on the selective medium. The interaction between CYP90A1 and CYP90B1 was also detected, but not as strong as their interaction to MSBP1 (Figure 3.4B). This result suggested that MSBP1 is likely act as a scaffold protein to strengthen the interaction between CYP90A1 and CYP90B1 and forms a biosynthetic complex to enhance **14** production. However, no interaction between MSBP1 with CYP90C1 or CYP90D1 was detected (Figure 3.4). Moreover, the interaction between MSBP1 and the cytochrome P450 reductase ATR1 was detected. Then the interaction between CYP90C1/CYP90D1 and

ATR1 was examined as well (Figure 3.4B). Interestingly, CYP90C1 had a clear physical interaction with ATR1 while CYP90D1 not (Figure 3.4B). Additionally, YYL94 harboring CYP90D1 grew signfiicantly slower than YYL94 harboring CYP90C1, although the production of compounds **15** and **16** was at the similar level. So far, whether the improvement of enzyme activity of CYP90C1 or CYP90D1 is due to the increase of substrate (compound **14**) supply or due to the expression of MSBP1 as a scaffold protein remains mysterious and requires further investigation.

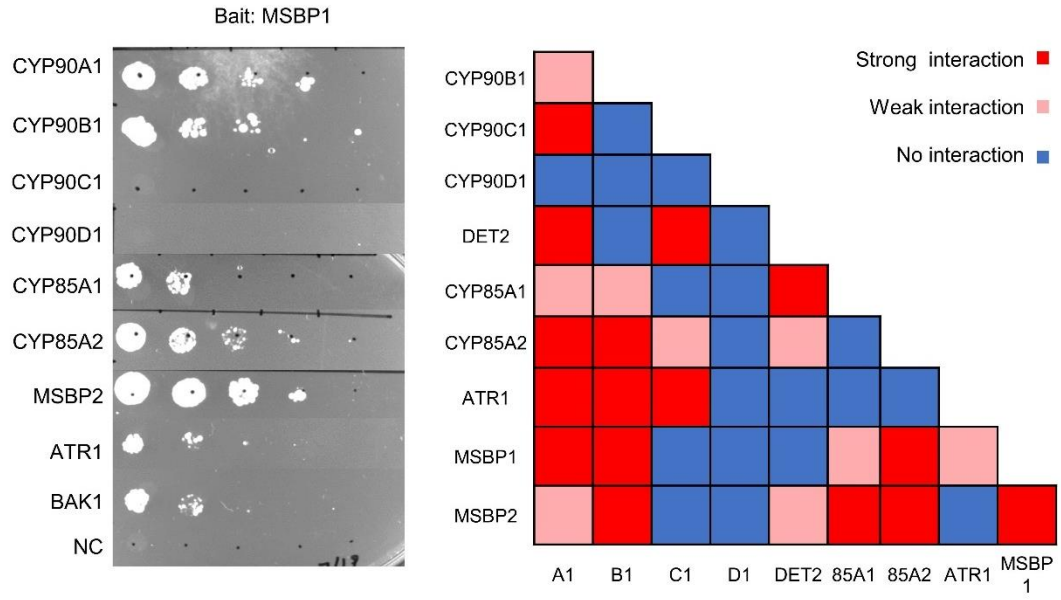


Figure 3.4 Physical interaction between MSBP1 and BR biosynthetic enzymes using Y2H system. (A) CYP90A1, CYP90B1, CYP90C1, CYP90D1, CYP85A1, CYP85A2, DET2, ATR1 were used as baits; MSBP1 were used as a prey. BAK1 with MSBP1 was used as the positive control. Empty vector (EV) was used as the negative control. The positive control and the negative control After serial dilution of 10^{-5} times (indicated at the top of the images), yeast diploid cells harbouring both bait and prey constructs grew on the growth (SC/−Trp (T), −Leu (L), −Ura (U)) and selective (SC/−Trp, −Leu, −Ade (A), −His (H), −Ura, −Met (M)) media, as indicated. Images were taken after incubation for 4 days at 30 °C. (B) The paired interaction map of BR biosynthetic enzymes and MSBP1.

Discussion

In this study, MSBP1, a non-enzymatic scaffold protein, was proved to involve in BR biosynthesis through the direct interaction with biosynthetic enzymes (CYP90A1, CYP90B1, CYP85A1, CYP85A2, ATR1) and enhance the bioconversion efficiency of the enzymes (CYP90A1, CYP90B1, CYP90C1, CYP90D1). MSBP1 plays multifaced roles in plants, also involved in lignin biosynthesis and BR signaling by interaction with BAK. It is very likely that MSBP1 still has more functions beyond our knowledge and worth further investigations.

DET2, a 5α reductase that is supposed to reduce the double carbon bond on substrates with C3-keto group (converted by CYP90A1) (Figure 3.1), was also reconstituted in our yeast platform. The expression of DET2 consumed almost all the substrate compound **14**, which is much more efficient than CYP90C1 to convert **14** (Figure S3.2). Unfortunately, the theoretical product of DET2, compound **17** (Figure 3.1), was not detected in our platform YYL94. The lack of authentic standard compound made it harder to characterize. However, it also led us thinking the possibility that DET2 is not a 5α reductase as reported before. Our collaborator, Dr. Takahito Nomura (Utsunomiya University in Japan), kindly shared their results on the minimal set of enzymes required for the synthesis of BR in *Nicotiana benthamiana*. However, expressing the minimal set of enzymes in yeast did not lead to the synthesis of the expected products, which further implies that BR biosynthesis likely requires a metabolon formed by catalyze enzymes and non-catalytic protein components, including but not limited to the MSBP1 that we discovered in this study.

Using mbSUS-Y2H system, we examined the interactions among the BR biosynthetic enzymes (DET2, CYP90A1, CYP90B1, CYP90C1, CYP90D1, CYP85A1, CYP85A2) (Figure 3.4B). Surprisingly, the results showed that CYP90D1 has no interaction with any other proteins, including ATR1 or MSBP1. In addition to the fact that *CYP90D1* is not present in the cDNA library of *Arabidopsis thaliana*, we strongly believe that *CYP90D1* is a redundant gene with overlapped function with CYP90C1 and is not an essential component in BR biosynthesis in *A. thaliana*. Other than CYP90D1, the rest of the five enzymes interacted with each other directly or indirectly: CYP90A1/CYP90C1 /DET2 interacted with each other pairwise; CYP90B1 interacted with CYP90A1; CYP85A2 interacted with all the other enzymes. To further understand BR biosynthesis in plant, we need to investigate whether BR biosynthetic enzymes form a complex and the components of the complex in plant, which will enable us to fully elucidate BR biosynthesis and reconstitute BR production in a microbial host, here, yeast.

Materials and Methods

Materials

Yeast strains were cultured at 30 ^{o}C in complex yeast extract peptone dextrose (YPD, all components from BD Diagnostics) medium or synthetic defined medium (SDM) containing yeast nitrogen base (YNB) (BD Diagnostics), ammonium sulfate (Fisher Scientific), 2% (w/v) glucose unless specified and the appropriate dropout (Takara Bio) solution for selection.

General Technique for DNA Manipulation

PCR reactions were performed with Expand high Fidelity system (Sigma-Aldrich), Phusion DNA Polymerase (NEB), Q5 High-Fidelity DNA Polymerase (New England Biolabs) or Taq Polymerase (NEB) according to manufacturer's protocols. PCR products were purified by Zymoclean Gel DNA Recovery Kit (Zymo Research). Plasmids were prepared with Econospin columns (Epoch Life Science) according to manufacturer's protocols. All DNA constructs were confirmed through DNA sequencing by Azenta life science Inc. BP Clonase II Enzyme Mix, Gateway pDONR221 Vector, LR Clonase II Enzyme Mix (Life Technologies), and the *S. cerevisiae* Advanced Gateway Destination Vector Kit [28] (Addgene) were used to perform the Gateway Cloning. The Gibson one-pot, isothermal DNA assembly[57] was conducted at the scale of 10 μL by incubating T5 exonuclease (NEB), Phusion polymerase (NEB), Taq ligase (NEB) and 50 ng of each DNA fragment at 50 ^{o}C for 1 hour to assemble multiple DNA fragments. Yeast strains are constructed through homologous recombination and DNA assembly[58]. Yeast strains, plasmids, and sequences of genes used in this work are listed in Table S3.1, Table S3.2, and Table S3.3, respectively.

Culture Conditions

For metabolite analysis, yeast strains were first cultured overnight in 500 μL SDM with 2% (w/v) glucose in 96-well plates (BD faulcon) at 30 ^{o}C overnight, at 250 r.p.m.. Appropriate volume of the overnight seed culture was inoculated in fresh 500 μL SDM with glucose to make OD_{600} around 0.1 and incubated at 30 ^{o}C for 72 hours before

metabolite analysis of the yeast pellets. To monitor the growth, yeast strains were grown in 2 mL SDM with 2% (w/v) glucose in test tubes. The yeast strains were first cultured overnight in 500 µL SDM with 2% (w/v) glucose. Appropriate volume of the overnight culture was inoculated in fresh 2 mL SDM with 2% (w/v) glucose to make OD_{600} around 0.1 and incubated at 30 ^{o}C for 72 hour, at 250 r.p.m.. Yeast growth was monitored through measuring OD_{600}: appropriate amount of yeast culture was collected into the cuvette and diluted with water as needed to make the reading within the linear range of the instrument (usually from OD_{600} 0.1 to 0.8). OD_{600} of the culture was measured by Nanodrop (Molecular Device^{TM}).

Analysis and Quantification of Phytosterols in Yeast

Brassinoteroid intermediates were collected and analyzed from yeast pellets. 100 µL acetone was added to yeast cell pellets (from 500 µL yeast culture), followed by the addition of 400 µL methanol with vigorous vortex to mix. The supernatant was transferred to new tubes. Repeat the extraction of the pellets by adding an additional 400 µL ethyl acetate. The supernatants were combined and dried in the vacufuge (Eppendorf). The dried samples were dissolved in 40 µL methanol and analyzed using LC-MS.

The samples were then analyzed by reverse phase LC-MS on a Shimadzu 2020 single quadrupole LC-MS (Poroshell 120, EC-C18, 3.0 × 100 mm, 2.7 µm) using positive ionization. To analyze sterols, the metabolites were separated under the linear gradient started from 80% methanol (v/v in water, 0.1% formic acid) to 100% methanol (v/v in water, 0.1% formic acid) for 8 min followed by 100% methanol (v/v in water, 0.1% formic

acid) for 12 min at the flow rate of 0.5 ml/min. Data was acquired by the electrospray ionization mode. The sterols were analyzed using selected ion monitoring (SIM) with either the characteristic or possible mass-to-charge ratio (m/z^+) according to the structures and molecular weights. The quantities of campesterol were estimated by comparing to the standard curves using authentic campesterol standard solutions.

Mating-Based Split-Ubiquitin Yeast Two-Hybrid Assay

For the yeast two-hybrid assay, we followed a standard protocol for the mating-based split-ubiquitin system14. Brassinosteroid biosynthetic genes were pDONR221 entry vector (Invitrogen) without the stop codons. These genes were cloned into the pMetYC–DEST vector to generate the bait constructs, and into the pXN22–DEST vector to generate prey constructs, using LR assembly reaction. The bait and prey constructs were transformed into THY.AP4 and THY.AP5 yeast strains and selected on −Leu and −Trp plates, respectively. The mating was conducted by mixing the concentrated culture of prey and bait in YPD medium then drop 2 µL on YPD agar plate. After overnight incubation at 30 oC, the diploid cells harbouring both constructs were selected on −Trp, −Leu, −Ura plates and let the diploid grew for 16 hours or overnight. Positive cells were serially diluted and plated on the selective medium −Trp, −Leu, −His, −Ura, −Met, −Ade.

Reference

1. De Bruyne, L.; Hofte, M.; De Vleesschauwer, D., Connecting growth and defense: the emerging roles of brassinosteroids and gibberellins in plant innate immunity. *Molecular plant* **2014,** *7* (6), 943-59.

2. Noguchi, T.; Fujioka, S.; Choe, S.; Takatsuto, S.; Yoshida, S.; Yuan, H.; Feldmann, K. A.; Tax, F. E., Brassinosteroid-insensitive dwarf mutants of Arabidopsis accumulate brassinosteroids. *Plant physiology* **1999,** *121* (3), 743-52.

3. Khripach, V.; Zhabinskii, V.; De Groot, A., Twenty years of brassinosteroids: Steroidal plant hormones warrant better crops for the XXI century. *Ann Bot-London* **2000,** *86* (3), 441-447.

4. Eremina, M.; Unterholzner, S. J.; Rathnayake, A. I.; Castellanos, M.; Khan, M.; Kugler, K. G.; May, S. T.; Mayer, K. F.; Rozhon, W.; Poppenberger, B., Brassinosteroids participate in the control of basal and acquired freezing tolerance of plants. *Proceedings of the National Academy of Sciences of the United States of America* **2016,** *113* (40), E5982-E5991.

5. Northey, J. G.; Liang, S.; Jamshed, M.; Deb, S.; Foo, E.; Reid, J. B.; McCourt, P.; Samuel, M. A., Farnesylation mediates brassinosteroid biosynthesis to regulate abscisic acid responses. *Nature plants* **2016,** *2*, 16114.

6. Romanutti, C.; Castilla, V.; Coto, C. E.; Wachsman, M. B., Antiviral effect of a synthetic brassinosteroid on the replication of vesicular stomatitis virus in Vero cells. *Int J Antimicrob Ag* **2007,** *29* (3), 311-316.

7. Liu, J. N.; Zhang, D.; Sun, X. Y.; Ding, T. L.; Lei, B. L.; Zhang, C. L., Structure-activity relationship of brassinosteroids and their agricultural practical usages. *Steroids* **2017,** *124*, 1-17.

8. Ohnishi, T.; Szatmari, A. M.; Watanabe, B.; Fujita, S.; Bancos, S.; Koncz, C.; Lafos, M.; Shibata, K.; Yokota, T.; Sakata, K.; Szekeres, M.; Mizutani, M., C-23 hydroxylation by Arabidopsis CYP90C1 and CYP90D1 reveals a novel shortcut in brassinosteroid biosynthesis. *The Plant cell* **2006,** *18* (11), 3275-88.

9. Kim, T. W.; Hwang, J. Y.; Kim, Y. S.; Joo, S. H.; Chang, S. C.; Lee, J. S.; Takatsuto, S.; Kim, S. K., Arabidopsis CYP85A2, a cytochrome P450, mediates the Baeyer-Villiger oxidation of castasterone to brassinolide in brassinosteroid biosynthesis. *The Plant cell* **2005,** *17* (8), 2397-412.

10. Nomura, T.; Kushiro, T.; Yokota, T.; Kamiya, Y.; Bishop, G. J.; Yamaguchi, S., The last reaction producing brassinolide is catalyzed by cytochrome P-450s, CYP85A3 in tomato and CYP85A2 in Arabidopsis. *The Journal of biological chemistry* **2005,** *280* (18), 17873-9.

11. Gou, M.; Ran, X.; Martin, D. W.; Liu, C. J., The scaffold proteins of lignin biosynthetic cytochrome P450 enzymes. *Nature plants* **2018,** *4* (5), 299-310.

12. Yang, X. H.; Xu, Z. H.; Xue, H. W., Arabidopsis membrane steroid binding protein 1 is involved in inhibition of cell elongation. *The Plant cell* **2005,** *17* (1), 116-31.

13. Song, L.; Shi, Q. M.; Yang, X. H.; Xu, Z. H.; Xue, H. W., Membrane steroid-binding protein 1 (MSBP1) negatively regulates brassinosteroid signaling by enhancing the endocytosis of BAK1. *Cell Res* **2009,** *19* (7), 864-76.

14. Grefen, C.; Lalonde, S.; Obrdlik, P., Split-ubiquitin system for identifying protein-protein interactions in membrane and full-length proteins. *Curr Protoc Neurosci* **2007,** *Chapter 5*, Unit 5 27.

Appendix I

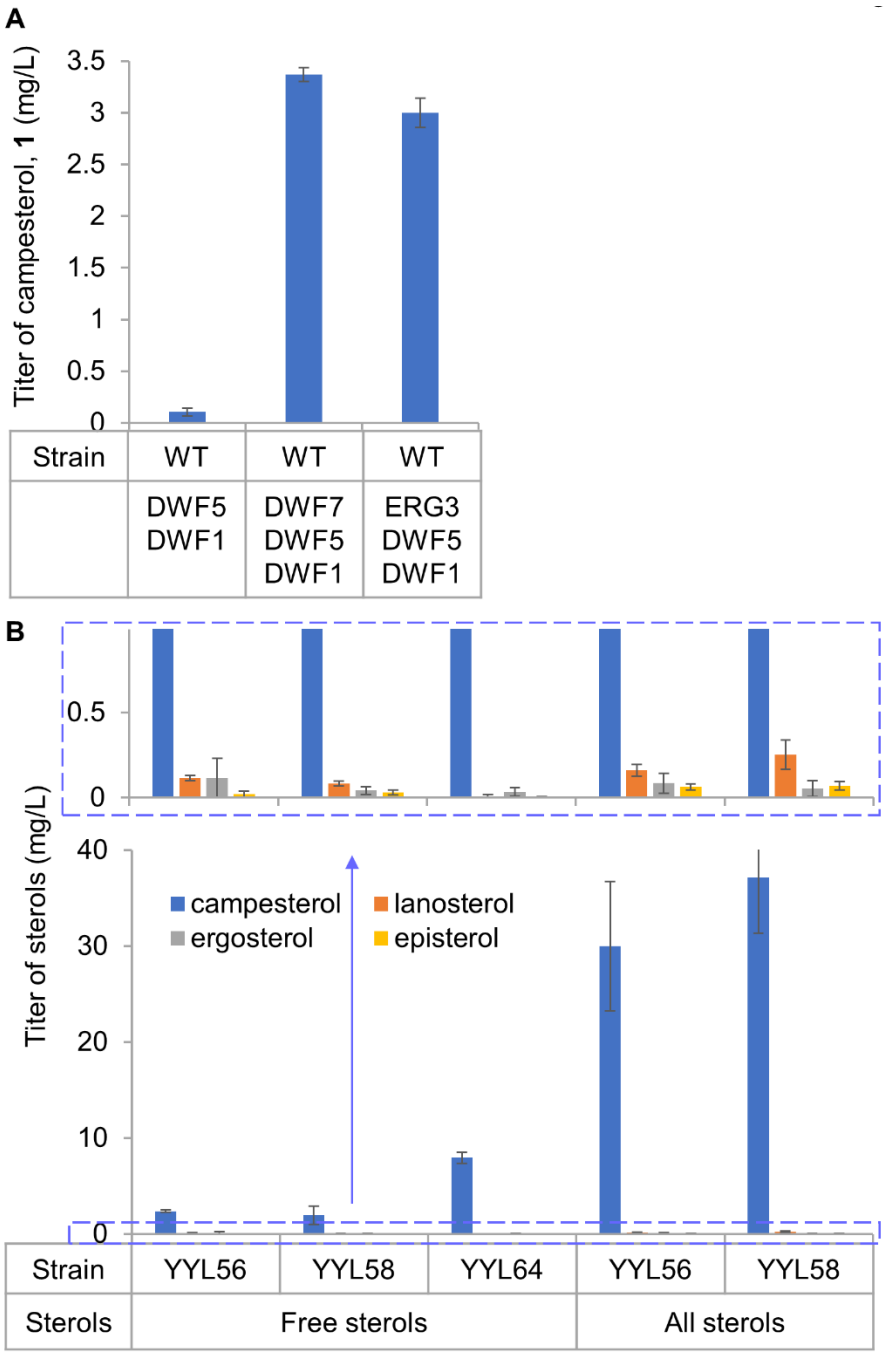

A
Titer of campesterol, 1 (mg/L)
3.5
3
2.5
2
1.5
1
0.5
0
Strain
WT
WT
WT
DWF5 DWF1
DWF7 DWF5 DWF1
ERG3 DWF5 DWF1
B
Titer of sterols (mg/L)
40
30
20
10
0
campesterol
lanosterol
ergosterol
episterol
Strain
YYL56
YYL58
YYL64
YYL56
YYL58
Sterols
Free sterols
All sterols

Figure S1.1. (A) campesterol production in yeast strains expressing DWF5/DWF1, DWF7/DWF5/DWF1, and ERG3/DWF5/DWF1 from low-copy number plasmids. WT stands for wild type yeast, as CENPK2. 1D. (C) Sterol profiling of YYL56, YYL58 and YYL64 (YYL56 with *are1*Δ*are2*Δ). Lanosterol, episterol, ergosterol and campesterol was quantified in free sterols and total sterols including sterol esters, by using unsaponified and saponified extraction methods. Bars represent mean values of three biological replicates, and the error bars represent the standard deviation of the replicates. The quantification of each sterol was measured by comparing the integrated peak area to a standard curve of the authentic standards of the corresponding sterols.

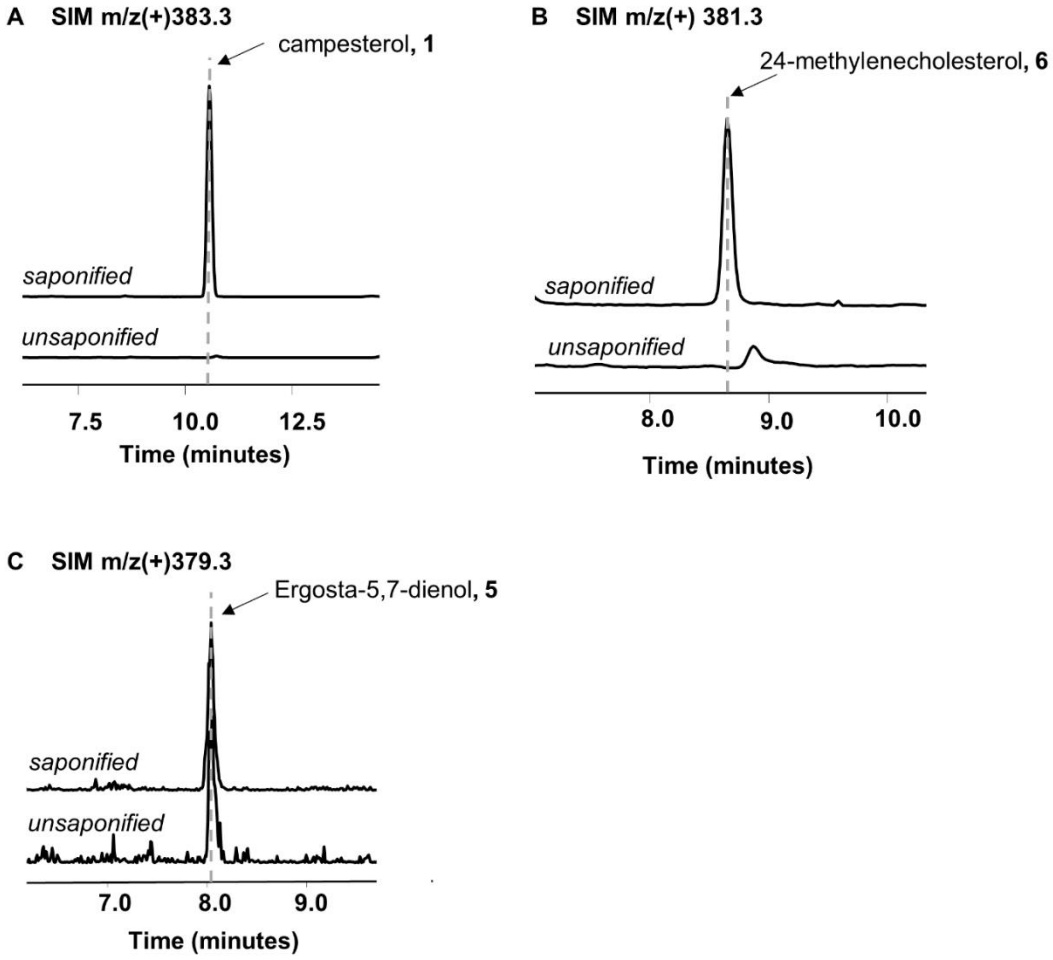


Figure S1.2. LCMS traces of campesterol, **1**, 24-methylenencholesterol, **6**, and ergosta-5,7-dienol, **5** in saponified and unsaponified extraction of YYL58. (A) EIC of campesterol characteristic m/z+ =383.3 of the saponified and unsaponified sterol extract of YYL58. (B) EIC of compound **6** characteristic m/z+ =381.3 of the saponified and unsaponified sterol extract of YYL58. (C) EIC of compound **5** characteristic m/z+ =379.3 of the saponified and unsaponified sterol extract of YYL60. All traces are representative of at least three biological replicates for each engineered yeast strain.

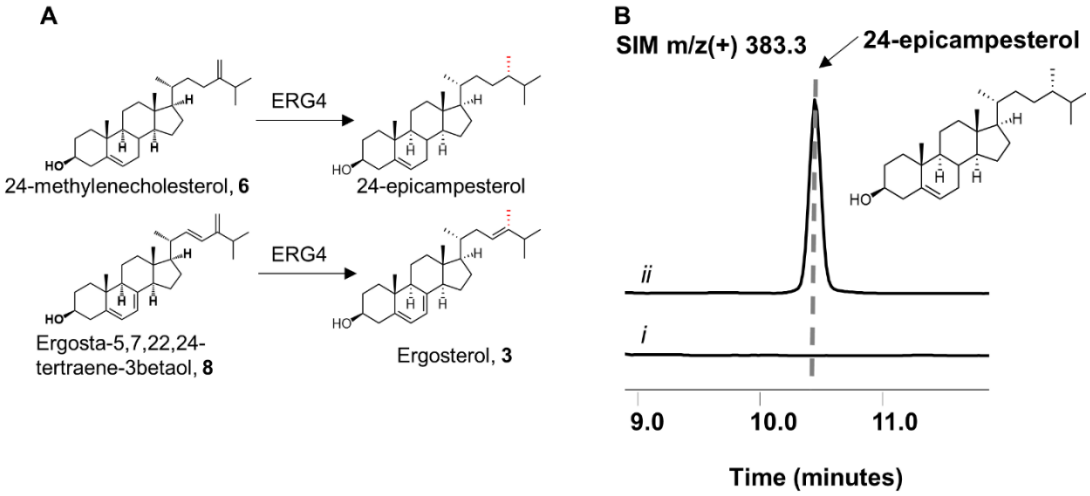


Figure S1.3. (A) ERG4 is likely a 24(*S*)-reductase functioning on sterols with 24-methylene group. (B) SIM EIC of characteristic campesterol or 24-epicampesterol characteristic m/z^{+}=383.3 of i) YYL66 (*are1*Δ*are2*Δ*erg4*Δ) expressing DWF5, ii) YYL63 (*are1*Δ*are2*Δ) expressing DWF5. Traces are representative of at least three biological replicates for each engineered yeast strain.

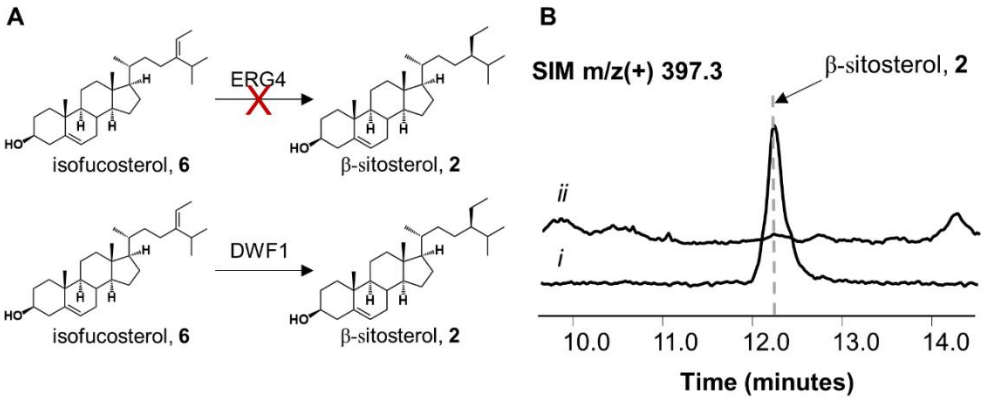


Figure S1.4. (A) DWF1 converts isofucosterol, **6**, to β-sitosterol, **2**. (B) EIC SIM using β-sitosterol, **2**, characteristic m/z$^+$ signal (MW=414.71, $[C_{29}H_{45}]^+$=397.4) in i) authentic standard of β-sitosterol, **2**, ii) YYL63 (*are1*Δ*are2*Δ) expressing DWF7, DWF5, and SMT2 from low-copy number plasmids. The traces and spectrums are representative of at least three biological replicates for each engineered yeast strain.

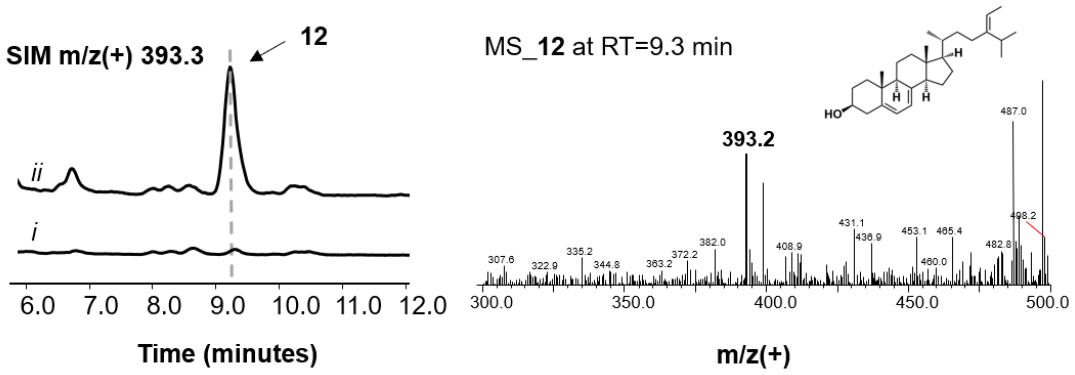


Figure S1.5. EIC SIM using Δ5,7-Avenasterol, **12**, characteristic m/z^+ signal (MW=410.69, $[C_{29}H_{45}]^+$=393.3) in i) **5**-accumulated strain YYL60, ii) YYL60 (*erg5*Δ*::dhcr7*) expressing SMT2 from a low-copy plasmid. The mass spectrum of the peak at retention time (RT) = 9.3 minutes is shown on the right. The traces and spectrums are representative of at least three biological replicates for each engineered yeast strain.

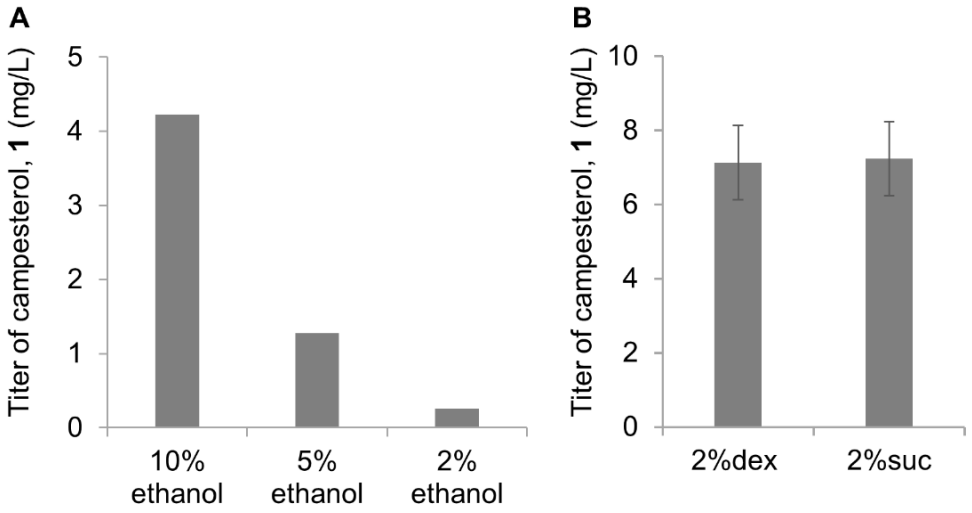


Figure S1.6. (A) Production of campesterol from YYL65 in medium with different concentrations of ethanol. YYL65 was cultured in SDM medium with 2% sucrose, adding 10%, 5% and 2% ethanol. (B) Production of campesterol from YYL67 when using 2% glucose (dex) or sucrose (suc) as carbon source in SDM. Bars represent mean values of three biological replicates, and the error bars represent the standard deviation of the replicates.

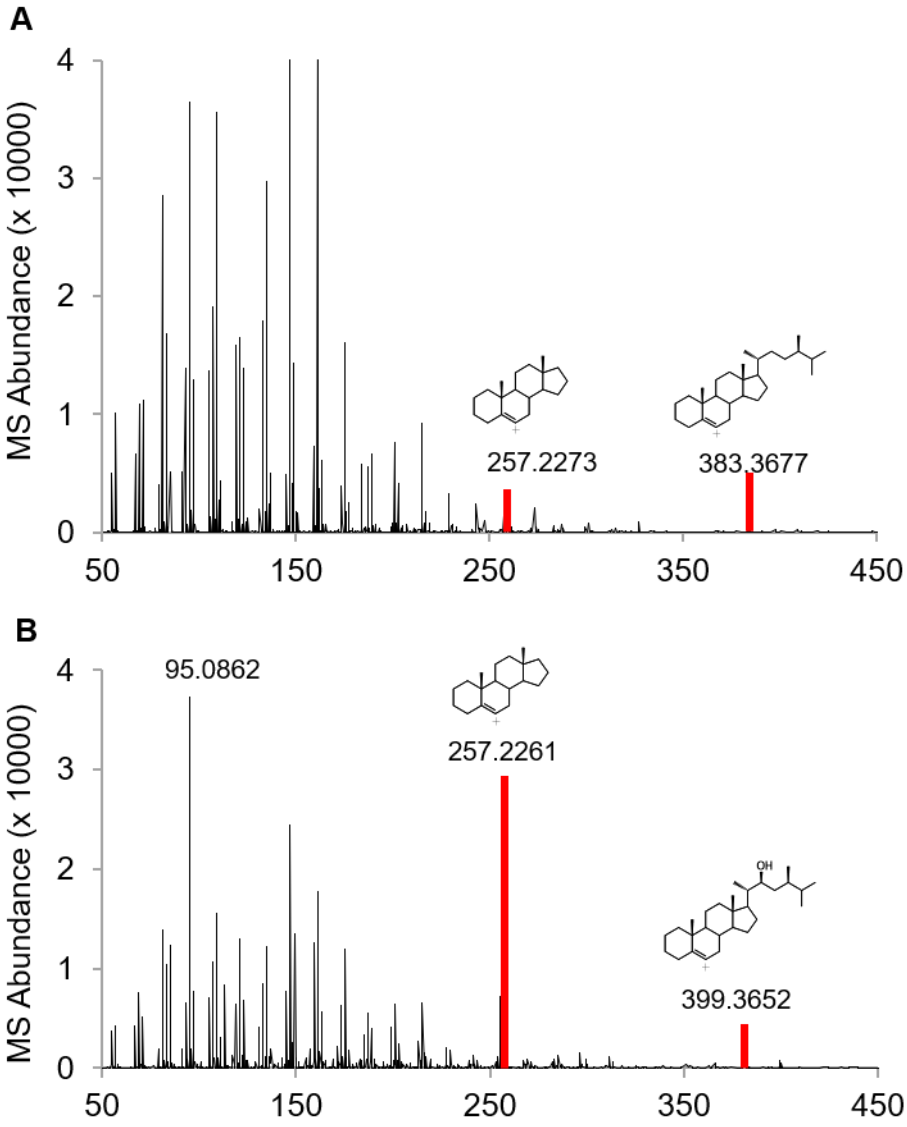


Figure S1.7. Mass spectrum of **(A)** campesterol, **1**, **(B)** 22-hydroxycampesterol, **13**. The spectrums are representative of at least three biological replicates for each engineered yeast strain.

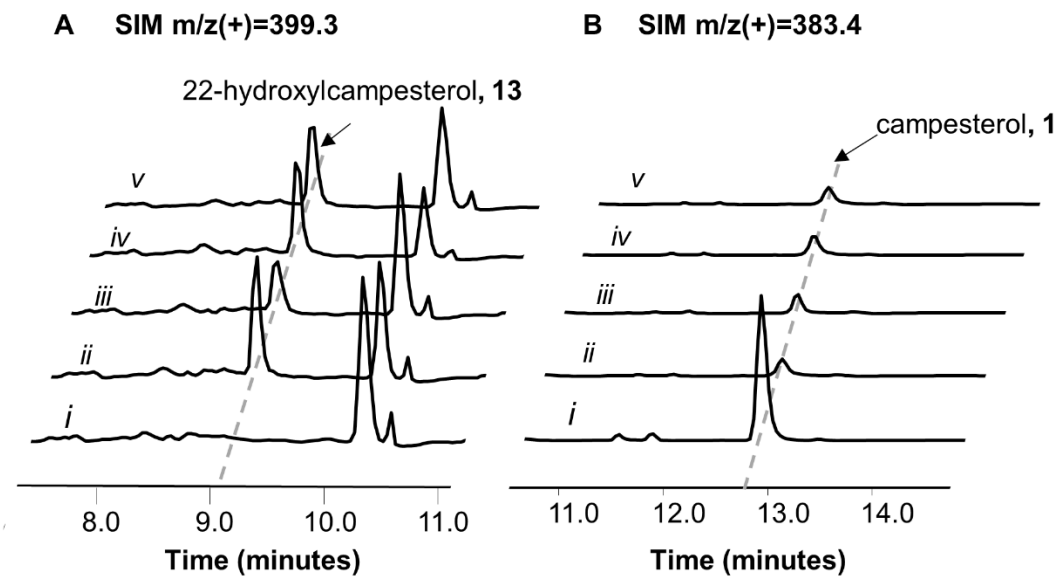
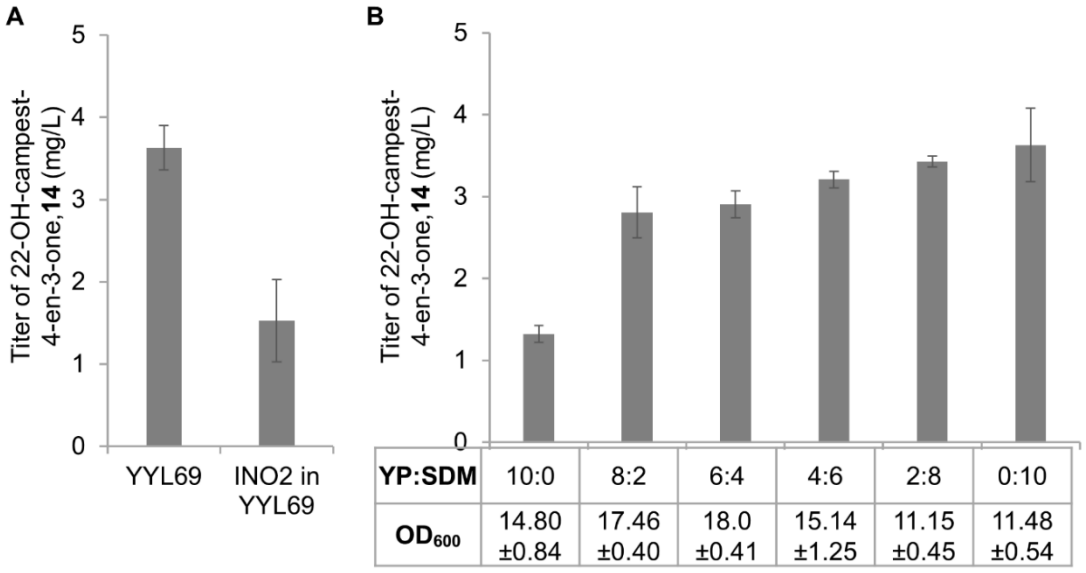


Figure S1.8. Characterization the activities of CYP90B1 variants and mutants. **(A)** EIC SIM using 22-hydroxycampesterol, **13**, characteristic m/z^+ signal (MW=416.69, $[C29H45]^+$=399.4) in YYL67 expressing i) empty vector, ii) CYP90B1, iii) CYP724B2, iv) CYP90B3, v) t30CYP90B1. **(B)** EIC SIM using campesterol's characteristic m/z^+ signal (MW=400.69, $[C29H45]^+$=383.4) in YYL67 expressing i) empty vector, ii) CYP90B1, iii) CYP724B2, iv) CYP90B3, v)t30CYP90B1. YYL67 harboring plasmids with CYP90B1 variants and mutants (pYL655-pYL663) were grown in selective SDM supplemented with 2% glucose at 30°C for 72 hours before analysis. Only the enzymes that converted campesterol were presented, and the full list of CYP90B1 variants tested are listed in Table S3.



YP:SDM	10:0	8:2	6:4	4:6	2:8	0:10
OD_{600}	14.80 ±0.84	17.46 ±0.40	18.0 ±0.41	15.14 ±1.25	11.15 ±0.45	11.48 ±0.54

Figure S1.9. Optimization of the production of **14**. **(A)**YYL69 harboring plasmid P_{GPD}-INO2 and YYL69. Yeast strains were grown in synthetic complete SDM medium at 30°C for 72 hours. **(B)** Titration of YP to SDM ratio with 2% (w/v) glucose. Bars represent mean values of three biological replicates, and the error bars represent the standard deviation of the replicates.

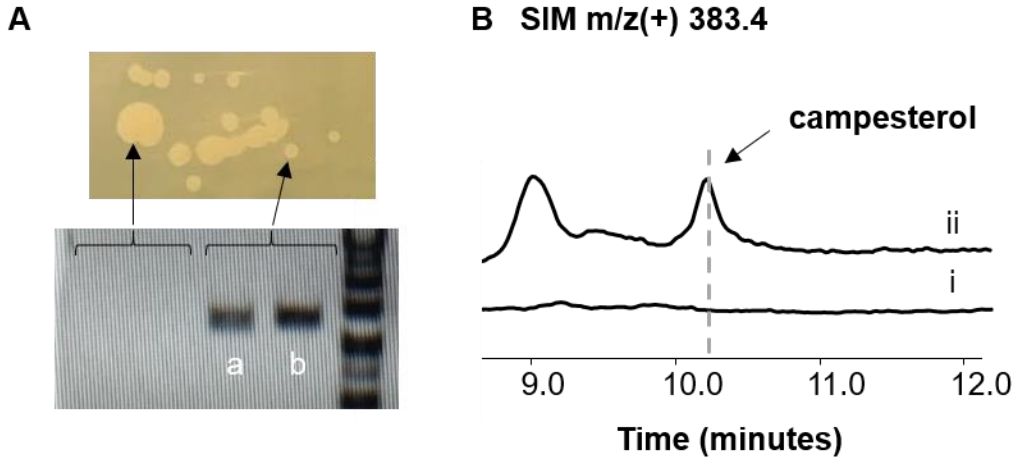


Figure S1.10. Identification of DWF5 in YYL65 big colony (BC) mutant. (A) colony PCR of DWF5 integration boundaries in YYL65 and YYL65-BC mutant indicating the loss of *dwf5*. a and b indicated the left and right boundaries of *dwf5*. (B) EIC SIM using campesterol's characteristic m/z^+ signal (MW=400.69, $[C29H45]^+$=383.4) in i) YYL65-BC and ii) YYL65-BC expressing DWF5.

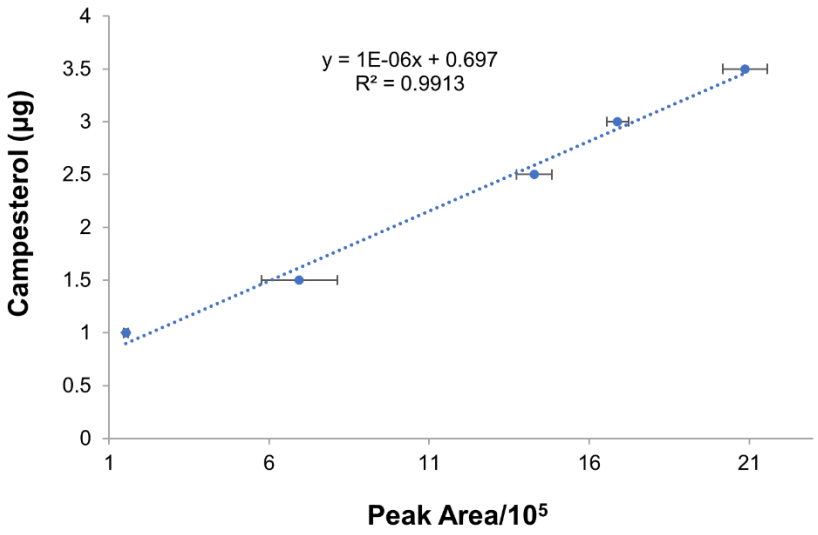


Figure S1.11. Standard curve of campesterol. Standard curve was plot by the mass of campesterol against peak area of campesterol using its characteristic m/z^+=383.3. The R^2=0.9931 shows the reliability of this standard curve. The peak areas of tested samples were located in the range of this standard curve. The peak area is calculated by Lab Solution, the software coming with the instrument Shimadzu LCMS 2020.

Table S1.1. List of sterol C22-hydrolases tested in YYL67.

Enzymes	Function	Strain	Campesterol as the substrate	Codon optimized	Ref
CYP90B3	C-22 hydroxylase	*Solanum lycopersicum*	Y	N	1
CYP724B2	C-22 hydroxylase	*Solanum lycopersicum*	Y	N	1
ADT1	arogenate dehydratase	*Arabidopsis thaliana*	N	Y	2
CYP90B27	C-22 hydroxylase	*Veratrum californicum*	N	Y	3
CYP302A1	C-22 hydroxylase	*Drosophila melanogaster*	N	Y	4
CYP718	unknown, putatively related to BR metabolism	*Arabidopsis thaliana*	N	Y	-
CYP716A1	C-28 hydroxylase	*Arabidopsis thaliana*	N	Y	5
CYP716A2	C-22 hydroxylase	*Arabidopsis thaliana*	N	Y	5
CYP722A1	unknown, putatively related to BR metabolism	*Arabidopsis thaliana*	N	Y	-

"Y" represents "yes". "N" represents "no".

Table S1.2. Plasmids used in this study

Plasmids #	Genotype	Reference
pAG413GAL-ccdB	*Centromeric HIS3, attR-P_{GAL}-ccdB-attR*	6
pAG414GAL-ccdB	*Centromeric TRP1, attR1-P_{GAL}-ccdB-attR2*	6
pAG415GAL-ccdB	*Centromeric LEU2, attR1-P_{GAL}-ccdB-attR2*	6
pAG416GAL-ccdB	*Centromeric URA3, attR1-P_{GAL}-ccdB-attR2*	6
pAG413GPD-ccdB	*Centromeric HIS3, attR1-P_{GPD}-ccdB-attR2*	6
pAG414GPD-ccdB	*Centromeric TRP1, attR1-P_{GPD}-ccdB-attR2*	6
pAG415GPD-ccdB	*Centromeric LEU2, attR1-P_{GPD}-ccdB-attR2*	6
pAG416GPD-ccdB	*Centromeric URA3, attR1-P_{GPD}-ccdB-attR2*	7
pYL21	*Centromeric HIS3, attR1-P_{HXT7}-ccdB-attR2*	7
pYL22	*Centromeric TRP1, attR1-P_{HXT7}-ccdB-attR2*	7
pYL20	*Centromeric URA3, attR1-P_{HXT7}-ccdB-attR2*	7
pYL188	*attL1-DWF1-attL2*	This work
pYL189	*attL1-DWF5-attL2*	This work
pYL190	*attL1-DWF7-attL2*	This work
pYL205	*attL1-SMT2-attL2*	This work
pYL593	*attL1-CYP90A1-attL2*	This work
pYL184	*attL1-CYP90B1-attL2*	This work
pYL595	*attL1-DET2-attL2*	This work
pYL237	*attL1-CYP90D1-attL2*	This work
pYL594	*attL1- CYP90C1-attL2*	This work
pYL165	*attL1- AtATR1-attL2*	This work
pYL223	*attL1-P_{TEF1}-ERG10-T_{CYC1}-attL2*	This work
pYL224	*attL1-P_{TEF1}-ERG19-T_{CYC1}-attL2*	This work
pYL225	*attL1-P_{TEF1}-ERG20-T_{CYC1}-attL2*	This work
pYL216	*attL1-P_{PGK1}-ERG12-T_{pho5}-attL2*	This work
pYL215	*attL1-P_{PGK1}-DWF5-T_{pho5}-attL2*	This work
pYL652	*attL1-P_{PGK1}-DET2-T_{pho5}-attL2*	This work
pYL226	*attL1-P_{TPI1}-ERG13-T_{Ste2}-attL2*	This work
pYL654	*attL1-P_{TPI1}-CYP90A1-T_{Ste2}-attL2*	This work
pYL228	*attL1-PTPI1-AtATR1-TSte2-attL2*	This work
pYL227	*attL1-PTPI1-ERG8-TSte2-attL2*	This work
pYL211	*attL1-P_{GPD}-IDI1-T_{ADH1}-attL2*	This work
pYL213	*attL1-P_{GPD}-UPC2-T_{ADH1}-attL2*	This work
pYL219	*attL1-P_{PYK1}-tHMG1-T_{MFa1}-attL2*	This work
pYL218	*attL1-P_{PYK1}-DWF1-T_{MFa1}-attL2*	This work

P, promoter; T, terminator

Table S1.3. DNA sequences of genes used in this study

Name	sequences
CYP90A1	ATGGCTTTCACTGCTTTCTTGCTGTTGTTGTCATCTATTGCTGCTGGTTTCTTGTTGCTGTTGAGA AGAACTAGATACAGAAGAATGGGTTTGCCACCAGGTTCTTTGGGTTTACCATTGATTGGTGAAAC CTTCCAATTGATCGGTGCTTACAAAACTGAAAACCCAGAACCATTCATTGACGAAAGAGTTGCTA GATACGGTTCTGTTTTCATGACTCATTTGTTCGGTGAACCTACCATTTTTTCTGCTGATCCAGAAA CTAACAGATTCGTCTTGCAAAATGAAGGCAAGTTGTTCGAATGTTCTTACCCAGCTTCTATCTGTA ACTTGTTGGGTAAACACTCCCTGTTGTTGATGAAGGGTTCATTGCATAAGAGGATGCATTCTTTG ACTATGTCCTTCGCTAACTCCTCCATTATCAAGGATCATTTGATGTTGGACATCGACAGATTGGT CAGATTCAATTTGGATTCCTGGTCCTCTAGAGTCTTGTTGATGGAAGAAGCTAAGAAGATCACTT TCGAATTGACCGTCAAGCAGTTGATGTCTTTTGATCCAGGTGAATGGTCTGAGTCCTTGAGAAAA GAATACTTGTTGGTCATCGAGGGTTTCTTCTCATTGCCATTGCCTTTGTTTTCTACCACTTACAGA AAGGCTATTCAGGCCAGAAGAAAAGTTGCTGAAGCTTTGACAGTTGTCGTCATGAAGAGAAGAG AAGAAGAGGAAGAAGGCGCTGAAAGAAAAAAAGATATGTTGGCTGCTTTGTTGGCAGCTGATGA TGGTTTTTCTGATGAAGAAATCGTTGATTTCTTGGTCGCCTTGTTGGTTGCTGGTTACGAAACTAC TTCTACCATTATGACTTTGGCCGTTAAGTTCTTGACTGAAACTCCATTGGCTTTGGCCCAATTGAA AGAAGAACACGAAAAAATCAGAGCCATGAAGTCCGATTCTTACTCTTTGGAATGGTCCGATTACA AGTCTATGCCATTCACTCAATGCGTTGTTAACGAAACTTTGAGAGTTGCCAAACATTATCGGTGGT GTTTTTAGAAGGGCTATGACCGATGTTGAAATCAAGGGTTACAAGATTCCAAAAGGCTGGAAGGT CTTTTCATCTTTCAGAGCTGTTCATTTGGACCCAAATCATTTCAAGGATGCCAGAACTTTTAATCC TTGGAGATGGCAATCTAACTCTGTTACTACTGGTCCATCTAACGTTTTTACTCCATTTGGTGGTG GTCCAAGATTGTGTCCAGGTTATGAATTGGCTAGAGTTGCTTTGTCTGTTTTCTTGCATAGATTG GTTACTGGTTTCTCTTGGGTTCCAGCTGAACAAGATAAGTTGGTATTTTTCCCAACCACTAGAAC CCAAAAGAGATACCCAATTTTCGTTAAGAGAAGGGATTTCGCTACCTAA
CYP90B1	ATGTTCGAGACCGAGCACCACACGCTGCTGCCCCTGCTGCTGCTCCCGAGCCTGCTGTCGCTG CTCCTGTTCCTGATCCTGCTGAAGCGCCGCAACCGGCGCACCCGCTTCAACCTGCCGCCGGGC AAGTCCGGCTGGCCGTTCCTGGGCGAGACCATCGGCTACCTGAAGCCGTACACCGCCACCACG CTGGGCGACTTCATGCAGCAGCACGTGTCCAAGTACGGCAAGATCTACCGCTCCAACCTGTTCG GCGAGCCGACCATCGTGTCCGCCGACGCCGGCCTGAACCGCTTCATCCTCCAGAACGAGGGC CGCCTGTTCGAGTGCTCGTACCCGCGCTCCATCGGCGGCATCCTCGGCAAGTGGTCCATGCTG GTCCTGGTCGGCGACATGCACCGCGACATGCGCTCGATCTCCCTGAACTTCCTGTCGCACGCC CGCCTGCGCACCATCCTGCTCAAGGACGTCGAGCGGCACACCCTGTTCGTCCTGGACTCCTGG CAGCAGAACTCCATCTTCTCGGCCCAGGACGAGGCCAAGAAGTTCACCTTCAACCTCATGGCCA AGCACATCATGTCGATGGACCCCGGCGAGGAGGAGACCGAGCAGCTGAAGAAGGAGTACGTCA CGTTCATGAAGGGCGTCGTCAGCGCCCCGCTGAACCTGCCGGGCACCGCCTACCACAAGGCC CTGCAGTCCCGCGCCACCATCCTCAAGTTCATCGAGCGCAAGATGGAGGAGCGGAAGCTGGAC ATCAAGGAGGAGGACCAGGAGGAGGAGGAGGTCAAGACCGAGGACGAGGCGGAGATGTCCAA GTCCGACCACGTCCGCAAGCAGCGCACCGACGACGACCTGCTCGGCTGGGTCCTGAAGCACA GCAACCTGTCCACCGAGCAGATCCTGGACCTGATCCTCTCGCTGCTGTTCGCCGGCCACGAGA CCTCGTCCGTCGCGATCGCCCTGGCGATCTTCTTCCTGCAGGCGTGCCCGAAGGCCGTCGAGG AGCTGCGCGAGGAGCACCTCGAGATCGCCCGCGCCAAGAAGGAGCTGGGCGAGTCCGAGCTG AACTGGGACGACTACAAGAAGATGGACTTCACCCAGTGCGTCATCAACGAGACCCTGCGCCTG GGCAACGTCGTCCGCTTCCTGCACCGGAAGGCCCTGAAGGACGTCCGCTACAAGGGCTACGAC ATCCCCTCCGGCTGGAAGGTCCTGCCGGTGATCTCCGCCGTCCACCTGGACAACTCCCGCTAC GACCAGCCGAACCTGTTCAACCCGTGGCGGTGGCAGCAGCAGAACAACGGCGCCTCCTCGTCC GGCTCCGGCTCCTTCTCCACCTGGGGCAACAACTACATGCCGTTCGGCGGCGGCCCGCGCCTG TGCGCCGGCTCCGAGCTGGCGAAGCTCGAGATGGCCGTGTTCATCCACCACCTGGTGCTGAAG TTCAACTGGGAGCTGGCGGAGGACGACAAGCCGTTCGCGTTCCCGTTCGTCGACTTCCCGAAC GGCCTGCCGATCCGCGTGTCCCGCATCCTGTGA
CYP90C1	ATGCAACCACCAGCTTCAGCTGGTTTGTTTAGATCTCCAGAAAATTTGCCATGGCCATACAACTA CATGGATTATTTGGTTGCTGGTTTCTTGGTTTTGACCGCCGGTATTTTGTTAAGACCATGGTTGT GGTTGAGACTGAGAAACTCTAAAACTAAGGATGGTGATGAAGAAGAGGACAACGAAGAAAAGAA AAAGGGTATGATTCCAAACGGTTCTTTAGGTTGGCCAGTTATTGGTGAAACCTTGAACTTTATTG CTTGCGGTTATTCCTCTAGACCAGTTACTTTTATGGACAAGAGGAAGTCCTTGTACGGTAAGGTT TTCAAGACCAACATTATCGGTACTCCCATTATCATTTCTACCGATGCTGAAGTTAACAAGGTCGTC TTGCAAAATCATGGTAACACTTTTGTTCCAGCTTACCCAAAGTCTATTACTGAGTTGTTGGGTGAG AACTCCATCTTGTCTATTAACGGTCCACATCAAAAGAGATTGCATACCTTGATTGGTGCCTTTTTG AGATCTCCACATTTGAAGGATAGAATCACCAGAGATATTGAAGCCTCTGTTGTTTTGACTTTGGC TTCTTGGGCTCAATTGCCATTGGTTCATGTTCAAGACGAAATCAAGAAGATGACCTTCGAAATCT TGGTCAAGGTTTTGATGTCTACTTCTCCAGGTGAAGATATGAACATCTTGAAGTTGGAATTCGAA GAGTTCATCAAGGGTTTGATCTGCATTCCAATCAAGTTTCCAGGTACTAGGTTGTACAAATCCTT

	GAAGGCTAAAGAAAGGCTGATCAAGATGGTTAAGAAGGTGGTTGAAGAAAGACAAGTTGCTATG ACTACTACCTCTCCAGCTAATGATGTTGTTGATGTCTTGTTGAGAGATGGTGGTGATTCCGAAAA ACAATCTCAACCATCTGATTTCGTGTCCGGTAAAATCGTTGAAATGATGATTCCTGGTGAAGAAA CTATGCCAACTGCTATGACATTGGCTGTTAAGTTCTTGTCTGATAACCCAGTTGCTTTGGCTAAG TTGGTAGAAGAAAACATGGAAATGAAGAGGCGTAAATTGGAATTGGGTGAAGAGTACAAATGGA CCGATTACATGTCTTTGTCCTTCACTCAAAACGTCATCAACGAAACTTTGAGAATGGCCAATATCA TCAACGGTGTTTGGAGAAAGGCTTTGAAAGACGTTGAAATCAAGGGCTACTTGATTCCTAAAGGT TGGTGTGTTTTGGCCTCTTTCATTAGCGTTCATATGGACGAAGATATCTACGACAATCCATACCA ATTTGATCCTTGGAGATGGGACAGAATCAATGGTTCTGCTAACTCTTCTATTTGCTTCACTCCATT TGGTGGTGGTCAAAGATTGTGTCCAGGTTTGGAATTGTCCAAGTTGGAAATCTCCATTTTCTTGC ATCACTTGGTCACTAGATATTCTTGGACTGCTGAAGAGGACGAAATCGTTTCTTTTCCAACTGTC AAAATGAAGCGTAGATTGCCAATTAGAGTTGCTACTGTTGATGATTCTGCCTCTCCAATTTCTTTG GAAGATCACTGA
CYP90D1	ATGGACACCTCTTCATCCTTGTTGTTCTTCAGCTTTTTCTTCTTCATTATCATCGTCATCTTCAACA AGATCAACGGCTTGAGATCTTCTCCAGCTTCTAAGAAGAAATTGAACGATCATCACGTCACCTCT CAATCTCATGGTCCAAAATTTCCACATGGTTCTTTAGGTTGGCCAGTTATTGGTGAAACTATCGA GTTTGTTTCCTCCGCTTATTCTGATAGACCAGAATCTTTCATGGACAAGAGAAGATTGATGTACG GTAGGGTTTTCAAGTCCCATATTTTTGGTACTGCTACCATCGTTTCTACCGATGCTGAAGTTAATA GAGCTGTCTTGCAATCTGATTCTACTGCTTTTGTTCCATTCTACCCAAAGACCGTTAGAGAATTGA TGGGTAAGAGTTCCATCCTGTTGATTAACGGTTCATTGCATAGAAGATTCCACGGTTTGGTTGGT TCCTTTTTGAAATCTCCTTTGTTGAAGGCCCAAATCGTTAGAGATATGCACAAGTTCTTGTCCGAA TCTATGGATTTGTGGTCTGAAGATCAACCAGTTTTGTTGCAAGATGTTTCTAAGACCGTTGCCTTT AAGGTTTTGGCTAAGGCTTTGATTTCTGTCGAAAAGGGTGAAGATTTGGAAGAGTTGAAGAGGG AATTCGAGAACTTCATTTCTGGCTTGATGTCTTTGCCAATCAATTTCCCAGGTACTCAATTGCACA GATCATTGCAAGCTAAAAAGAACATGGTCAAGCAGGTCGAAAGAATCATTGAAGGTAAGATCAGA AAGACCAAGAACAAAGAAGAGGACGACGTTATTGCTAAGGATGTTGTTGACGTGTTGCTGAAGG ATTCTTCTGAACATTTGACCCATAACTTGATCGCCAACAACATGATCGATATGATGATTCCAGGTC ACGATTCTGTTCCAGTTTTGATTACTTTGGCCGTGAAGTTTTTGTCTGATTCACCAGCTGCTTTGA ACTTGTTGACTGAAGAAAACATGAAGCTGAAGTCCCTGAAAGAATTGACTGGTGAACCATTATAC TGGAACGACTATTTGTCATTGCCATTCACTCAAAAGGTTATCACCGAAACTTTGAGAATGGGTAA CGTTATCATCGGTGTTATGAGAAAGGCTATGAAGGATGTAGAAATCAAGGGTTACGTTATCCCAA AAGGTTGGTGTTTTTTGGCCTACTTGAGATCAGTTCACTTGGACAAGTTGTACTACGAATCTCCC TACAAGTTTAATCCTTGGAGATGGCAAGAAAGGGATATGAACACTTCTTCTTTTTCTCCATTTGGT GGTGGTCAAAGATTGTGTCCAGGTTTGGATTTGGCTAGATTGGAAACATCTGTTTTCTTGCATCA TTTGGTCACCAGATTCAGATGGATTGCTGAAGAAGATACCATCATTAACTTCCCAACTGTCCACA TGAAGAACAAGTTGCCAATTTGGATCAAGAGGATCTAA
det2	ATGGAAGAAATTGCCGATAAGACCTTCTTCAGATACTGTTTGTTGACCTTGATTTTTGCTGGTCCA CCAACTGCTGTTTTGTTGAAGTTTTTACAAGCCCCATACGGTAAGCACAATAGAACTGGTTGGGG TCCAACTGTTTCTCCACCAATAGCTTGGTTTGTTATGGAATCTCCAACTCTGTGGTTGACGTTGTT GTTGTTTCCATTTGGTAGACATGCTCTGAACCCAAAGTCTTTGTTGTTATTCTCCCCATACTTGAT CCACTACTTCCATAGAACTATCATCTACCCATTGAGGTTGTTCAGATCTTCATTTCCAGCTGGTAA GAACGGTTTCCCAATTACTATTGCTGCTTTGGCTTTCACCTTCAACTTGTTGAATGGTTACATTCA AGCCAGATGGGTTTCTCACTACAAGGATGATTATGAAGATGGTAATTGGTTCTGGTGGCGTTTCG TTATTGGTATGGTTGTTTTCATTACCGGCATGTACATTAACATCACCTCTGATAGAACCTTGGTCA GGTTGAAGAAAGAAAACAGAGGTGGTTACGTTATTCCAAGAGGTGGATGGTTTGAATTGGTTTCT TGTCCAAATTACTTCGGTGAAGCCATAGAATGGTTAGGTTGGGCTGTTATGACTTGGTCTTGGGC TGGTATTGGTTTTTTCTTGTACACCTGTTCTAACTTGTTCCCAAGAGCTAGAGCTTCTCATAAGTG GTACATTGCCAAGTTCAAAGAAGAATACCCAAAGACCAGAAAGGCTGTTATTCCATTCGTTTACT AA
dwf1	ATGTCGGATCTTCAGACACCGCTTGTGAGGCCCAAGAGGAAGAAGACTTGGGTTGATTACTTTGT CAAGTTCAGATGGATCATTGTCATCTTCATCGTCCTTCCATTCTCAGCCACATTCTACTTCCTCAT CTACCTCGGGGACATGTGGTCAGAGTCCAAGTCCTTTGAGAAACGTCAGAAGGAACACGACGA GAATGTCAAGAAAGTCATCAAAAGGCTTAAGGGTAGGGATGCTTCCAAGGACGGGCTTGTCTGC ACTGCTCGTAAGCCCTGGATCGCTGTTGGAATGAGGAACGTTGACTACAAGAGAGCCCGGCATT TCGAGGTTGACTTGGGGGAGTTCCGTAACATCCTTGAGATCAACAAGGAGAAGATGACTGCTAG AGTGGAGCCTCTTGTTAACATGGGACAGATTTCCCGTGCTACCGTCCCAATGAACCTGTCTCTC GCTGTTGTTGCTGAGCTTGATGACCTTACCGTTGGTGGACTTATCAATGGATATGGTATTGAAGG AAGCTCTCACATCTACGGTTTGTTTGCTGATACCGTTGAGGCTTACGAGATTGTTCTTGCGGGTG GAGAGCTTGTCCGCGCCACAAGGGATAATGAGTATTCTGATCTTTACTACGCAATCCCGTGGTC GCAAGGAACTCTTGGACTCCTTGTAGCTGCTGAGATCAGGCTTATTAAAGTCAAGGAGTACATGA GACTCACTTACATACCAGTCAAGGGTGATCTTCAAGCCTTAGCTCAAGGTTACATTGATTCTTTTG CTCCCAAAGACGGTGACAAGTCGAAAATCCCGGATTTCGTCGAAGGCATGGTTTACAATCCAAC

	GGAAGGAGTGATGATGGTTGGAACATATGCATCTAAAGAAGAGGCAAAGAAGAAAGGGAACAAA ATCAACAATGTGGGATGGTGGTTCAAGCCGTGGTTCTACCAGCACGCGCAGACCGCCCTGAAAA AGGGACAGTTTGTTGAGTACATCCCAACTCGTGAATACTACCACAGGCACACAAGGTGCTTGTA CTGGGAAGGGAAGCTTATTCTTCCATTTGGTGATCAGTTCTGGTTTAGGTACCTCTTAGGTTGGT TGATGCCTCCAAAGGTCTCTCTTCTTAAGGCCACTCAAGGTGAAGCTATCAGGAACTATTACCAT GATATGCATGTTATTCAGGATATGCTTGTTCCTCTTTACAAGGTTGGCGATGCACTCGAATGGGT CCACCGCGAAATGGAGGTGTATCCAATTTGGCTTTGCCCACACAAACTCTTCAAGCAGCCAATC AAAGGCCAAATCTACCCAGAGCCAGGCTTCGAGTACGAAAACAGACAAGGAGACACAGAAGATG CACAGATGTACACTGATGTTGGAGTCTACTACGCACCTGGCTGTGTCCTAAGAGGTGAAGAGTT TGATGGATCAGAAGCAGTGCGTAGGATGGAGAAATGGCTGATAGAGAACCATGGATTCCAGCCT CAGTACGCGGTGTCTGAGCTCGACGAGAAGAGCTTCTGGAGAATGTTTAATGGTGAATTGTATG AGGAGTGCCGCAAGAAGTATAGAGCTATTGGAACGTTCATGAGTGTTTACTACAAGTCCAAGAAA GGAAGGAAGACTGAGAAAGAAGTTAGAGAAGCCGAACAAGCTCATCTCGAAACTGCTTATGCCG AGGCAGATTAA
dwf5	ATGGCGGAGACTGTACATTCTCCGATCGTTACTTACGCATCGATGTTATCGCTTCTCGCCTTCTG TCCACCTTTCGTCATTCTCCTATGGTACACAATGGTTCATCAGGATGGTTCTGTTACTCAGACCTT TGGCTTCTTTTGGGAGAATGGAGTTCAAGGACTTATCAACATATGGCCAAGACCCACTTTGATTG CTTGGAAAATTATATTTTGCTATGGAGCATTTGAAGCTATTCTTCAGCTGCTTCTGCCTGGTAAAA GAGTTGAGGGTCCAATATCTCCAGCCGGAAACCGACCAGTTTACAAGGCCAATGGTCTGGCTGC TTACTTTGTGACACTAGCAACCTATCTTGGTCTTTGGTGGTTTGGAATCTTCAACCCTGCAATTGT CTATGATCACTTGGGTGAAATATTTTCGGCACTAATATTCGGAAGCTTCATATTTTGTGTTTTGTT GTACATAAAAGGCCATGTTGCACCTTCATCAAGTGACTCTGGTTCATGTGGTAACCTAATAATTG ACTTCTATTGGGGCATGGAGTTGTACCCTCGAATTGGTAAGAGCTTTGACATCAAGGTGTTTACT AATTGCAGATTCGGAATGATGTCTTGGGCAGTTCTTGCAGTCACGTACTGCATAAAACAGTATGA AATAAATGGCAAAGTATCTGATTCAATGCTGGTGAACACCATCCTGATGCTGGTGTATGTCACAA AATTCTTCTGGTGGGAAGCTGGTTATTGGAACACCATGGACATTGCACATGACCGAGCTGGATT CTATATATGCTGGGGTTGTCTAGTGTGGGTGCCTTCTGTCTACACTTCTCCAGGCATGTACCTTG TGAACCACCCCGTCGAACTCGGAACTCAGTTGGCAATATACATTCTCGTTGCAGGAATTCTGTGC ATTTACATAAACTATGACTGTGATAGACAAAGGCAAGAGTTCAGGAGGACAAACGGGAAATGTTT GGTTTGGGGAAGGGCCCCGTCAAAGATTGTGGCGTCGTATACTACAACATCTGGTGAAACTAAA ACTAGTCTTCTCTTAACGTCTGGATGGTGGGGATTGGCTCGTCATTTCCATTATGTTCCTGAGAT CTTAAGTGCTTTCTTCTGGACCGTACCGGCTCTCTTCGATAACTTCTTGGCATACTTCTACGTCAT ATTTCTCACCCTTCTTCTCTTTGATCGAGCCAAGAGAGACGATGACCGATGCCGATCAAAGTATG GGAAATATTGGAAGCTGTATTGTGAGAAAGTCAAATACAGGATCATTCCGGGAATTTATTGA
dwf7	ATGGCTGCTGATAATGCTTACTTGATGCAATTCGTTGACGAAACCTCATTCTACAACAGAATCGTT TTGTCCCATTTGTTGCCAGCTAATTTGTGGGAACCATTGCCACATTTTTTACAAACCTGGTTGAGA AACTACTTGGCCGGTACTTTGTTGTACTTCATTTCTGGTTTTCTGTGGTGCTTCTACATCTACTAC TTGAAGATCAACGTCTACTTGCCAAAGGATGCTATTCCAACTATTAAGGCCATGAGATTGCAAAT GTTCGTTGCTATGAAGGCTATGCCATGGTATACTTTGTTACCAACCGTTTCCGAATCCATGATTG AACGTGGTTGGACAAAATGCTTTGCCTCTATTGGTGAATTTGGCTGGATCTTGTACTTCGTTTAC ATTGCCATCTACTTGGTGTTCGTTGAATTTGGTATCTACTGGATGCATAGAGAATTGCATGATATC AAGCCCCTGTACAAATACTTGCATGCTACTCATCACATCTACAACAAGCAAAACACTTTGTCTCCA TTCGCTGGTTTGGCTTTTCATCCAGTTGATGGTATTTTACAAGCCGTTCCACATGTTATTGCCTTG TTCATAGTTCCAATCCATTTCACTACCCACATCGGTTTGTTGTTCATGGAAGCTATTTGGACCGCT AACATTCATGATTGCATTCATGGTAACATCTGGCCAGTTATGGGTGCTGGTTATCATACAATTCAT CACACCACCTACAAGCACAATTACGGTCATTATACCATTTGGATGGACTGGATGTTTGGTTCTTT GAGAGATCCTTTGTTGGAAGAGGATGATAACAAGGACTCTTTTAAGAAGGCCGAGTAG
smt2	ATGGACTCTTTAACACTCTTCTTCACCGGTGCACTCGTCGCCGTCGGTATCTACTGGTTCCTCTG CGTTCTCGGTCCAGCAGAGCGTAAAGGCAAACGAGCCGTAGATCTCTCTGGTGGCTCAATCTCC GCCGAGAAAGTCCAAGACAACTACAAACAGTACTGGTCTTTCTTCCGCCGTCCAAAAGAAATCG AAACCGCCGAGAAAGTTCCAGACTTCGTCGACACATTCTACAATCTCGTCACCGACATATACGAG TGGGGATGGGGACAATCCTTCCACTTCTCACCATCAATCCCCGGAAAATCTCACAAAGACGCCA CGCGCCTCCACGAAGAGATGGCCGTAGATCTGATCCAAGTCAAACCTGGTCAAAAGATCCTAGA CGTCGGATGCGGTGTCGGCGGTCCGATGCGAGCGATTGCATCTCACTCGCGAGCTAACGTAGT CGGGATTACAATAAACGAGTATCAGGTGAACAGAGCTCGTCTCCACAATAAGAAAGCTGGTCTC GACGCGCTTTGCGAGGTCGTGTGTGGTAACTTCCTCCAGATGCCGTTCGATGACAACAGTTTCG ACGGTGCTTATTCCATCGAAGCCACGTGTCACGCGCCGAAGCTGGAAGAAGTGTACGCAGAGA TCTACAGGGTGTTGAAACCCGGATCTATGTATGTGTCGTACGAGTGGGTTACGACGGAGAAATT TAAGGCGGAGGATGACGAACACGTGGAGGTAATCCAAGGGATTGAGAGAGGCGATGCGTTACC AGGGCTTAGGGCTTACGTGGATATAGCTGAGACGGCTAAAAAGGTTGGGTTTGAGATAGTGAAG GAGAAGGATCTGGCGAGTCCACCGGCTGAGCCGTGGTGGACTAGGCTTAAGATGGGTAGGCTT GCTTATTGGAGGAATCACATTGTGGTTCAGATTTTGTCAGCGGTTGGAGTTGCTCCTAAAGGAAC

	TGTTGATGTTCATGAGATGTTGTTTAAGACTGCTGATTATTTGACCAGAGGAGGTGAAACCGGAA TATTCTCTCCGATGCATATGATTCTCTGCAGAAAACCGGAGTCACCGGAGGAGAGTTCTTGA
are1	ATGACGGAGACTAAGGATTTGTTGCAAGACGAAGAGTTTCTTAAGATCCGCAGACTCAATTCCGC AGAAGCCAACAAACGGCATTCGGTCACGTACGATAACGTGATCCTGCCACAGGAGTCCATGGAG GTTTCGCCACGGTCGTCTACCACGTCGCTGGTGGAGCCAGTGGAGTCGACTGAAGGAGTGGAG TCGACTGAGGCGGAACGTGTGGCAGGGAAGCAGGAGCAGGAGGAGGAGTACCCTGTGGACGC CCACATGCAAAAGTACCTTTCACACCTGAAGAGCAAGTCTCGGTCGAGGTTCCACCGAAAGGAT GCTAGCAAGTATGTGTCGTTTTTTGGGGACGTGAGTTTTGATCCTCGCCCCACGCTCCTGGACA GCGCCATCAACGTGCCCTTCCAGACGACTTTCAAAGGTCCGGTGCTGGAGAAACAGCTCAAAAA TTTACAGTTGACAAAGACCAAGACCAAGGCCACGGTGAAGACTACGGTGAAGACTACGGAGAAA ACGGACAAGGCAGATGCCCCCCCAGGAGAAAAACTGGAGTCGAACTTTTCAGGGATCTACGTGT TCGCATGGATGTTCTTGGGCTGGATAGCCATCAGGTGCTGCACAGATTACTATGCGTCGTACGG CAGTGCATGGAATAAGCTGGAAATCGTGCAGTACATGACAACGGACTTGTTCACGATCGCAATG TTGGACTTGGCAATGTTCCTGTGCACTTTCTTCGTGGTTTTCGTGCACTGGCTGGTGAAAAAGCG GATCATCAACTGGAAGTGGACTGGGTTCGTTGCAGTGAGCATCTTCGAGTTGGCTTTCATCCCC GTGACGTTCCCCATTTACGTCTACTACTTTGATTTCAACTGGGTCACGAGAATCTTCCTGTTCCT GCACTCCGTGGTGTTTGTTATGAAGAGCCACTCGTTTGCCTTTTACAACGGGTATCTTTGGGACA TAAAGCAGGAACTCGAGTACTCTTCCAAACAGTTGCAAAAATACAAGGAATCTTTGTCCCCAGAG ACCCGCGAGATTCTGCAAAAAAGTTGCGACTTTTGCCTTTTCGAATTGAACTACCAGACCAAGGA TAACGACTTCCCCAACAACATCAGTTGCAGCAATTTCTTCATGTTCTGTTTGTTCCCCGTCCTCGT GTACCAGATCAACTACCCAAGAACGTCGCGCATCAGATGGAGGTATGTGTTGGAGAAGGTGTGC GCCATCATTGGCACCATCTTCCTCATGATGGTCACGGCACAGTTCTTCATGCACCCGGTGGCCA TGCGCTGTATCCAGTTCCACAACACGCCCACCTTCGGCGGCTGGATCCCCGCCACGCAAGAGT GGTTCCACCTGCTCTTCGACATGATTCCGGGCTTCACTGTTCTGTACATGCTCACGTTTTACATG ATATGGGACGCTTTATTGAATTGCGTGGCGGAGTTGACCAGGTTTGCGGACAGATATTTCTACG GCGACTGGTGGAATTGCGTTTCGTTTGAAGAGTTTAGCAGAATCTGGAACGTCCCCGTTCACAA ATTTTTACTAAGACACGTGTACCACAGCTCCATGGGCGCATTGCATTTGAGCAAGAGCCAAGCTA CATTATTTACTTTTTTCTTGAGTGCCGTGTTCCACGAAATGGCCATGTTCGCCATTTTCAGAAGGG TTAGAGGATATCTGTTCATGTTCCAACTGTCGCAGTTTGTGTGGACTGCTTTGAGCAACACCAAG TTTCTACGGGCAAGACCGCAGTTGTCCAACGTTGTCTTTTCGTTTGGTGTCTGTTCAGGGCCCA GTATCATTATGACGTTGTACCTGACCTTATGA
are2	ATGGACAAGAAGAAGGATCTACTGGAGAACGAACAATTTCTCCGCATCCAAAAGCTCAACGCTG CCGATGCGGGCAAAAGACAATCTATAACAGTGGACGACGAGGGCGAACTATATGGGTTAGACAC CTCCGGCAACTCACCAGCCAATGAACACACAGCTACCACAATTACACAGAATCACAGCGTGGTG GCCTCAAACGGAGACGTCGCATTCATCCCAGGAACTGCTACCGAAGGCAATACAGAGATTGTAA CTGAAGAAGTGATTGAGACCGATGATAACATGTTCAAGACCCATGTGAAGACTTTAAGCTCCAAA GAGAAGGCACGGTATAGGCAAGGGTCCTCTAACTTTATATCGTATTTCGATGATATGTCATTTGA ACACAGGCCCAGTATATTAGATGGGTCAGTTAACGAGCCCTTCAAGACCAAATTCGTGGGACCT ACTTTAGAAAAGGAGATCAGAAGAAGGGAGAAAGAGCTAATGGCCATGCGCAAAAATTTACACC ACCGCAAGTCCTCCCCAGATGCTGTCGACTCAGTAGAGAAAAATGATGGCGCCGCCCCAACTAC TGTTCCAACTGCCGCCACCTCAGAAACGGTGGTCACGGTTGAAACCACCATACTTTCATCCAATT TCTCCGGGTTGTACGTGGCGTTTTGGATGGCTATTGCATTTGGTGCTGTCAAGGCTTTAATAGAC TATTATTACCAGCATAATGGTAGCTTCAAGGATTCGGAGATCTTGAAATTTATGACTACGAATTTG TTCACTGTGGCATCTGTAGATCTTTTGATGTATTTGAGCACTTATTTTGTCGTTGGAATACAATAC TTATGCAAGTGGGGGGTCTTGAAATGGGGCACTACCGGCTGGATCTTCACCTCAATTTACGAGT TTTTGTTTGTTATCTTCTACATGTATTTAACAGAAAACATCCTAAAACTACACTGGCTGTCCAAGAT CTTCCTTTTTTTGCATTCTTTAGTTTTATTGATGAAAATGCATTCTTTCGCCTTCTACAATGGCTAT CTATGGAGTATAAAGGAAGAACTACAATTTTCCAAAAGCGCTCTTGCCAAATACAAGGATTCTATA AATGATCCAAAAGTTATTGGTGCTCTTGAGAAAAGTTGTGAGTTTTGTAGTTTTGAATTGAGCTCT CAGTCTTTAAGCGACCAAACTCAAAAATTCCCCAACAATATCAGTGCAAAAAGCTTTTTTTGGTTC ACCATGTTTCCAACCCTAATTTACCAAATTGAATATCCAAGAACTAAGGAAATCAGATGGAGCTAC GTATTAGAAAAGATCTGCGCCATCTTCGGTACCATTTTCTTAATGATGATAGATGCTCAAATCTTG ATGTATCCTGTAGCAATGAGAGCATTGGCTGTGCGCAATTCTGAATGGACTGGTATATTGGATAG ATTATTGAAATGGGTTGGATTGCTCGTTGATATCGTCCCAGGGTTTATCGTGATGTACATCTTGG ACTTCTATTTGATTTGGGATGCCATTTTGAACTGTGTGGCTGAATTGACAAGATTTGGCGACAGA TATTTCTACGGTGACTGGTGGAATTGTGTTAGTTGGGCAGACTTCAGTAGAATTTGGAACATCCC AGTGCATAAGTTTTTGTTAAGACATGTTTACCATAGTTCAATGAGTTCATTCAAATTGAACAAGAG TCAAGCAACTTTGATGACCTTTTTCTTAAGTTCCGTCGTTCATGAATTAGCAATGTACGTTATCTT CAAGAAATTGAGGTTTTACTTGTTCTTCTTCCAAATGCTGCAAATGCCATTAGTAGCTTTAACAAA TACTAAATTCATGAGGAACAGAACCATAATCGGAAATGTTATTTTCTGGCTTGGTATCTGCATGG GACCAAGTGTCATGTGTACGTTGTACTTGACATTCTAA

erg3	ATGGATTTGGTCTTAGAAGTCGCTGACCATTATGTCTTAGACGACTTGTACGCTAAAGTTCTGCC CGCTTCGTTGGCAGCTAATATTCCTGTCAAGTGGCAGAAATTGCTAGGGTTGAACAGTGGGTTC AGCAATTCTACGATTTTGCAGGAGACTTTGAACTCCAAGAATGCCGTCAAAGAATGTAGAAGGTT CTACGGGCAGGTGCCATTCCTGTTTGATATGTCGACGACGTCTTTTGCATCGCTATTGCCTCGTT CCAGCATCTTGAGAGAATTCCTCTCACTATGGGTTATTGTTACGATCTTTGGTTTACTACTTTACT TATTCACGGCTAGTCTCAGCTACGTGTTTGTGTTTGACAAGTCGATTTTCAACCATCCTCGTTACT TGAAAAACCAAATGGCAATGGAAATCAAGTTGGCAGTCAGTGCTATCCCATGGATGTCGATGTTG ACCGTTCCATGGTTTGTTATGGAATTGAACGGCCATTCTAAACTATACATGAAGATTGATTATGAA AACCACGGTGTAAGGAAGCTCATTATCGAGTACTTCACTTTCATCTTTTTCACTGATTGCGGTGT GTATTTAGCGCACAGATGGTTGCATTGGCCAAGGGTCTACCGTGCTCTGCACAAGCCTCATCAC AAGTGGCTGGTCTGCACACCTTTCGCATCTCATTCTTTCCATCCTGTAGACGGGTTTTTGCAATC CATCTCGTACCACATCTACCCATTGATTCTGCCATTACACAAGGTTTCTTATTTGATTCTGTTCAC TTTTGTTAACTTTTGGACTGTTATGATTCATGACGGTCAATACCTATCAAACAATCCTGCCGTCAA CGGTACTGCCTGCCACACGGTTCACCATCTATATTTCAACTACAACTACGGTCAATTCACCACTC TGTGGGACAGACTAGGGGGTTCTTACCGTAGACCAGATGACTCATTGTTTGATCCTAAGTTAAGA GATGCTAAGGAGACCTGGGACGCTCAAGTTAAGGAAGTTGAACATTTCATCAAGGAGGTCGAAG GTGATGATAATGATAGAATCTATGAAAACGACCCAAATACCAAGAAGAACAACTGA
erg4	ATGGCAAAGGATAATAGTGAGAAGCTGCAGGTGCAGGGAGAGGAGAAAAAGTCCAAGCAACCG GTTAATTTCCTGCCTCAGGGTAAATGGCTGAAGCCAAATGAAATCGAATATGAGTTTGGTGGGAC TACTGGTGTTATTGGTATGCTGATCGGGTTTCCACTGCTAATGTACTATATGTGGATTTGTGCGG AATTTTATCACGGTAAGGTTGCCCTACCCAAGGCTGGTGAATCGTGGATGCACTTTATCAAGCAC CTATACCAGTTAGTCTTGGAGAACGGTATCCCAGAAAAGTATGACTGGACTATTTTCTTAACATTT TGGGTGTTTCAGATCATTTTCTACTATACGTTGCCCGGGATTTGGACAAAAGGTCAACCATTGTC TCATTTGAAGGGAAAACAATTGCCTTACTTTTGTAATGCCATGTGGACCTTGTATGTAACTACCAC TTTGGTCTTGGTTTTGCACTTTACCAATCTTTTTAGATTGTATGTCATTATTGACCGTTTTGGGAG GATCATGACATGTGCCATTATTTCAGGGTTTGCCTTCTCCATCATATTGTACTTATGGACTTTATT TATCTCACATGACTATCATAGAATGACAGGAAACCATCTATATGATTTCTTCATGGGAGCTCCACT AAACCCTAGGTGGGGGATTTTGGACTTGAAGATGTTTTTCGAGGTTAGATTACCTTGGTTCACCC TTTACTTTATCACTTTGGGTGCCTGTTTGAAGCAGTGGGAGACTTACGGCTATGTGACACCACAA TTGGGGGTTGTCATGTTAGCTCATTGGTTGTACGCGAACGCATGTGCTAAAGGTGAAGAATTGAT TGTTCCAACCTGGGACATGGCTTACGAAAAGTTTGGATTTATGCTGATCTTCTGGAATATTGCCG GTGTCCCATACACTTACTGTCATTGTACGTTGTATTTGTACTACCATGACCCATCTGAATATCACT GGTCTACACTGTACAATGTTTCGCTGTACGTTGTTCTATTATGCGCCTACTACTTCTTTGACACGG CAAATGCTCAGAAAAATGCCTTCAGAAAGCAAATGTCTGGTGACAAGACAGGTAGGAAGACTTTC CCATTTTTGCCATACCAAATTTTGAAGAATCCAAAGTATATGGTTACCTCCAATGGATCGTACCTA TTGATTGATGGTTGGTACACTTTGGCTAGAAAAATTCACTACACTGCCGATTGGACTCAATCTCT CGTTTGGGCCTTGTCTTGCGGGTTCAACTCGGTGTTCCCATGGTTTTTCCCAGTATTCTTCCTTG TTGTCCTGATTCACAGAGCCTTCAGAGACCAAGCAAAATGTAAGAGAAAGTACGGAAAAGATTG GGATGAGTATTGTAAACATTGCCCTTACGTCTTTATTCCTTATGTTTTCTAG
erg8	ATGTCAGAGTTGAGAGCCTTCAGTGCCCCAGGGAAAGCGTTACTAGCTGGTGGATATTTAGTTTT AGATACAAAATATGAAGCATTTGTAGTCGGATTATCGGCAAGAATGCATGCTGTAGCCCATCCTT ACGGTTCATTGCAAGGGTCTGATAAGTTTGAAGTGCGTGTGAAAAGTAAACAATTTAAAGATGGG GAGTGGCTGTACCATATAAGTCCTAAAAGTGGCTTCATTCCTGTTTCGATAGGCGGATCTAAGAA CCCTTTCATTGAAAAAGTTATCGCTAACGTATTTAGCTACTTTAAACCTAACATGGACGACTACTG CAATAGAAACTTGTTCGTTATTGATATTTTCTCTGATGATGCCTACCATTCTCAGGAGGATAGCGT TACCGAACATCGTGGCAACAGAAGATTGAGTTTTCATTCGCACAGAATTGAAGAAGTTCCCAAAA CAGGGCTGGGCTCCTCGGCAGGTTTAGTCACAGTTTTAACTACAGCTTTGGCCTCCTTTTTTGTA TCGGACCTGGAAAATAATGTAGACAAATATAGAGAAGTTATTCATAATTTAGCACAAGTTGCTCAT TGTCAAGCTCAGGGTAAAATTGGAAGCGGGTTTGATGTAGCGGCGGCAGCATATGGATCTATCA GATATAGAAGATTCCCACCCGCATTAATCTCTAATTTGCCAGATATTGGAAGTGCTACTTACGGC AGTAAACTGGCGCATTTGGTTGATGAAGAAGACTGGAATATTACGATTAAAAGTAACCATTTACC TTCGGGATTAACTTTATGGATGGGCGATATTAAGAATGGTTCAGAAACAGTAAAACTGGTCCAGA AGGTAAAAAATTGGTATGATTCGCATATGCCAGAAAGCTTGAAAATATATACAGAACTCGATCAT GCAAATTCTAGATTTATGGATGGACTATCTAAACTAGATCGCTTACACGAGACTCATGACGATTA CAGCGATCAGATATTTGAGTCTCTTGAGAGGAATGACTGTACCTGTCAAAAGTATCCTGAAATCA CAGAAGTTAGAGATGCAGTTGCCACAATTAGACGTTCCTTTAGAAAAATAACTAAAGAATCTGGT GCCGATATCGAACCTCCCGTACAAACTAGCTTATTGGATGATTGCCAGACCTTAAAAGGAGTTCT TACTTGCTTAATACCTGGTGCTGGTGGTTATGACGCCATTGCAGTGATTACTAAGCAAGATGTTG ATCTTAGGGCTCAAACCGCTAATGACAAAAGATTTTCTAAGGTTCAATGGCTGGATGTAACTCAG GCTGACTGGGGTGTTAGGAAAGAAAAAGATCCGGAAACTTATCTTGATAAATAA

erg10

ATGTCTCAGAACGTTTACATTGTATCGACTGCCAGAACCCCAATTGGTTCATTCCAGGGTTCTCTA
TCCTCCAAGACAGCAGTGGAATTGGGTGCTGTTGCTTTAAAAGGCGCCTTGGCTAAGGTTCCAG
AATTGGATGCATCCAAGGATTTTGACGAAATTATTTTTGGTAACGTTCTTTCTGCCAATTTGGGCC
AAGCTCCGGCCAGACAAGTTGCTTTGGCTGCCGGTTTGAGTAATCATATCGTTGCAAGCACAGT
TAACAAGGTCTGTGCATCCGCTATGAAGGCAATCATTTTGGGTGCTCAATCCATCAAATGTGGTA
ATGCTGATGTTGTCGTAGCTGGTGGTTGTGAATCTATGACTAACGCACCATACTACATGCCAGCA
GCCCGTGCGGGTGCCAAATTTGGCCAAACTGTTCTTGTTGATGGTGTCGAAAGAGATGGGTTGA
ACGATGCGTACGATGGTCTAGCCATGGGTGTACACGCAGAAAAGTGTGCCCGTGATTGGGATAT
TACTAGAGAACAACAAGACAATTTTGCCATCGAATCCTACCAAAAATCTCAAAAATCTCAAAAGGA
AGGTAAATTCGACAATGAAATTGTACCTGTTACCATTAAGGGATTTAGAGGTAAGCCTGATACTC
AAGTCACGAAGGACGAGGAACCTGCTAGATTACACGTTGAAAAATTGAGATCTGCAAGGACTGT
TTTCCAAAAAGAAAACGGTACTGTTACTGCCGCTAACGCTTCTCCAATCAACGATGGTGCTGCAG
CCGTCATCTTGGTTTCCGAAAAAGTTTTGAAGGAAAAGAATTTGAAGCCTTTGGCTATTATCAAAG
GTTGGGGTGAGGCCGCTCATCAACCAGCTGATTTTACATGGGCTCCATCTCTTGCAGTTCCAAA
GGCTTTGAAACATGCTGGCATCGAAGACATCAATTCTGTTGATTACTTTGAATTCAATGAAGCCTT
TTCGGTTGTCGGTTTGGTGAACACTAAGATTTTGAAGCTAGACCCATCTAAGGTTAATGTATATG
GTGGTGCTGTTGCTCTAGGTCACCCATTGGGTTGTTCTGGTGCTAGAGTGGTTGTTACACTGCT
ATCCATCTTACAGCAAGAAGGAGGTAAGATCGGTGTTGCCGCCATTTGTAATGGTGGTGGTGGT
GCTTCCTCTATTGTCATTGAAAAGATATGA

erg12

ATGTCATTAC CGTTCTTAAC TTCTGCACCG GGAAAGGTTA TTATTTTTGG TGAACACTCT 61
GCTGTGTACA ACAAGCCTGC CGTCGCTGCT AGTGTGTCTG CGTTGAGAAC CTACCTGCTA 121
ATAAGCGAGT CATCTGCACC AGATACTATT GAATTGGACT TCCCGGACAT TAGCTTTAAT 181
CATAAGTGGT CCATCAATGA TTTCAATGCC ATCACCGAGG ATCAAGTAAA CTCCCAAAAA 241
TTGGCCAAGG CTCAACAAGC CACCGATGGC TTGTCTCAGG AACTCGTTAG TCTTTTGGAT 301
CCGTTGTTAG CTCAACTATC CGAATCCTTC CACTACCATG CAGCGTTTTG TTTCCTGTAT 361
ATGTTTGTTT GCCTATGCCC CCATGCCAAG AATATTAAGT TTTCTTTAAA GTCTACTTTA 421
CCCATCGGTG CTGGGTTGGG CTCAAGCGCC TCTATTTCTG TATCACTGGC CTTAGCTATG 481
GCCTACTTGG GGGGGTTAAT AGGATCTAAT GACTTGGAAA AGCTGTCAGA AAACGATAAG 541
CATATAGTGA ATCAATGGGC CTTCATAGGT GAAAAGTGTA TTCACGGTAC CCCTTCAGGA 601
ATAGATAACG CTGTGGCCAC TTATGGTAAT GCCCTGCTAT TTGAAAAAGA CTCACATAAT 661
GGAACAATAA ACACAAACAA TTTTAAGTTC TTAGATGATT TCCCAGCCAT TCCAATGATC 721
CTAACCTATA CTAGAATTCC AAGGTCTACA AAAGATCTTG TTGCTCGCGT TCGTGTGTTG 781
GTCACCGAGA AATTTCCTGA AGTTATGAAG CCAATTCTAG ATGCCATGGG TGAATGTGCC 841
CTACAAGGCT TAGAGATCAT GACTAAGTTA AGTAAATGTA AAGGCACCGA TGACGAGGCT 901
GTAGAAACTA ATAATGAACT GTATGAACAA CTATTGGAAT TGATAAGAAT AAATCATGGA 961
CTGCTTGTCT CAATCGGTGT TTCTCATCCT GGATTAGAAC TTATTAAAAA TCTGAGCGAT 1021
GATTTGAGAA TTGGCTCCAC AAAACTTACC GGTGCTGGTG GCGGCGGTTG CTCTTTGACT
1081 TTGTTACGAA GAGACATTAC TCAAGAGCAA ATTGACAGCT TCAAAAAGAA ATTGCAAGAT
1141 GATTTTAGTT ACGAGACATT TGAAACAGAC TTGGGTGGGA CTGGCTGCTG TTTGTTAAGC
1201 GCAAAAAATT TGAATAAAGA TCTTAAAATC AAATCCCTAG TATTCCAATT ATTTGAAAAT
1261 AAAACTACCA CAAAGCAACA AATTGACGAT CTATTATTGC CAGGAAACAC GAATTTACCA
1321 TGGACTTCAT AA

erg13

ATGAAACTCTCAACTAAACTTTGTTGGTGTGGTATTAAAGGAAGACTTAGGCCGCAAAAGCAACA
ACAATTACACAATACAAACTTGCAAATGACTGAACTAAAAAAACAAAAGACCGCTGAACAAAAAAC
CAGACCTCAAAATGTCGGTATTAAAGGTATCCAAATTTACATCCCAACTCAATGTGTCAACCAATC
TGAGCTAGAGAAATTTGATGGCGTTTCTCAAGGTAAATACACAATTGGTCTGGGCCAAACCAACA
TGTCTTTTGTCAATGACAGAGAAGATATCTACTCGATGTCCCTAACTGTTTTGTCTAAGTTGATCA
AGAGTTACAACATCGACACCAACAAAATTGGTAGATTAGAAGTCGGTACTGAAACTCTGATTGAC
AAGTCCAAGTCTGTCAAGTCTGTCTTGATGCAATTGTTTGGTGAAAACACTGACGTCGAAGGTAT
TGACACGCTTAATGCCTGTTACGGTGGTACCAACGCGTTGTTCAACTCTTTGAACTGGATTGAAT
CTAACGCATGGGATGGTAGAGACGCCATTGTAGTTTGCGGTGATATTGCCATCTACGATAAGGG
TGCCGCAAGACCAACCGGTGGTGCCGGTACTGTTGCTATGTGGATCGGTCCTGATGCTCCAATT
GTATTTGACTCTGTAAGAGCTTCTTACATGGAACACGCCTACGATTTTTACAAGCCAGATTTCACC
AGCGAATATCCTTACGTCGATGGTCATTTTTCATTAACTTGTTACGTCAAGGCTCTTGATCAAGTT
TACAAGAGTTATTCCAAGAAGGCTATTTCTAAAGGGTTGGTTAGCGATCCCGCTGGTTCGGATGC
TTTGAACGTTTTGAAATATTTCGACTACAACGTTTTCCATGTTCCAACCTGTAAATTGGTCACAAA
ATCATACGGTAGATTACTATATAACGATTTCAGAGCCAATCCTCAATTGTTCCCAGAAGTTGACG
CCGAATTAGCTACTCGCGATTATGACGAATCTTTAACCGATAAGAACATTGAAAAAACTTTTGTTA
ATGTTGCTAAGCCATTCCACAAAGAGAGAGTTGCCCAATCTTTGATTGTTCCAACAAACACAGGT
AACATGTACACCGCATCTGTTTATGCCGCCTTTGCATCTCTATTAAACTATGTTGGATCTGACGAC
TTACAAGGCAAGCGTGTTGGTTTATTTTCTTACGGTTCCGGTTTAGCTGCATCTCTATATTCTTGC
AAAATTGTTGGTGACGTCCAACATATTATCAAGGAATTAGATATTACTAACAAATTAGCCAAGAGA

	ATCACCGAAACTCCAAAGGATTACGAAGCTGCCATCGAATTGAGAGAAAATGCCCATTTGAAGAA GAACTTCAAACCTCAAGGTTCCATTGAGCATTTGCAAAGTGGTGTTTACTACTTGACCAACATCG ATGACAAATTTAGAAGATCTTACGATGTTAAAAAATAA
erg19	ATGACCGTTTACACAGCATCCGTTACCGCACCCGTCAACATCGCAACCCTTAAGTATTGGGGGA AAAGGGACACGAAGTTGAATCTGCCCACCAATTCGTCCATATCAGTGACTTTATCGCAAGATGAC CTCAGAACGTTGACCTCTGCGGCTACTGCACCTGAGTTTGAACGCGACACTTTGTGGTTAAATG GAGAACCACACAGCATCGACAATGAAAGAACTCAAAATTGTCTGCGCGACCTACGCCAATTAAG AAAGGAAATGGAATCGAAGGACGCCTCATTGCCCACATTATCTCAATGGAAACTCCACATTGTCT CCGAAAATAACTTTCCTACAGCAGCTGGTTTAGCTTCCTCCGCTGCTGGCTTTGCTGCATTGGTC TCTGCAATTGCTAAGTTATACCAATTACCACAGTCAACTTCAGAAATATCTAGAATAGCAAGAAAG GGGTCTGGTTCAGCTTGTAGATCGTTGTTTGGCGGATACGTGGCCTGGGAAATGGGAAAAGCTG AAGATGGTCATGATTCCATGGCAGTACAAATCGCAGACAGCTCTGACTGGCCTCAGATGAAAGC TTGTGTCCTAGTTGTCAGCGATATTAAAAAGGATGTGAGTTCCACTCAGGGTATGCAATTGACCG TGGCAACCTCCGAACTATTTAAAGAAAGAATTGAACATGTCGTACCAAAGAGATTTGAAGTCATG CGTAAAGCCATTGTTGAAAAAGATTTCGCCACCTTTGCAAAGGAAACAATGATGGATTCCAACTC TTTCCATGCCACATGTTTGGACTCTTTCCCTCCAATATTCTACATGAATGACACTTCCAAGCGTAT CATCAGTTGGTGCCACACCATTAATCAGTTTTACGGAGAAACAATCGTTGCATACACGTTTGATG CAGGTCCAAATGCTGTGTTGTACTACTTAGCTGAAAATGAGTCGAAACTCTTTGCATTTATCTATA AATTGTTTGGCTCTGTTCCTGGATGGGACAAGAAATTTACTACTGAGCAGCTTGAGGCTTTCAAC CATCAATTTGAATCATCTAACTTTACTGCACGTGAATTGGATCTTGAGTTGCAAAAGGATGTTGCC AGAGTGATTTTAACTCAAGTCGGTTCAGGCCCACAAGAAACAAACGAATCTTTGATTGACGCAAA GACTGGTCTACCAAAGGAATAA
erg20	ATGGCTTCAGAAAAAGAAATTAGGAGAGAGAGATTCTTGAACGTTTTCCCTAAATTAGTAGAGGA ATTGAACGCATCGCTTTTGGCTTACGGTATGCCTAAGGAAGCATGTGACTGGTATGCCCACTCAT TGAACTACAACACTCCAGGCGGTAAGCTAAATAGAGGTTTGTCCGTTGTGGACACGTATGCTATT CTCTCCAACAAGACCGTTGAACAATTGGGGCAAGAAGAATACGAAAAGGTTGCCATTCTAGGTT GGTGCATTGAGTTGTTGCAGGCTTACTTCTTGGTCGCCGATGATATGATGGACAAGTCCATTACC AGAAGAGGCCAACCATGTTGGTACAAGGTTCCTGAAGTTGGGGAAATTGCCATCAATGACGCAT TCATGTTAGAGGCTGCTATCTACAAGCTTTTGAAATCTCACTTCAGAAACGAAAAATACTACATAG ATATCACCGAATTGTTCCATGAGGTCACCTTCCAAACCGAATTGGGCCAATTGATGGACTTAATC ACTGCACCTGAAGACAAAGTCGACTTGAGTAAGTTCTCCCTAAAGAAGCACTCCTTCATAGTTAC TTTCAAGACTGCTTACTATTCTTTCTACTTGCCTGTCGCATTGGCCATGTACGTTGCCGGTATCAC GGATGAAAAGGATTTGAAACAAGCCAGAGATGTCTTGATTCCATTGGGTGAATACTTCCAAATTC AAGATGACTACTTAGACTGCTTCGGTACCCCAGAACAGATCGGTAAGATCGGTACAGATATCCAA GATAACAAATGTTCTTGGGTAATCAACAAGGCATTGGAACTTGCTTCCGCAGAACAAAGAAAGAC TTTAGACGAAAATTACGGTAAGAAGGACTCAGTCGCAGAAGCCAAATGCAAAAAGATTTTCAATG ACTTGAAAATTGAACAGCTATACCACGAATATGAAGAGTCTATTGCCAAGGATTTGAAGGCCAAA ATTTCTCAGGTCGATGAGTCTCGTGGCTTCAAAGCTGATGTCTTAACTGCGTTCTTGAACAAAGT TTACAAGAGAAGCAAATAG
idi1	ATGACTGCCGACAACAATAGTATGCCCCATGGTGCAGTATCTAGTTACGCCAAATTAGTGCAAAA CCAAACACCTGAAGACATTTTGGAAGAGTTTCCTGAAATTATTCCATTACAACAAAGACCTAATAC CCGATCTAGTGAGACGTCAAATGACGAAAGCGGAGAAACATGTTTTTCTGGTCATGATGAGGAG CAAATTAAGTTAATGAATGAAAATTGTATTGTTTTGGATTGGGACGATAATGCTATTGGTGCCGGT ACCAAGAAAGTTTGTCATTTAATGGAAAATATTGAAAAGGGTTTACTACATCGTGCATTCTCCGTC TTTATTTTCAATGAACAAGGTGAATTACTTTTACAACAAAGAGCCACTGAAAAAATAACTTTCCCT GATCTTTGGACTAACACATGTTGCTCTCATCCACTATGTATTGATGACGAATTAGGTTTGAAGGG TAAGCTAGACGATAAGATTAAGGGCGCTATTACTGCGGCGGTGAGAAAACTAGATCATGAATTA GGTATTCCAGAAGATGAAACTAAGACAAGGGGTAAGTTTCACTTTTTAAACAGAATCCATTACAT GGCACCAAGCAATGAACCATGGGGTGAACATGAAATTGATTACATCCTATTTTATAAGATCAACG CTAAAGAAAACTTGACTGTCAACCCAAATGTCAATGAAGTTAGAAACTTCAAATGGGTTTCACCAA ATGATTTGAAAACTATGTTTGCTGACCCAAGTTACAAGTTTACGCCTTGGTTTAAGATTATTTGCG AGAATTACTTATTCAACTGGTGGGAGCAATTAGATGACCTTTCTGAAGTGGAAAATGACAGACAA ATTCATAGAATGCTATAA

upc2	ATGAGCGAAGTCGGTATACAGAATCACAAGAAAGCGGTGACAAAACCCAGAAGAAGAGAAAAAG TCATCGAGCTAATTGAAGTGGACGGCAAAAAGGTGAGTACGACTTCAACCGGTAAACGTAAATT CCATAACAAATCAAAGAATGGGTGCGATAACTGTAAAAGAAGAAGAGTTAAGTGTGATGAAGGG AAGCCAGCCTGTAGGAAGTGCACAAATATGAAGTTGGAATGTCAGTATACACCAATCCATTTAAG GAAAGGTAGAGGAGCAACAGTAGTGAAGTATGTCACGAGAAAGGCAGACGGTAGCGTGGAGTC TGATTCATCGGTAGATTTACCTCCTACGATCAAGAAGGAGCAGACACCGTTCAATGATATCCAAT CAGCGGTAAAAGCTTCAGGCTCATCCAATGATTCCTTTCCATCAAGCGCCTCTACAACTAAGAGT GAGAGCGAGGAAAAGTCATCGGCCCCTATAGAGGACAAAAACAATATGACTCCTCTAAGTATGG GCCTCCAGGGTACCATCAATAAGAAAGATATGATGAATAACTTTTTCTCTCAAAATGGCACTATTG GTTTTGGTTCTCCTGAAAGATTGAATTCAGGTATCGATGGCTTACTATTACCGCCATTGCCTTCTG GAAATATGGGTGCGTTCCAACTTCAGCAACAGCAGCAAGTGCAGCAGCAATCTCAACCACAGAC CCAAGCGCAGCAAGCAAGTGGAACTCCAAACGAGAGATATGGTTCATTCGATCTTGCGGGTAGT CCTGCATTGCAATCCACGGGAATGAGCTTATCAAATAGTCTAAGCGGGATGTTACTATGTAACAG GATTCCTTCCGGCCAAAACTACACTCAACAACAATTACAATATCAATTACACCAGCAGCTGCAATT GCAACAGCATCAGCAAGTTCAGCTGCAGCAGTATCAACAATTACGTCAGGAACAACACCAACAA GTTCAGCAACAACAACAGGAACAACTCCAGCAATACCAACAACATTTTTTGCAACAGCAGCAACA AGTACTGCTTCAGCAAGAGCAACAACCTAACGATGAGGAAGGTGGCGTTCAGGAAGAAAACAGC AAAAAGGTAAAGGAAGGGCCTTTACAATCACAAACAAGCGAAACTACTTTAAACAGCGATGCTGC TACATTACAAGCTGATGCATTATCTCAGTTAAGTAAGATGGGGCTAAGCCTAAAGTCGTTAAGTA CCTTTCCAACAGCTGGTATTGGTGGTGTTTCCTATGACTTTCAGGAACTGTTAGGTATTAAGTTTC CAATAAATAACGGCAATTCAAGAGCTACTAAGGCCAGCAACGCAGAGGAAGCTTTGGCCAATAT GCAAGAGCATCATGAACGTGCAGCTGCTTCTGTAAAGGAGAATGATGGTCAGCTCTCTGATACG AAGAGTCCAGCGCCATCGAATAACGCCCAAGGGGGAAGTGCTAGTATTATGGAACCTCAGGCG GCTGATGCGGTTTCGACAATGGCGCCTATATCAATGATTGAAAGAAACATGAACAGAAACAGCAA CATTTCTCCATCAACGCCCTCTGCAGTGTTGAATGATAGGCAAGAGATGCAAGATTCTATAAGTT CTCTAGGAAATCTGACAAAAGCAGCCTTGGAGAACAACGAACCAACGATAAGTTTACAAACATCA CAGACAGAGAATGAAGACGATGCATCGCGGCAAGACATGACCTCAAAAATTAATAACGAAGCTG ACCGAAGTTCTGTTTCTGCTGGTACCAGTAACATCGCTAAGCTTTTAGATCTTTCTACCAAAGGC AATCTGAACCTGATAGACATGAAACTGTTTCATCATTATTGCACAAAGGTCTGGCCTACGATTACA GCGGCCAAAGTTTCTGGGCCTGAAATATGGAGGGACTACATACCGGAGTTAGCATTTGACTATC CATTTTTAATGCACGCTTTGTTGGCATTCAGTGCCACCCATCTTTCGAGGACTGAAACTGGACTG GAGCAATACGTTTCATCTCACCGCCTAGACGCTCTGAGATTATTAAGAGAAGCTGTTTTAGAAAT ATCTGAGAATAACACCGATGCGCTAGTTGCCAGCGCCCTGATACTAATCATGGACTCGTTAGCA AATGCTAGTGGTAACGGCACTGTAGGAAACCAAAGTTTGAATAGCATGTCACCAAGCGCTTGGA TCTTTCATGTCAAAGGTGCTGCAACAATTTTAACCGCTGTGTGGCCTTTGAGTGAAAGATCTAAA TTTCATAACATTATATCTGTTGATCTTAGCGATTTAGGCGATGTCATTAACCCTGATGTTGGAACA ATTACTGAATTGGTATGTTTTGATGAAAGTATTGCCGATTTGTATCCTGTCGGCTTAGATTCGCCA TATTTGATAACACTAGCTTATTTAGATAAATTGCACCGTGAAAAAAACCAGGGTGATTTTATTCTG CGGGTATTTACATTTCCAGCATTGCTAGACAAGACATTCCTGGCATTACTGATGACAGGTGATTT AGGTGCAATGAGAATTATGAGATCATATTATAAACTACTTCGAGGATTTGCCACAGAGGTCAAGG ATAAAGTCTGGTTTCTCGAAGGAGTCACGCAGGTGCTGCCTCAAGATGTTGACGAATACAGTGG AGGTGGTGGTATGCATATGATGCTAGATTTCCTCGGTGGCGGATTACCATCGATGACAACAACA AATTTCTCTGATTTTTCGTTATGA
thmg1	ATGGCAAAGGATAATAGTGAGAAGCTGCAGGTGCAGGGAGAGGAGAAAAAGTCCAAGCAACCG GTTAATTTCCTGCCTCAGGGTAAATGGCTGAAGCCAAATGAAATCGAATATGAGTTTGGTGGGAC TACTGGTGTTATTGGTATGCTGATCGGGTTTCCACTGCTAATGTACTATATGTGGATTTGTGCGG AATTTTATCACGGTAAGGTTGCCCTACCCAAGGCTGGTGAATCGTGGATGCACTTTATCAAGCAC CTATACCAGTTAGTCTTGGAGAACGGTATCCCAGAAAAGTATGACTGGACTATTTTCTTAACATTT TGGGTGTTTCAGATCATTTTCTACTATACGTTGCCCGGGATTTGGACAAAAGGTCAACCATTGTC TCATTTGAAGGGAAAACAATTGCCTTACTTTTGTAATGCCATGTGGACCTTGTATGTAACTACCAC TTTGGTCTTGGTTTTGCACTTTACCAATCTTTTTAGATTGTATGTCATTATTGACCGTTTTGGGAG GATCATGACATGTGCCATTATTTCAGGGTTTGCCTTCTCCATCATATTGTACTTATGGACTTTATT TATCTCACATGACTATCATAGAATGACAGGAAACCATCTATATGATTTCTTCATGGGAGCTCCACT AAACCCTAGGTGGGGGATTTTGGACTTGAAGATGTTTTTCGAGGTTAGATTACCTTGGTTCACCC TTTACTTTATCACTTTGGGTGCCTGTTTGAAGCAGTGGGAGACTTACGGCTATGTGACACCACAA TTGGGGGTTGTCATGTTAGCTCATTGGTTGTACGCGAACGCATGTGCTAAAGGTGAAGAATTGAT TGTTCCAACCTGGGACATGGCTTACGAAAAGTTTGGATTTATGCTGATCTTCTGGAATATTGCCG GTGTCCCATACACTTACTGTCATTGTACGTTGTATTTGTACTACCATGACCCATCTGAATATCACT GGTCTACACTGTACAATGTTTCGCTGTACGTTGTTCTATTATGCGCCTACTACTTCTTTGACACGG CAAATGCTCAGAAAAATGCCTTCAGAAAGCAAATGTCTGGTGACAAGACAGGTAGGAAGACTTTC CCATTTTTGCCATACCAAATTTTGAAGAATCCAAAGTATATGGTTACCTCCAATGGATCGTACCTA TTGATTGATGGTTGGTACACTTTGGCTAGAAAAATTCACTACACTGCCGATTGGACTCAATCTCT CGTTTGGGCCTTGTCTTGCGGGTTCAACTCGGTGTTCCCATGGTTTTTCCCAGTATTCTTCCTTG TTGTCCTGATTCACAGAGCCTTCAGAGACCAAGCAAAATGTAAGAGAAAGTACGGAAAAGATTG

	GGATGAGTATTGTAAACATTGCCCTTACGTCTTTATTCCTTATGTTTTCTAG
ADTO1	ATGGCTGAATCTGCTGGTGAATCCTACAGATTATTGTCCGTTTCTTCTTCTACTACCTTCTTGGCC TTCATCATCATTTTTTTGTTGGCTGGTATTGCCAGAAGAAAAAGAAGGGCTCCACATAGATTGCC ACCAGGTTCTAGAGGTTGGCCATTGATTGGTGATACTTTTGCTTGGTTGAATGCTGTTGCTGGTT CTCATCCATCTTCATTTGTTGAAAAGCAGATCAAGAAGTACGGCAGAATTTTCTCTTGTTCCTTGT TTGGTAAATGGGCTGTTGTTTCTGCTGATCCAGATTTCAACAGATTCATCATGCAAAACGAAGGC AAGCTGTTCCAATCTTCTTACCCAAAGTCTTTCAGAGATTTGGTTGGTAAGGATGGTGTTATTACC GTTCATGGTGACCAACAGAGAAGATTGCATTCTATTGCTAGCTCCATGATGAGACACGATCAATT GAAAACCCACTTCTTGGAAGTTATCCCAGTTGTTATGTTGCAGACCTTGTCTAATTTCAAGGACG GTGAAGTTGTCTTGTTGCAAGATATCTGTAGAAAGGTTGCCATCCACTTGATGGTTAATCAATTAT TGGGTGTCTCCTCCGAATCTGAAGTTGACGAAATGTCTCAGTTGTTCTCCGATTTTGTTGATGGT TGTTTGTCCGTTCCAATTGATTTGCCAGGTTTCACTTACAACAAAGCCATGAAGGCTAGGAAAGA GATCATCAGAAAGATCAACAAGACCATCGAAAAGAGGTTGCAAAACAAAGCTGCTTCTGATACTG CTGGTAATGGTGTTTTGGGTAGACTATTGGAAGAGGAATCTTTGCCAAATGAATCCATGGCCGAT TTCATTATCAACTTGTTGTTTGCTGGTAACGAAACCACTTCTAAGACTATGTTGTTCGCCGTTTAC TTTTTGACCCATTGTCCAAAAGCTATGACGCAGTTGTTGGAAGAACATGATAGATTGGCTGGTGG TATGTTGACTTGGCAAGATTACAAGACTATGGATTTCACCCAATGCGTTATTGACGAAACCTTGA GATTAGGTGGTATTGCTATTTGGTTGATGAGGGAAGCTAAAGAGGATGTTTCTTACCAAGATTAC GTTATCCCAAAGGGTTGTTTCGTTGTTCCATTTTTGTCTGCCGTTCACTTGGATGAGTCTTACTAC AAAGAATCCTTGTCTTTCAATCCTTGGAGATGGTTGGATCCAGAAACTCAACAAAAGAGAAATTG GAGAACCTCTCCATTCTACTGTCCATTTGGTGGTGGTACTAGATTTTGTCCAGGTGCTGAATTGG CTAGATTGCAAATTGCTTTGTTCCTGCATTACTTCATCACTACTTACAAGTGGACCCAATTGAAAG AGGACAGAATTTCATTTTTCCCATCCGCCAGATTGGTTAACGGTTTCAAAATCCAATTGAACAGAA GAGATAGCGATCCACCAAACCAGTAA
CYP722A1	ATGGAACATTTGTGCTTGTGCTTGGTTTTGTGTGCTGCTATGTTGACTTTGGGCAAGTTCTTGAAG ATCATGTTCCAAGACAGAAAAAAGTCTACTGCTGGTGTTCCACCAGGTTCTGATGGTTTTCCAGT TATTGGTGAAACCTTGCAGTTCATGTTGTCTGTTAATTCTGGTAAGGGTTTCTACGAGTTCGTCA GATCTAGAAGAATTAGATACGGTTCTTGCTTCAGGACTTCTTTGTTCGGTGAAACTCATGTTTTCT TGTCTACTACCGAATCCGCTAGATTGGTTTTGAACAATGATTCTGGTATGTTCACCAAGAGGTAC ATCAAATCCATTGGTGAATTGGTTGGTGACAGATCTTTGTTGTGTGCTCCACAACATCATCATAA GATCTTGAGATCCCGTCTGATCAACTTGTTCTCTAAAAGATCTACCGCCTTGATGGTTAGACACT TTGATGAATTAGTTGTCGATGCTTTAGGTGGTTGGGAACATAGAGGTACTGTTGTTTTGTTGACC GACTTGTTGCAAATTACCTTCAAGGCTATGTGCAAGATGTTGGTGTCTTTGGAAAAAGAAGAGGA ATTGGGTTCCATGCAAAGAGATGTAGGTTTTGTTTGTGAAGCCATGTTGGCTTTTCCATTGAATTT GCCATGGACCAGATTCCATAAGGGTATTATGGCTAGAGGTAGAGTCATGGAAATGTTGGAAAAG ATCATCAGAGAGAGGCGTAACGAAATCAACTCTCATAACAACCATCACGAGGATTTCTTGCAACA ATTATTGGCCGTTGATAACGATACTCCACAATTGACTGATGCTGAAATCAAGGATAACATCCTGA CCATGATTATTGCTGGTCAAGATACAACTGCTTCTGCTTTGACTTGGATGGTTAAGTACTTGGGT GAAAACCAGAAGGTTTTGGACATCTTGATCGAAGAACAATCCCAAATTACCAAGAAGGCTTCTAA CAAGCCATTCTTGGAACTGGAAGATTTGTCTGAAATGCCATACGCTTCTAAGATGGTCAAAGAAA GTTTGAGAATGGCCTCTGTTGTTCCATGGTTTCCAAGATTGGTATTGCAAGATTGTGAGATGGAA GGCTACAAGATTAAGAAAGGTTGGAACATTAACATCGACGCCAGATCCATTCATTTGGATCCTAC TGTTTATTCCGAACCACATAAGTTCAACCCATTGAGGTTTGAAGAAGAAGCTAAGGCTAATTCCTT TTTGGCTTTTGGTATGGGTGGTAGAACTTGTTTGGGTTTAGCTTTGGCTAAAGCCATGATGTTGG TTTTCTTGCATAGATTCATTACCACCTACAGATGGGAAGTTGTTGATGAAGATCCATCCATTGAAA AGTGGACTTTGTTTGCCAGATTGAAGTCTGGTTACCCAATCAGAGTTTCTAGAAGGTTGTGA
CYP724A1	aATGTTCATCTTGCAAAACGAGGGCAAGTTGTTCACTTCTGATTATCCAAAAGCCATGCACGACAT TTTGGGCAAGTATTCTTTGTTGTTGGCTACCGGTGAAATCCACAGAAAATTGAAGAACGTGATCA TCTCCTTCATCAACTTGACTAAGTCTAAGCCAGATTTCTTGCATTGCGCTGAAAACTTGTCCATCT CCATTTTGAAATCTTGGAAGAACTGCAGGGAAGTCGAATTTCACAAAGAAGTTAAGATCTTCACC CTGTCCGTTATGGTTAATCAGTTGTTGTCTATCAAGCCAGAAGATCCAGCTAGATTATATGTGTTG CAAGACTTCCTGTCTTACATGAAGGGTTTCATCTCATTGCCAATTCCATTGCCAGGTACTGGTTAT ACCAATGCTATTAAGGCTAGAAAGAGATTGTCCGCTAGAGTTATGGGTATGATCAAAGAAAGGGA ACGTGAAGAAGAGGATATGAACAACGCTATCAGAGAAGAGGACTTCTTGGATTCCATTATCTCCA ATGAGGACTTGAACTACGAAGAGAAGGTTTCCATCGTTTTGGACATTTTGTTAGGTGGTTTTGAA ACTTCTGCTACCACCTTGTCTTTGGTTGTTTACTTTTTGGCTAAGAGCCCAAACCTGTTGCACAAG TTGAAAGAAGAACACGCTGCTATTAGAGCCAAAAAAGGTGATGGTGAACTGTTGAATTGGGAAG ATTACCAGAAGATGGAATTCACCCAATGCGTTATTTCTGAAGCTTTGAGATGCGAATACGTCATT CCAAAAGGTTGGAAGGTTTTCCCAATTTTCACTGCCGTTCATTTGGATCCATCATTGCACGAAAA TCCATTCGAATTCAATCCTATGAGATGGACCGATAAGGCCAAAATGAACAAAAAGACTACTGCTT TTGGTGGTGGTGTTAGAGTTTGTCCAGGTGGTGAATTGGGTAAATTGCAAATTGCTTTCTTCTTG

	CACCACTTGGTCTTGTCTTATAGATGGAAGATTAAGTCCGACGAAATGCCAATTGCTCATCCATA CGTTGAGTTTAAGAGAGGTATGTTGTTGGAAATCGAACCCACTAAGTTCTTGGAGGATTGA
CYP716A1	ATGTACATGGCAATCATGATTATCCTCTTCCTTAGCTCCATTCTTCTTTCTCTCCTATTGCTCCTTA GAAAACATTTGTCACACTTCTCCTATCCCAACCTTCCTCCCGGAAATACCGGCCTTCCCTTAATC GGAGAGAGTTTCTCCTTCCTCTCTGCTGGCCGTCAAGGCCATCCAGAGAAGTTCATCACTGACC GAGTTCGTCGTTTCTCCTCCTCCTCCTCATGTGTCTTCAAGACCCACCTCTTTGGGTCTCCCACC GCGGTGGTGACTGGTGCATCTGGGAACAAGTTTCTATTCACTAACGAGAACAAGCTTGTGGTCT CGTGGTGGCCAGATTCCGTTAATAAGATCTTCCCTTCTTCAATGCAGACGAGCTCCAAAGAAGAA GCCAGGAAGCTGAGGATGCTTCTTTCGCAGTTCATGAAGCCCGAGGCTTTGAGGAGGTATGTTG GTGTTATGGATGAGATTGCTCAAAGACATTTTGAGACGGAATGGGCCAATCAAGATCAAGTCATT GTCTTCCCTCTTACCAAAAAGTTCACGTTTTCAATAGCATGCCGTTCGTTCCTGAGCATGGAAGA TCCCGCAAGAGTAAGGCAACTAGAAGAGCAATTCAATACGGTAGCGGTAGGGATCTTCTCAATC CCAATAGACTTGCCAGGAACACGGTTTAACCGAGCCATCAAGGCGTCGAGGTTACTCAGAAAAG AGGTTTCCGCTATCGTAAGGCAGAGGAAAGAAGAGCTCAAAGCCGGGAAAGCATTAGAGGAGC ACGACATATTATCTCACATGTTGATGAATATAGGAGAGACCAAAGACGAGGATTTGGCTGATAAA ATTATTGGATTGTTAATCGGAGGACATGACACAGCTAGTATCGTATGCACTTTCGTTGTCAATTAT CTTGCTGAATTCCCTCATGTCTACCAACGTGTGCTACAAGAGCAAAAGGAGATACTAAAGGAGAA AAAAGAAAAGGAAGGATTAAGGTGGGAGGACATTGAGAAAATGAGATATTCATGGAATGTTGCAT GTGAAGTGATGAGAATTGTTCCTCCTCTTTCTGGCACTTTTCGTGAGGCCATTGATCACTTCTCTT TTAAGGGTTTTTACATTCCCAAAGGATGGAAGTTATATTGGAGTGCCACCGCGACACATATGAAT CCAGACTACTTCCCAGAACCAGAGAGATTTGAGCCAAACCGTTTCGAAGGAAGTGGTCCGAAGC CTTATACCTATGTTCCATTTGGAGGAGGACCAAGAATGTGTCCAGGGAAAGAGTATGCTCGTCTA GAGATTCTTATATTCATGCACAATCTTGTTAATAGATTTAAGTGGGAAAAAGTGTTTCCAAATGAA AATAAAATAGTTGTTGATCCCTTACCAATACCAGACAAAGGTCTCCCTATAAGAATTTTTCCTCAA TCTTGA
CYP716A2	ATGTACCTGACGATCATCTTCCTGTTCATCTCTTCCATTATCTTCCCTCTGTTGTTCTTTTTGGGTA AGCACTTGTCTAACTTCAGATACCCAAATTTGCCACCAGGTAAGATTGGTTTTCCATTGATTGGT GAAACCCTGTCTTTTTTGTCTGCTGGTAGACAAGGTCATCCAGAAAAGTTTGTTACCGATAGAGT TAGGCATTTCTCCTCCGGTATTTTCAAAACTCATTTGTTCGGTTCTCCATTCGCTGTTGTTACTGG TGCTTCTGGTAACAAGTTTTTGTTCACCAACGAAAACAAGCTGGTTATTTCTTGGTGGCCAGATT CCGTTAACAAGATTTTTCCTTCTTCTACCCAAACCTCCTCCAAAGAAGAAGCTATTAAGACCAGAA TGCTGCTGATGCCATCTATGAAGCCTGAAGCTTTGAGAAGATATGTTGGTGTTATGGACGAAATT GCCCAAAAGCACTTTGAAACTGAATGGGCCAATCAAGACCAATTGATTGTCTTTCCACTGACCAA GAAGTTCACCTTCTCTATTGCTTGTAGGCTGTTCTTGTCTATGGATGATTTGGAAAGAGTCCGTA AGTTGGAAGAACCATTCACTACTGTTATGACCGGTGTTTTCTCTATCCCTATTGATTTGCCAGGTA CTAGATTCAACAGAGCCATTAAGGCTTCCAGGTTGTTGTCTAAAGAAGTCTCCACCATTATCAGG CAGAGGAAAGAAGAATTGAAAGCCGGTAAAGTTTCCGTCGAACAAGATATTTTGTCCCACATGTT GATGAACATCGGTGAAACTAAGGATGAAGATTTGGCCGATAAGATTATCGCCTTGTTAATTGGTG GTCATGATACCACTTCTATTGTCTGTACTTTCGTCGTTAACTACTTGGCTGAATTCCCACATATCT ACCAAAGGGTTTTAGAAGGTATGCAAATCCCCTTGTTGTAA
CYP302A1	ATGCTGACCAAGTTGTTGAAGATTTCTTGCACCTCTAGACAATGTACTTTCGCTAAACCATACCAAG CTATTCCAGGTCCAAGAGGTCCATTTGGTATGGGTAACTTGTATAATTACTTGCCAGGTATCGGTTC CTACTCTTGGTTGAGATTGCATCAAGCTGGTCAAGATAAGTACGAAAAGTATGGTGCTATCGTCAGA GAAACTATCGTACCAGGTCAAGACATAGTTTGGTTGTACGATCCAAAGGATATTGCCTTGTTGTTGA ACGAAAGAGATTGCCCACAAAGAAGATCTCATTTGGCTTTGGCTCAGTACAGAAAATCTAGACCAG ATGTTTACAAGACCACTGGTTTGTTGCCAACTAATGGTCCAGAATGGTGGCGTATTAGAGCACAAGT TCAAAAAGAATTGTCCGCTCCAAAGTCCGTTAGAAACTTTGTTAGACAAGTTGACGGTGTCACCAAA GAATTCATCAGATTCCTACAAGAATCCAGAAACGGTGGTGCTATTGATATGTTGCCAAAGTTGACTA GGTTGAACTTGGAATTGACTTGCTTGTTGACTTTCGGTGCTAGATTGCAAAGTTTCACTGCTCAAGA ACAAGACCCAAGATCAAGATCTACTAGATTGATGGATGCTGCTGAAACTACCAACTCTTGTATTTTG CCAACCGACCAAGGTTTACAACTTTGGAGATTTTTGGAAACCCCATCCTTCAGAAAATTGTCCCAAG CTCAATCTTACATGGAATCCGTTGCTTTGGAATTGGTTGAAGAAAACGTCAGAAATGGTTCCGTTGG TTCCTCTTTGATTTCTGCTTATGTTAAGAACCCAGAGTTGGATAGATCTGATGTTGTTGGTACTGCTC CTGATTTGTTGTTGGCTGGTATTGATACAACTTCTTACGCTTCTGCTTTCCTGTTGTACCATATTGCT AGAAACCCAGAAGTCCAACAAAAGTTGCATGAAGAAGCTAGAAGAGTTTTGCCATCTGCTAAGGAT GAATTGTCTATGGATGCTTTGAGAACCGATATTACTTACACCAGAGCCGTCTTGAAAGAATCCTTGA GATTGAATCCAATTGCCGTTGGTGTTGGTAGAATTTTGAATCAAGATGCCATCTTCTCCGGTTACTT TGTTCCAAAAGGTACTACTGTTGTCACCCAAAACATGGTTGCTTGTAGATTGGAACAACACTTCCAA GATCCATTGAGATTCCAACCAGATAGATGGTTGCAACATAGATCTGCTTTGAACCCATATTTGGTCT TGCCTTTTGGTCATGGTATGAGAGCTTGTATTGCAAGAAGATTGGCTGAACAAAACATGCACATCTT

	GTTGCTAAGGTTGTTGAGGGAATACGAATTGATTTGGTCTGGTTCTGATGACGAAATGGGTGTTAA GACCTTGTTGATTAACAAACCTGATGCTCCAGTTTTGATCGACTTGAGATTAAGAAGGGAATAG
CYP90B27	ATGGCTATGGAACTGTTGTTGTTGATTCCAGCTTTCATCGTTGCCATCATCATTTTCTTCAGCTTC AAGTCTACCAACGGTACTTCTACTAAGCCATTGAAATTGCCACCAGGTCAAATGGGTTGGCCTTT TATTGGTCATACAATCCCTTTTATGCAACCCCATTCTTCTGCTTCATTGGGTCCATATATTGATTT GAACACTGCCAGATACGGCACCATTTTTAGAATGAATTTGTTGGCCAAGCCAACCATCGTTTCTG CTGATCCAGAATTCAATAGGTACATCTTGCAAAACGAAGGCAGGTTGTTTGAAAACTCTTGTCCA ACTTCCATTGCCGAAATTATGGGTAGATGGTCTATGTTGGCTTTGACTGGTGATGTTCATAGAGA AATGAGATCTATCGCCGTGTCTTTTATGTCCAACGTTAAGTTGAGAACCTACTTCATCGGTGACA TTGAACAACAAGCTATTAAGGTTTTGGCTTCTTGGGCTGGTAGAGATGCTCCATTTTCTGCTCAA GATGAAGGTAAGAAGTTCGCCTTTAACTTGATGGTTAAGCACCTGATGTCTATGGAACCAGGTAT GAAGGAAACTGAACAGTTGAGATCTGAATACCACGCTTTTATGAAAGGCATGGCTTCTATTCCAA TCAACTTGCCAGGTACTGCTTATAGAAAGGCCTTGCAATCTAGATCCATCATCTTGAAGATTATG GGCGAAAAATTGGACGAAAGAATCAAGCAAGTAAAAGAGGGTTGCGAAGGTTTGGAACAAGATG ATTTGTTAGCCTCTGTTTCTAAGCACCCAAACTTGGCCAAAGAACAAATCTTGGATCTGATCCTGT CAATGTTGTTTGCTGGTCACGAAACTTCATCTGCTGCTATTGCTTTGGCTATCTACTTCTTGGAAT CATGTCCAAAGGCTGTCGAACAATTGAGAGAAGAACACAAAGAAATCGCCAGGCAAAAGAAAGA AAGAGGTGAAACTGGTTTGAACTGGGATGACTACAAAAAGATGGAATTCACCCATTGCGTTATCA ACGAAACTTTGAGAATGGGTAACATCGTCAAGTTCTTGCATAGAAGGGCTATCAAGGATGTTCAA TTCAAGGGTTACGATATTCCATGTGGTTGGGAAGTTGTTCCAATTATTTCTGCTGCTCACTTGGA CTCTTCTATCTATGATGATCCACAAAGATACGATCCTTGGAGATGGCAAGCTATTTTGGCTGGTA ATACCAAGAACAACAACGTCACCTCTATCATGTCTTTTTCAGGTGGTCCAAGATTGTGTCCAGGT GCTGAATTGGCTAAATTGGAAATTGCTGTTTTCTTGCATCACCTGGTCCAAAAGTATCAATGGGA AATGGCTGAACATGACTACCCAGTTTCTTTTCCATTTTTGGGTTTCCCAAAGAGGTTGCCAATCAA AGTTAGACCATTGGGTGATTAA
CYP90B3	ATGTCTGACTTAGAGTTTTTTCTTTTTCTTATTCCTCCAATCTTAGCAGTACTTATAATTCTTAATCT ATTCAAAAGAAAACACAACTTTCAAAATCTTCCACCAGGGGATATGGGTTGGCCTTTTCTTGGTG AAACTATTGGTTATTTGAGACCTTATTCAGCTACTACTATTGGAGATTTCATGCAAGATCATATTTC TAGGTATGGGAAAATTTTCAAGTCAAATTTGTTTGGAGAGCCAACAATAGTTTCAGCAGATGCAG GGCTAAACAGATACATTCTGCAGAATGAAGGGAGATTATTTGAGTGTAATTATCCAAGAAGTATA GGTGGGATACTTGGTAAATGGTCTATGTTAGTTCAAGTTGGACAAATGCATAGAGATATGAGGAT GATTTCTCTGAATTTTTTGAGCAATGCTAGGCTAAGGAATCAACTTTTAAGTGAAGTTGAAAAGCA TACTTTGCTTGTTCTTGGCTCTTGGAAACAGGATTCTGTTGTTTGTGCACAAGATGAAGCAAAGA AGTTAACATTCAACTTTATGGCAGAGCATATCATGAGTCTACAACCTGGAAATCCAGAGACAGAG AAGCTGAAAAAAGAGTACATCACATTTATGAAAGGAGTGGTTTCTGCTCCATTGAATTTTCCAGG AACAGCTTACAGAAAGGCCTTACAGTCTCGATCAACAATTCTTGGATTTATTGAGAGAAAAATGG AGGAGAGGCTTAAGGAAATGAACAGAAACGAAAACGACCTTCTAGGTTGGGTTCTGAAGAATTC AAATCTCTCAAAAGAGCAAATTCTTGATTTGCTACTGAGTTTGCTCTTTGCTGGCCATGAAACTTC ATCAGTAGCAATAGCTCTGTCTATTTTCTTACTCGAAAGCTGTCCTGCTGCTGTTCAACAATTAAC AGAAGAGCACTTGGAGATTTCCCGGGCAAAAAAACAGTCAGGAGAAACAGAATTAAATTGGGAT GACTACAAGAAAATGGAATTCACCCAATGTGTTATTAATGAGACTCTAAGACTTGGGAATGTAGT GAGGTTTCTGCACAGGAAAGCTGTGAAAGATGTTCGATATAAAGGTTATGATATTCCATGTGGAT GGAAAGTGTTGCCGGTGATTTCAGCAGCGCATTTAGATCCTTCACTTTTCGACCGACCTCACGA CTTTGATCCTTGGAGATGGCAGAACGCAGAAGAGTCGCCTTCAGGTAAAGGAGGAAGCACAGG CACAAGCAGCACAACAAAAAGTAGTAATAATTTCATGCCATTTGGGGGAGGTCCACGTCTATGTG CAGGATCTGAACTGGCCAAACTTGAGATGGCCATTTTCATTCACTATCTTGTTCTTAATTTTCACT GGAAATTAGCTGCAACTGATCAGGCTTTTGCCTATCCTTACGTAGATTTTCCCAATGCCCTACCT ATCAATATCCAACATCGATCGTTAAATAAATTACATGACTAG

CYP724B2	ATGGGTGAAGAAGGTAGTCTTTTGATAATTGTTATCACCCTAGTTTTTAGTTTTGTAATTGGCATAA CATTAAACCATTTTTGGCCTTTGTTCTTCAATAATTATGGTACTACTCTTCATGTTATTCCCAAGGG CACTTTTGGATGGCCTTTACTTGGTGAAACCCTTTCCTTTTTGAAGCCTCATCCTTCTAATTCTAT TGGTACTTTCCTTCAACAACATTGTTCTAGGTATGGGAAAGTGTTCAAGTCACATTTATTTTTCTC CCCAACAGTGGTGTCATGTGACCAAGACCTTAATTACTTCATATTACAAAACGAAGATAAGTTATT TCAGTGTAGTTATCCAAAGCCAATTCATGGTATACTTGGCAAAGTTTCATTGCTTGTGGCTGTTG GCGACACACATAAAAGGCTTAGGAATGTTTCATTATCACTAATCAGCACCATTAAGTCTAAACCT GAGTTTATTAATGATGTTGAAACATTAGCACTTCAGATTCTCCAATCATGGAAAGATAAACATCAA GTCAGATACTGGGAAGAGGCAAGAAAGTTTTCATTCAATGTGATAGTGAAGCAAGTACTTGGATT AACTCCAGATAATCCACAAAGTGCATTAATTCTTCAAGATTTTCTTGCTTTTATGAGAGGATTAATT TCTTTACCATTATACATACCTGGAACTCCATATGCAAGAGCAGTGCAGGCTAGAAGTAGAATATC TTCAACTATCAAAGCAATTATAGAGGAAAGAAGAAGAAAACATGTTGTTGATGGTGATGGTAAAA AAAATGATTTTCTAGAGATACTTCTTTGTGTTGATACCTTATCTGAAGAGGAAAAAGTTAGCTTTG TTCTTGATTCTTTACTTGGTGGTTATGAGACCACTTCTCTCCTTATGGCTATGGTTGTCTTTTTTAT TGGTCAATCACAAACTGCTTTCGATCGACTCAAGGAAGAGCATGATAACATACGAAGCACGAAG GAAAAGGAGTTATTGAATTGGGAAGATTACCAAAAGATGGACTTCACTCAAAAGGTAATAAATGA AGCTTTAAGATATGGGAATGTTGTCAAATTTGTGCACCGAAAGGCACTTAAAGATGTCAAATTTAA AGATTATGTGATTCCAGCGGGTTGGAAGGTCCTACCAGTATTCAGTGCTGTTCATTTGGACCCAT CAGTTCATCCTAATGCACTCCACTTTAATCCTTGGAGATGGGAGAGTGATGAGCAAATAAGCAAG AAGTTGACTCCTTTTGGGGGAGGATCAAGATGTTGTCCTGGATTTGAACTTGCAAAGGTTGAAGT AGCCTTCTTCCTTCACCACCTTGTACAAAAATACAGATGGGAGGTAGAAGAAGGAGAACAACCC ATTGCTTATCCATATGTGGAGTTCAAAAATGGTTTAACAATCAGACTTCATAAAAATTCTACTTAG

Reference:

1. Ohnishi, T.; Watanabe, B.; Sakata, K.; Mizutani, M., CYP724B2 and CYP90B3 function in the early C-22 hydroxylation steps of brassinosteroid biosynthetic pathway in tomato. *Biosci Biotechnol Biochem* **2006,** *70* (9), 2071-80.

2. Chen, Q.; Man, C.; Li, D.; Tan, H.; Xie, Y.; Huang, J., Arogenate Dehydratase Isoforms Differentially Regulate Anthocyanin Biosynthesis in Arabidopsis thaliana. *Molecular plant* **2016,** *9* (12), 1609-1619.

3. Augustin, M. M.; Ruzicka, D. R.; Shukla, A. K.; Augustin, J. M.; Starks, C. M.; O'Neil-Johnson, M.; McKain, M. R.; Evans, B. S.; Barrett, M. D.; Smithson, A.; Wong, G. K.; Deyholos, M. K.; Edger, P. P.; Pires, J. C.; Leebens-Mack, J. H.; Mann, D. A.; Kutchan, T. M., Elucidating steroid alkaloid biosynthesis in Veratrum californicum: production of verazine in Sf9 cells. *The Plant journal : for cell and molecular biology* **2015,** *82* (6), 991-1003.

4. Yamazaki, Y.; Kiuchi, M.; Takeuchi, H.; Kubo, T., Ecdysteroid biosynthesis in workers of the European honeybee Apis mellifera L. *Insect Biochem Mol Biol* **2011,** *41* (5), 283-93.

5. Yasumoto, S.; Fukushima, E. O.; Seki, H.; Muranaka, T., Novel triterpene oxidizing activity of Arabidopsis thaliana CYP716A subfamily enzymes. *FEBS Lett* **2016,** *590* (4), 533-40.

6. Galanie, S.; Thodey, K.; Trenchard, I. J.; Filsinger Interrante, M.; Smolke, C. D., Complete biosynthesis of opioids in yeast. *Science* **2015,** *349* (6252), 1095-100.

7. Li, Y.; Smolke, C. D., Engineering biosynthesis of the anticancer alkaloid noscapine in yeast. *Nat Commun* **2016,** *7*, 12137.

Appendix II

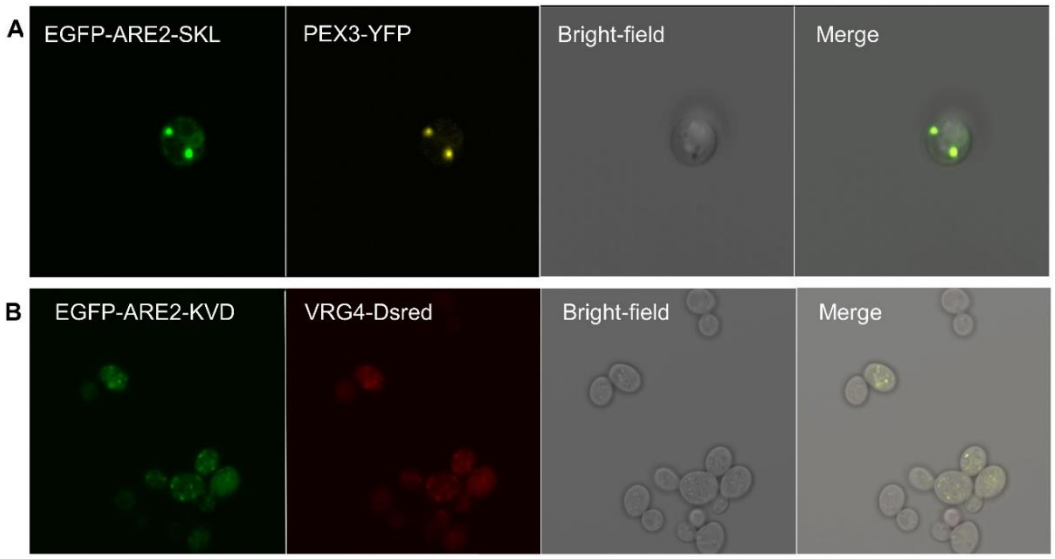


Figure S2.1. Localization of ARE2-SKL and ARE2-KVD by confocal microscopy imaging. (A) EGFP was fused to the N-terminal of ARE2-SKL and co-expressed with PEX3-YFP in YYL67. (B) EGFP was fused to the N-terminal of ARE2-KVD and co-expressed with VRG4-Dsred in YYL67.

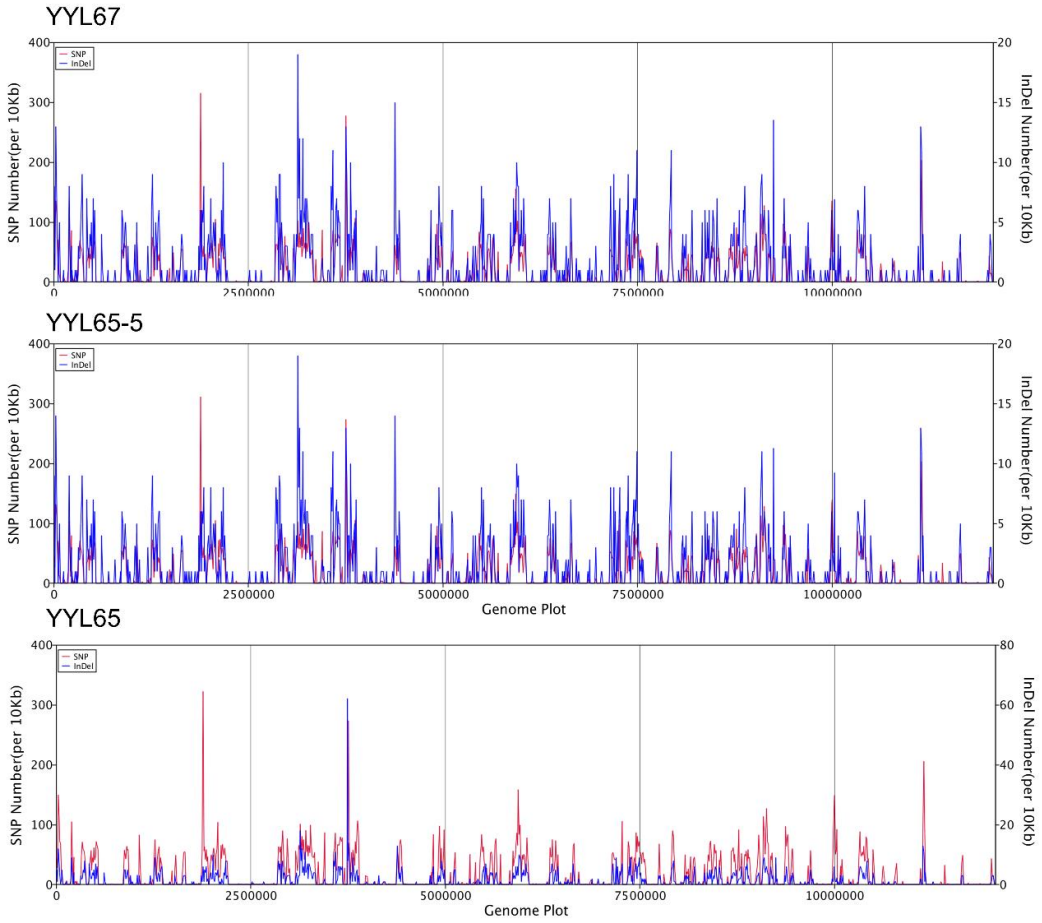


Figure S2.2. Distribution of InDels and SNPs through the genomes of YYL67, YYL65-5 and YYL65.

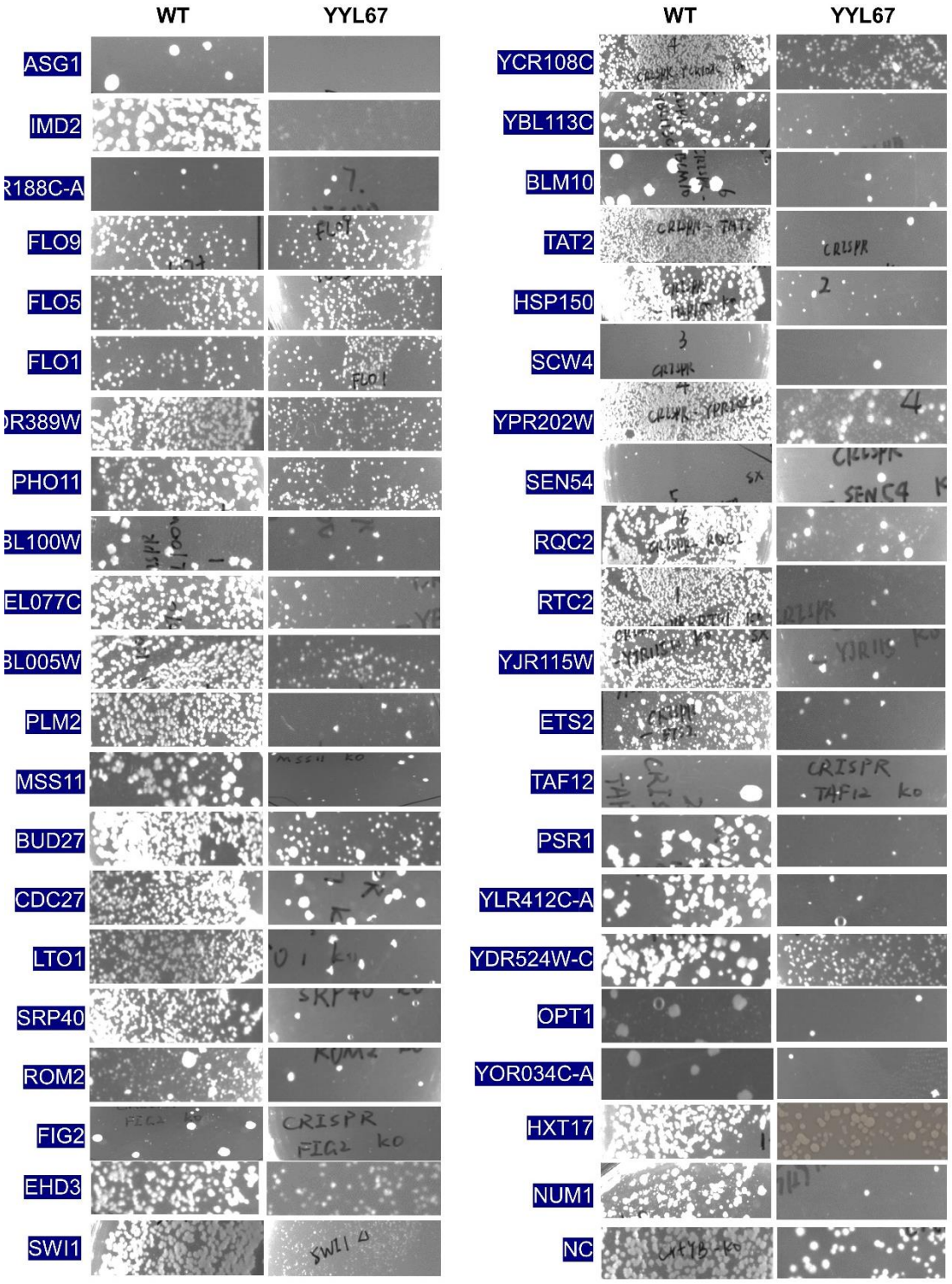


Figure S2.3. Inactivation of genes screened from InDel and SNP analysis in wildtype (WT) and YYL67.

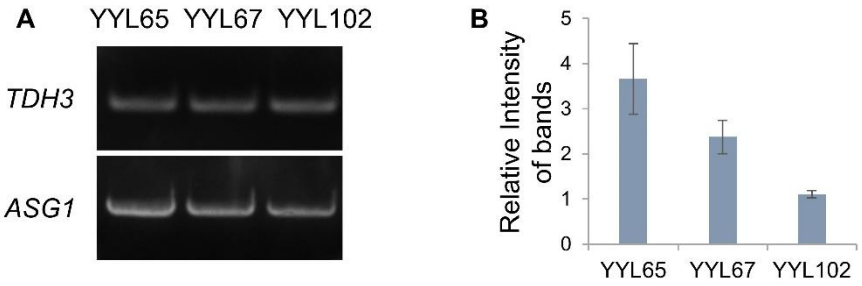


Figure S2.4. Expression of *ASG1* in different campesterol-producing strains. (A) RT-PCR of *ASG1* in YYL65, YYL67 and YYL102. *TDH3* is a constitutively expressed protein in yeast and used as an internal standard for quantification. (B) The bar graph represents the intensity of bands shown in (A). Error bars represent standard deviation of three biological replicates.

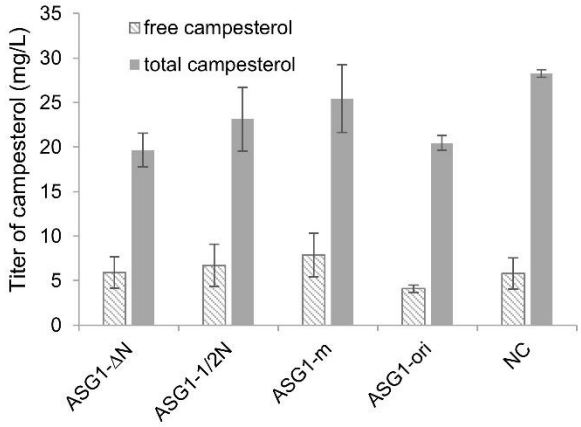


Figure S2.5. Production of total and free campesterol in YYL102 expressing ASG1 mutants. Free campesterol was extracted by *Extraction Method A* without saponification and total campesterol was extracted by *Extraction Method B* with KOH saponification. Error bars represent standard deviation of three biological replicates.

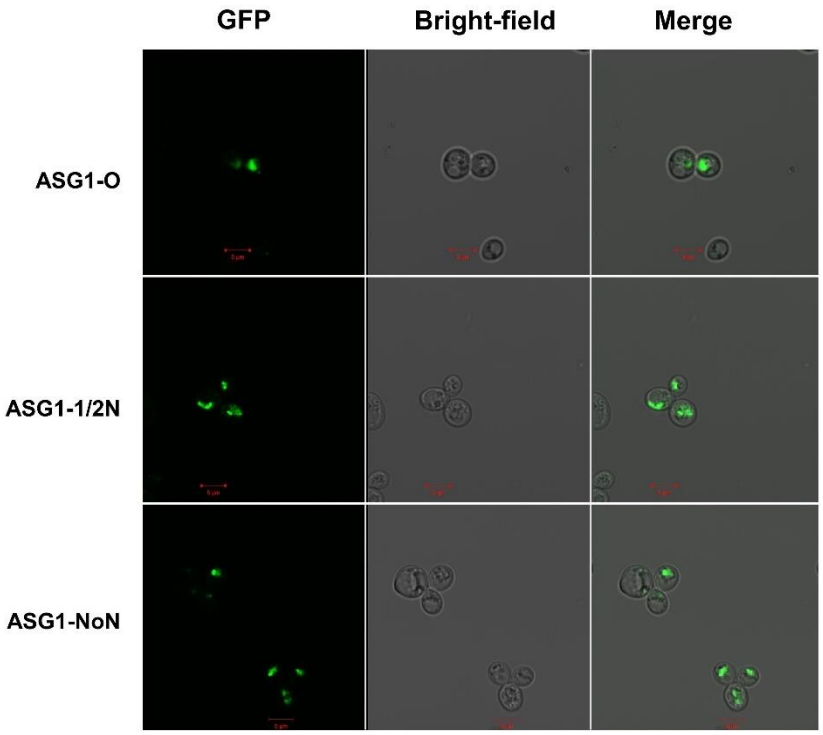


Figure S2.6. Localization and protein aggregation of $ASG1^{ori}$, $ASG1^{1/2N}$ and $ASG1^{\Delta N}$ in yeast.

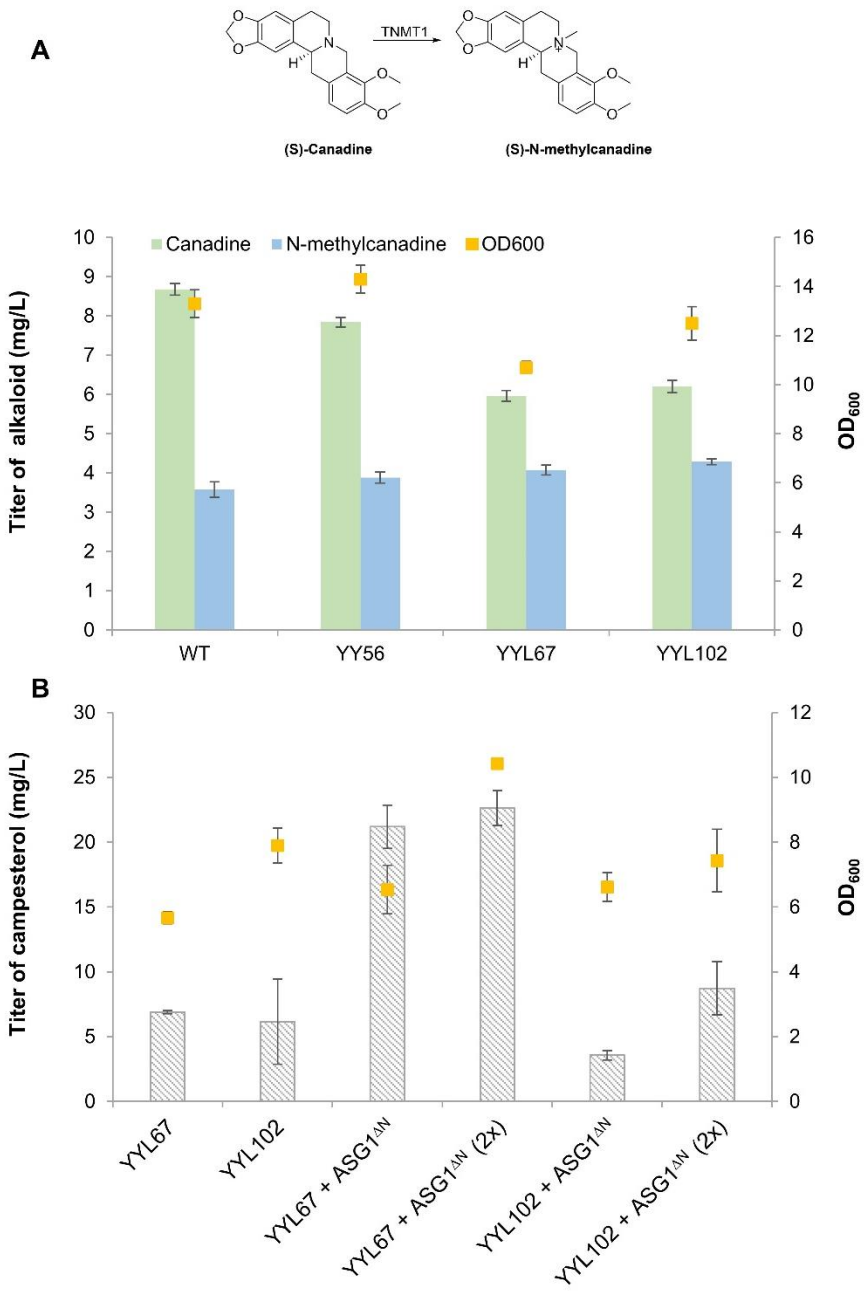


Figure S2.7. (A) Production of N-methylcanadine by TNMT1 in WT, YYL56, YYL67, YYL102 fed with canadine. Error bars represent standard deviation of three biological replicates. (B) The production of free campesterol in YYL67, YYL102, YYL67 expressing $ASG1^{\Delta N}$, YYL102 expressing $ASG1^{\Delta N}$ cultured in standard SDM; YYL67 expressing $ASG1^{\Delta N}$ cultured in 2 × SDM, and YYL102 expressing $ASG1^{\Delta N}$ cultured in 2 × SDM. Error bars represent standard deviation of three biological replicates.

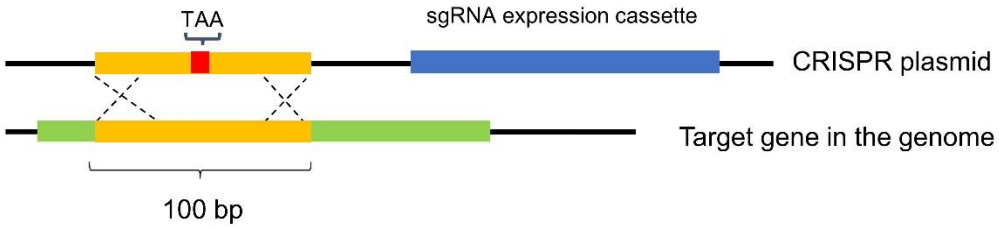


Figure S2.8. Scheme of the CRISPR/CAS9 construct for gene inactivation.

Table S2.1. Genotypes of yeast strains utilized in this study.

Strain	Genotype
YYL56	*leu2*Δ::P_{PYK1}-*dwf1*-T_{MFA1}, P_{TPI1}-*atr1*-T_{STE2}, P_{PGK1}-*dwf5*-T_{PHO5}, P_{GPD}-*dwf7*-T_{CYC1} *ybl059w*Δ::P_{PGK1}-*erg12*-T_{PHO5}, P_{TEF1}-*erg10*-T_{CYC1}, P_{PYK1}-*thmg1*-T_{MFA1}, P_{TPI1}-*erg13*-T_{STE2}
YYL63	*are1*Δ, *are2*Δ
YYL65	*leu2*Δ::P_{PYK1}-*dwf1*-T_{MFA1}, P_{TPI1}-*atr1*-T_{STE2}, P_{PGK1}-*dwf5*-T_{PHO5}, P_{GPD}-*dwf7*-T_{CYC1} *ybl059w*Δ::P_{PGK1}-*erg12*-T_{PHO5}, P_{TEF1}-*erg10*-T_{CYC1}, P_{PYK1}-*thmg1*-T_{MFA1}, P_{TPI1}-*erg13*-T_{STE2} *are1*Δ, *are2*Δ *erg4*Δ::*ura3*
YYL67	*leu2*Δ::P_{PYK1}-*dwf1*-T_{MFA1}, P_{TPI1}-*atr1*-T_{STE2}, P_{PGK1}-*dwf5*-T_{PHO5}, P_{GPD}-*dwf7*-T_{CYC1} *ybl059w*Δ::P_{PGK1}-*erg12*-T_{PHO5}, P_{TEF1}-*erg10*-T_{CYC1}, P_{PYK1}-*thmg1*-T_{MFA1}, P_{TPI1}-*erg13*-T_{STE2} *are1*Δ, *are2*Δ *erg4*Δ
YYL102	*leu2*Δ::P_{PYK1}-*dwf1*-T_{MFA1}, P_{TPI1}-*atr1*-T_{STE2}, P_{PGK1}-*dwf5*-T_{PHO5}, P_{GPD}-*dwf7*-T_{CYC1} *ybl059w*Δ::P_{PGK1}-*erg12*-T_{PHO5}, P_{TEF1}-*erg10*-T_{CYC1}, P_{PYK1}-*thmg1*-T_{MFA1}, P_{TPI1}-*erg13*-T_{STE2} *ymr206w*Δ::P_{TEF1}-*erg20-A99G*-T_{CYC1}, *HIS*, P_{PYK1}-*are2*-T_{MFA1} *are1*Δ, *are2*Δ *erg4*Δ

Table S2.2. Plasmids used in this study (P, promoter; T, terminator)

Plasmids #	Genotype	Reference
pAG413GPD-ccdB	*Centromeric HIS3*, *attR1*-P_{GPD}-*ccdB-attR2*	1*
pAG414GPD-ccdB	*Centromeric TRP1*, *attR1*-P_{GPD}-*ccdB-attR2*	1*
pAG415GPD-ccdB	*Centromeric LEU2*, *attR1*-P_{GPD}-*ccdB-attR2*	1*
pAG416GPD-ccdB	*Centromeric URA3*, *attR1*-P_{GPD}-*ccdB-attR2*	1*
pAG414GPD-ccdB-EGFP	*Centromeric TRP1*, *attR1*-P_{GPD}-*ccdB-EGFP-attR2*	1*
pAG414GPD-EGFP-ccdB	*Centromeric TRP1*, *attR1*-P_{GPD}-*EGFP-ccdB-attR2*	1*
pAG416GPD-ccdB-EYFP	*Centromeric URA3*, *attR1*-P_{GPD}-*ccdB-EYFP-attR2*	1*
pAG416GPD-ccdB-Dsred	*Centromeric URA3*, *attR1*-P_{GPD}-*ccdB-Dsred-attR2*	1*
pYL18	*Centromeric TRP1*, *attR1-ccdB-attR2*	2*
pYL660	*attL1*-P_{PGK1}-*ARE1*-T_{pho5}-*attL2*	This work
pYL659	*attL1*-P_{PGK1}-*ARE2*-T_{pho5}-*attL2*	This work
pYL657	*attL1*-P_{PGK1}-*ARE1-SKL*-T_{pho5}-*attL2*	This work
pYL658	*attL1*-P_{PGK1}-*ARE2-SKL*-T_{pho5}-*attL2*	This work
pYL1207	*attL1*-P_{PGK1}-*ARE1-KVD*-T_{pho5}-*attL2*	This work
pYL1208	*attL1*-P_{PGK1}-*ARE2-KVD*-T_{pho5}-*attL2*	This work
pYL1200	*attL1*-P_{PGK1}-*SAT1*-T_{pho5}-*attL2*	This work
pYL225	*attL1*-P_{TEF1}-*ERG20*-T_{CYC1}-*attL2*	This work
pYL1211	*attL1*-P_{TEF1}-*ERG20-A99G*-T_{CYC1}-*attL2*	This work
pYL645	*attL1*-P_{TEF1}-*ERG20*^{F96C}-T_{CYC1}-*attL2*	This work
pYL91	*attL1*-P_{TEF1}-*BTS1*-T_{CYC1}-*attL2*	This work
pYL1651	*attL1- ASG1*^{ori}*-attL2*	This work
pYL1218	*attL1- ASG1*^{m}*-attL2*	This work
pYL1692	*attL1-ASG1*$^{\Delta N}$*-attL2*	This work
pYL1693	*attL1-ASG1*$^{1/2N}$*-attL2*	This work
pYL1694	*attL1-IMD2*^{ori}*-attL2*	This work
pYL1695	*attL1-IMD2*^{m}*-attL2*	This work
pYL1696	*attL1- TNMT1-attL2*	This work
pYL6	*attL1*-P_{TPI1}-*CYP82Y1*-T_{Ste2}-*attL2*	This work
pYL1407	*attL1-PEX3-attL2*	This work
pYL1406	*attL1-VRG4-attL2*	This work
pYL1228	CRISPRa-*crtYB* inactivation	This work

1* Galanie, S., Thodey, K., Trenchard, I. J., Interrante, M. F., & Smolke, C. D. (2015). Complete biosynthesis of opioids in yeast. Science, 349(6252), 1095-1100.

2* Li, Y., & Smolke, C. D. (2016). Engineering biosynthesis of the anticancer alkaloid noscapine in yeast. Nature communications, 7, 12137.

Table S2.3. Summary of genomic sequencing information

(A) Sequencing quality information.

Strain	Average depth	Coverage ≥1X	Coverage ≥4X	Coverage ≥10X	Coverage ≥20X	Map rate
YYL65	112	98.54	98.28	97.57	95.71	98.27
YYL65-5	71	98.57	98.46	98.3	98.05	98.43
YYL67	53	98.55	98.43	98.23	97.87	98.13

(B) Category of genes with unique InDels in YYL67 and YYL65-5 comparing with YYL65

Transcription		**Transporter**	**Signaling**	**Stress Response**	**Others**		
taf12	*cyc8*	*chc1*	*psr1*	*rtc1*	*srp40*	*flo1*	*lto1*
mss11	*cdc39*	*sal1*	*arg82*	*cos12*	*est2*	*flo9*	*rom2*
swi1	*hmo1*	*opt1*	*ctk1*	*asg1*	*sen54*	*cdc27*	*cbf5*
pdr1			*hrk1*		*rqc2*	*hsm3*	*gid8*
					bud27	*fig2*	*bni1*

(C) SNPs in ORFs

	YYL67	YYL65-5
total SNPs	842	839
SNPs in ORFs	337	286
Genes with SNPs (at amino acid level)	19	22

Table S2.4. Summary of genes with InDels in adapted *S. cerevisiae* strains

Strains	Genes	Descriptions of genes
	EST2	Reverse transcriptase subunit of the telomerase holoenzyme
	RTC1	Subunit of SEACAT, a subcomplex of the SEA complex; complex that associates dynamically with the vacuole
	RQC2	Ribosome Quality control Complex; Component of RQC, which mediates nascent chain degradation
	SEN54	Subunit of the tRNA splicing endonuclease; tRNA splicing endonuclease
	FLO1	Lectin-like protein involved in flocculation
	FLO9	Lectin-like protein involved in flocculation
	YIL134C-A	uncharacterized protein
	YJR115W	uncharacterized protein
	YOR034C-A	uncharacterized protein
InDel found in both YYL67 and YYL65-7	CDC27	involved in Cell Division Cycle
	HSM3	involved in DNA mismatch repair during slow growth
	FIG2	Cell wall adhesin, expressed specifically during mating
	TAF12	TATA binding protein-Associated Factor; involved in RNA polymerase II transcription initiation and in chromatin modification
	BUD27	involved in translation initiation; mutants have inappropriate expression of nutrient sensitive genes
	OPT1	OligoPeptide Transporter ; Proton-coupled oligopeptide transporter of the plasma membrane; also transports glutathione and phytochelatin; member of the OPT family
	PSR1	Plasma membrane-associated protein phosphatase; involved in the general stress response
	MSS11	Transcription factor; involved in regulation of invasive growth and starch degradation; controls the activation of FLO11 and STA2 in response to nutritional signals;
	LTO1	Substrate-specific adaptor protein involved in apo-Rli1p maturation;
	SWI1	Subunit of the SWI/SNF chromatin remodeling complex; regulates transcription by remodeling chromatin; required for transcription of many genes, including ADH1, ADH2, GAL1, HO, INO1 and SUC2
	SRP40	Nucleolar serine-rich protein; role in preribosome assembly or transport
	ROM2	Guanine nucleotide exchange factor (GEF) for Rho1p and Rho2p
	ASG1	Activator of Stress Genes; Zinc cluster protein proposed to be a transcriptional regulator; regulator involved in the stress response; null mutants have a respiratory deficiency, calcofluor white sensitivity and slightly increased cycloheximide resistance
	YCR108C	uncharacterized protein

	YDR524W-C	uncharacterized protein
	YER188C-A	uncharacterized protein
	YEL077C	uncharacterized protein
	YLR412C-A	uncharacterized protein
YYL65-7	CYC8	General transcriptional co-repressor
	CDC39	involved in regulation of transcription and destabilization of mRNA by deadenylation; basal transcription factor that increases initiation and elongation
	ARG82	diphosphoinositol polyphosphate synthase activity; regulates arginine-, phosphate-, and nitrogen-responsive genes
	HMO1	Chromatin associated high mobility group (HMG) family member
	COS12	Endosomal protein involved in turnover of plasma membrane proteins; required for the multivesicular vesicle body sorting pathway that internalizes plasma membrane proteins for degradation
	PDR1	Transcription factor that regulates the pleiotropic drug response; zinc cluster protein
	CHC1	Clathrin heavy chain; subunit of the major coat protein involved in intracellular protein transport and endocytosis
	CTK1	Catalytic (alpha) subunit of C-terminal domain kinase I (CTDK-I)
	CBF5	Pseudouridine synthase catalytic subunit of box H/ACA snoRNPs; acts on large and small rRNAs, on snRNA U2, and on some mRNAs
	GID8	Subunit of GID Complex, binds strongly to central component Vid30p; GID Complex is involved in proteasome-dependent catabolite inactivation
	BNI1	nucleates the formation of linear actin filaments, involved in cell processes such as budding and mitotic spindle orientation
	SAL1	ADP/ATP transporter
	HRK1	Protein kinase; implicated in activation of the plasma membrane H(+)-ATPase Pma1p in response to glucose metabolism; plays a role in ion homeostasis; protein abundance increases in response to DNA replication stress
	YLL066W-B	uncharacterized protein
	YML100W-A	uncharacterized protein
	YOR034C-A	uncharacterized protein
	YOR012W	uncharacterized protein

Table S2.5. Summary of genes with SNPs in adapted *S. cerevisiae* strains

	Genes	Number of SNPs	Description of genes
YYL67	BLM10	1	proteasome activator
	HSP150	1	heat shock protein
	PLM2p	1	putative transcription factor
	TAT2	1	aromatic amino acid transmembrane transporter
	YPR202W	1	similar to telomere-encoded helicases
Genes with SNPs in both YYL67 and YYL65-5	FLO5	3	flocculation
	FLO9	3	flocculation
	PHO11	6	one of three repressible acid phosphatases
	EHD3	2	3-hydroxyisobutyryl-CoA hydrolase
	HXT17	1	hexose transporter HXT17
	NUM1	3	nuclear Migration
	IMD2	1	Inosine monophosphate dehydrogenase
	SCW4	1	putative family 17 glucosidase
	YBL100W-B	2	gag-pol fusion protein
	YBL005W-B	1	uncharacterized protein
	YBL113C	3	uncharacterized protein
	YEL077C	2	uncharacterized protein
	YOR389W	4	uncharacterized protein
	YCR108C	3	uncharacterized protein
YYL65-5	COS7	9	endosomal protein involved in turnover of plasma membrane proteins
	FLO1	2	flocculation
	BUL1	1	ubiquitin-ubiquitin ligase
	TIR4	2	Cell wall mannoprotein
	YHL008C	1	uncharacterized protein
	YHR219W	3	uncharacterized protein
	YAR064W	1	pseudogenic fragment with similarity to flocculins
	YPL277C	1	uncharacterized protein

Table S2.6. Sequences of sgRNA by CRISPR-Cas9 to inactivate genes screened by SNPs and InDels

Gene	sgRNA
ASG1	ATCAGATGAGAATATCACTG
BLM10	TGATTGTCGTCTATTCCCAG
BUD27	TAGAAAAAGAACCTGTTGTG
CDC27	CAGTTCGGCCAAAAACTCTG
EHD3	AACTCTAGCTGTATCCTGAA
EST2	ACGTAGTTAGAATTTCATCG
FIG2	CCGCCAATTTTAGCACACAA
FLO1	CAGTACCACTACAACTACAA
FLO5	AGTACCGGTCCATGGCTCGG
FLO9	ACTACAACGGAACCATGGAC
HSM3	*sgRNA expression cassette cannot be synthesized commercially
HSP150	ATCACCGATCTGAGAAACGG
HXT17	TCAGCTTGTTTGAGTACCTA
IMD2	AATACTTTTTAACTGTCATA
LTO1	AAAGACTTCAAATGTCACGG
MSS11	AGATTTCATAACTTTCTGAG
NUM1	TCGACTAGGTATGCTAGTGA
OPT1	GTGGCACTGACATTCACATG
PHO11	ACTTATTATCAGACAGGACC
PLM2p	CGAGATTGAATTAGAAGCAT
PSR1	GAGTGTTGTAACCTGGTGCA
ROM2	AATGGCGCACATTATCAAGT
RQC2	ATTAAAAGTGAACAAAAAAA
RTC1	AAAAAATTAAATGCTCACAC
SCW4	TATACCCCATACGAGTCCAG
SEN54	CTTGGAAATCAGGAAGACCG
SRP40	GGAATCCGAATCTGAATCCG
SWI1	TGATAACAATAACAGTAACA
TAF12	ACTGCCACCAATAATCCCAG
TAT2	AGATGGCTACCCAAATTACG
YBL005W-B	ATAAGATGTCTTTAACCCAG
YBL100W-B	CTTGCCAACCTTGAGTACGA
YBL113C	AGACCTGTCTGCTGACACAG
YCR108C	ATTGATCGGGCATACGCACA

YDR524W-C	AGGAAAAAAAATATGTTGGA
YEL077C	TCGTTCCTGCAATAAAGTGG
YER188C-A	GCAAAGGGTATAGACCGCTG
YIL134C-A	*sgRNA score too low
YJR115W	AATAGACGAAATCGACAGCG
YLR412C-A	AAAAAAACGGGTCCTAAATG
YOR034C-A	TCAAAGGTGCCTCGGTCCGA
YOR389W	CCACCAGACGAAAAAGATCG
YPR202W	CTTTTCGTTTTCAATTTCCA

Table S2.7. Amino acid sequences of ASG1 and IMD2 variants used in this study

Gene	Amino acid sequence
ASG1^{ori}	MPEQAQQGEQSVKRRRVTRACDECRKKKVKCDGQQPCIHCTVYSYECTYK KPTKRTQNSGNSGVLTLGNVTTGPSSSTVVAAAASNPNKLLSNIKTERAILPG ASTIPASNNPSKPRKYKTKSTRLQSKIDRYKQIFDEVFPQLPDIDNLDIPVFLQI FHNFKRDSQSFLDDTVKEYTLIVNDSSSPIQPVLSSNSKNSTPDEFLPNMKSD SNSASSNREQDSVDTYSNIPVGREIKIILPPKAIALQFVKSTWEHCCVLLRFYH RPSFIRQLDELYETDPNNYTSKQMQFLPLCYAAIAVGALFSKSIVSNDSSREK FLQDEGYKYFIAARKLIDITNARDLNSIQAILMLFIFLQCSARLSTCYTYIGVAMR SALRAGFHRKLSPNSGFSPIEIEMRKRLFYTIYKLDVYINAMLGLPRSTSPDDF DQTLPLDLSDENITEVAYLPENQHSVLSSTGISNEHTKLFLILNEIISELYPIKKT SNIISHETVTSLELKLRNWLDSLPKELIPNAENIDPEYERANRLLHLSFLHVQIIL YRPFIHYLSRNMNAENVDPLCYRRARNSIAVARTVIKLAKEMVSNNLLTGSY WYACYTIFYSVAGLLFYIHEAQLPDKDSAREYYDILKDAETGRSVLIQLKDSSM AASRTYNLLNQIFEKLNSKTIQLTALHSSPSNESSSLVTNNSSALKPHLRDSLQ PPVFFSSQDTKNSFSLAKSEESTNDYAMANYLNNTPISENPLNEAQQQDQVS QGTTNMSNERDPNNFLSTDIRLDNNGQSNILDATDDVFIRNDGDIPTNSAFDF SSSKSNASNNSNPDTINNYNNVSGKNNNNNNITNNSNNNHNNNNNDNNNN NNNNNNNNNNNNNNNNNNNNSGNSSNNNNNSNNNKNNNNDFGIKIDNNSP SYEGFPQLQIPLSQDNLNIEDKEEMSPNIEIKNEQNMTDSNDILGVFDQLDAQ LFGKYLPLNYPSE*
ASG1^{m}	MPEQAQQGEQSVKRRRVTRACDECRKKKVKCDGQQPCIHCTVYSYECTYK KPTKRTQNSGNSGVLTLGNVTTGPSSSTVVAAAASNPNKLLSNIKTERAILPG ASTIPASNNPSKPRKYKTKSTRLQSKIDRYKQIFDEVFPQLPDIDNLDIPVFLQI FHNFKRDSQSFLDDTVKEYTLIVNDSSSPIQPVLSSNSKNSTPDEFLPNMKSD SNSASSNREQDSVDTYSNIPVGREIKIILPPKAIALQFVKSTWEHCCVLLRFYH RPSFIRQLDELYETDPNNYTSKQMQFLPLCYAAIAVGALFSKSIVSNDSSREK FLQDEGYKYFIAARKLIDITNARDLNSIQAILMLFIFLQCSARLSTCYTYIGVAMR SALRAGFHRKLSPNSGFSPIEIEMRKRLFYTIYKLDVYINAMLGLPRSISPDDF DQTLPLDLSDENITEVAYLPENQHSVLSSTGISNEHTKLFLILNEIISELYPIKKT SNIISHETVTSLELKLRNWLDSLPKELIPNAENIDPEYERANRLLHLSFLHVQIIL YRPFIHYLSRNMNAENVDPLCYRRARNSIAVARTVIKLAKEMVSNNLLTGSY WYACYTIFYSVAGLLFYIHEAQLPDKDSAREYYDILKDAETGRSVLIQLKDSSM AASRTYNLLNQIFEKLNSKTIQLTALHSSPSNESSSLVTNNSSALKPHLGDSLQ PPVFFSSQDTKNSFSLAKSEESTNDYAMANYLNNTPISENPLNEAQQQDQVS QGTTNMSNERDPNNFLSTDIRLDNNGQSNILDATDDVFIRNDGDIPTNSAFDF SSSKSNASNNSNPDTINNYNNVSGKNNNNNNITNNSNNNHNNNNNDNNNN NNNNNNNNNNNNNNNNNNNNSGNSSNNNNNSNNNKNNNNDFGIKIDNNSPS YEGFPQLQIPLSQDNLNIEDKEEMSPNIEIKNEQNMTDSNDILGVFDQLDAQL FGKYLPLNYPSE*

ASG1$^{\Delta N}$	MPEQAQQGEQSVKRRRVTRACDECRKKKVKCDGQQPCIHCTVYSYECTYK KPTKRTQNSGNSGVLTLGNVTTGPSSSTVVAAAASNPNKLLSNIKTERAILPG ASTIPASNNPSKPRKYKTKSTRLQSKIDRYKQIFDEVFPQLPDIDNLDIPVFLQI FHNFKRDSQSFLDDTVKEYTLIVNDSSSPIQPVLSSNSKNSTPDEFLPNMKSD SNSASSNREQDSVDTYSNIPVGREIKIILPPKAIALQFVKSTWEHCCVLLRFYH RPSFIRQLDELYETDPNNYTSKQMQFLPLCYAAIAVGALFSKSIVSNDSSREK FLQDEGYKYFIAARKLIDITNARDLNSIQAILMLFIFLQCSARLSTCYTYIGVAMR SALRAGFHRKLSPNSGFSPIEIEMRKRLFYTIYKLDVYINAMLGLPRSISPDDF DQTLPLDLSDENITEVAYLPENQHSVLSSTGISNEHTKLFLILNEIISELYPIKKT SNIISHETVTSLELKLRNWLDSLPKELIPNAENIDPEYERANRLLHLSFLHVQIIL YRPFIHYLSRNMNAENVDPLCYRRARNSIAVARTVIKLAKEMVSNNLLTGSY WYACYTIFYSVAGLLFYIHEAQLPDKDSAREYYDILKDAETGRSVLIQLKDSSM AASRTYNLLNQIFEKLNSKTIQLTALHSSPSNESSSLVTNNSSALKPHLGDSLQ PPVFFSSQDTKNSFSLAKSEESTNDYAMANYLNNTPISENPLNEAQQQDQVS QGTTNMSNERDPNNFLSTDIRLDNNGQSNILDATDDVFIRNDGDIPTNSAFDF SSSKSNASNNSNPDTINNYNNVSGKDFGIKIDNNSPSYEGFPQLQIPLSQDNL NIEDKEEMSPNIEIKNEQNMTDSNDILGVFDQLDAQLFGKYLPLNYPSE*
ASG1^{½N}	MPEQAQQGEQSVKRRRVTRACDECRKKKVKCDGQQPCIHCTVYSYECTYK KPTKRTQNSGNSGVLTLGNVTTGPSSSTVVAAAASNPNKLLSNIKTERAILPG ASTIPASNNPSKPRKYKTKSTRLQSKIDRYKQIFDEVFPQLPDIDNLDIPVFLQI FHNFKRDSQSFLDDTVKEYTLIVNDSSSPIQPVLSSNSKNSTPDEFLPNMKSD SNSASSNREQDSVDTYSNIPVGREIKIILPPKAIALQFVKSTWEHCCVLLRFYH RPSFIRQLDELYETDPNNYTSKQMQFLPLCYAAIAVGALFSKSIVSNDSSREK FLQDEGYKYFIAARKLIDITNARDLNSIQAILMLFIFLQCSARLSTCYTYIGVAMR SALRAGFHRKLSPNSGFSPIEIEMRKRLFYTIYKLDVYINAMLGLPRSISPDDF DQTLPLDLSDENITEVAYLPENQHSVLSSTGISNEHTKLFLILNEIISELYPIKKT SNIISHETVTSLELKLRNWLDSLPKELIPNAENIDPEYERANRLLHLSFLHVQIIL YRPFIHYLSRNMNAENVDPLCYRRARNSIAVARTVIKLAKEMVSNNLLTGSY WYACYTIFYSVAGLLFYIHEAQLPDKDSAREYYDILKDAETGRSVLIQLKDSSM AASRTYNLLNQIFEKLNSKTIQLTALHSSPSNESSSLVTNNSSALKPHLGDSLQ PPVFFSSQDTKNSFSLAKSEESTNDYAMANYLNNTPISENPLNEAQQQDQVS QGTTNMSNERDPNNFLSTDIRLDNNGQSNILDATDDVFIRNDGDIPTNSAFDF SSSKSNASNNSNPDTINNYNNVSGKNNNNNNNITNNSNNNHNNNNNDNNNN NNNNNNNNNNNNDFGIKIDNNSPSYEGFPQLQIPLSQDNLNIEDKEEMSPNIE IKNEQNMTDSNDILGVFDQLDAQLFGKYLPLNYPSE*
IMD2^{ori}	MAAIRDYKTALDFTKSLPRPDGLSVQELMDSKIRGGLTYNDFLILPGLVDFAS SEVSLQTKLTRNITLNIPLVSSPMDTVTESEMATFMALLGGIGFIHHNCTPEDQ ADMVRRVKNYENGFINNPIVISPTTTVGEVKSMKEKYGFAGFPVTEEGKRNA RLVGVITSRDIQFVEDNSLLVQDVMTKNPVTGAQGITLSEGNEILKKIKKGRLL VVDEKGNLVSMLSRTDLMKNQNYPLASKSANTKQLLCGASIGTMDADKERL RLLVKAGLDVVILDSSQGNSIFELNMLKWVKESFPGLEVIAGNVVTREQAANLI AAGADGLRIGMGTGSICITQEVMACGRPQGTAVYNVCEFANQFSVPCMADG GVQNIGHIIKALALGSSTVMMGGMLAGTTESPGEYFYQDGKRLKAYRGMGSI DAMQKTGTKGNASTSRYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGLQ HSCQDIGCKSLSLLKENVQRGKVRFEFRTASAQLEGGVHNLHSYEKRLHN*
IMD2^{m}	MAAIRDYKTALDFTKSLPRPDGLSVQELMDSKIRGGLTYNDFLILPGLVDFAS SEVSLQTKLTRNITLNIPLVSSPMDTVTESEMATFMALLGGIGFIHHNCTPEDQ ADMVRRVKNYENGFINNPIVISPTTTVGEAKSMREKYGFAGFPVTEDGKRNA KLVGVITSRDIQFVEDNSLLVQDVMTKNPVTGAQGITLSEGNEILKKIKKGRLL VVDEKGNLVSMLSRTDLMKNQNYPLASKSANTKQLLCGASIGTMDADKERL RLLVKAGLDVVILDSSQGNSIFELNMLKWVKESFPGLEVIAGNVVTREQAANLI AAGADGLRIGMGTGSICITQEVMACGRPQGTAVYNVCEFANQFGVPCMADG

	GVQNIGHITKALALGSSTVMMGGMLAGTTESPGEYFYQDGKRLKAYRDMGS IDAMQKTGTKGNASTSRYFSESDSVLVAQGVSGAVVDKGSIKKFIPYLYNGL QHSCQDIGCKSLSLLKENVQRGKVRFEFRTASAQLEGGVHNLHSYEKRLHN*

Appendix III

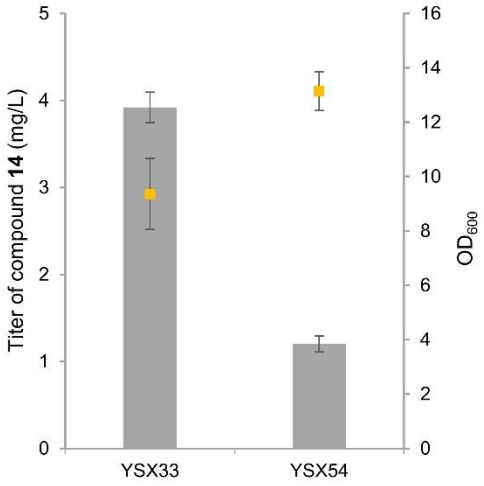


Figure S3.1 Production of 22-hydroxy-campest-4-en-3-one, **14**, in YYL102, the optimized campesterol-producing strain. Bars represent mean values of three biological replicates, and the error bars represent the standard deviation of the replicates.

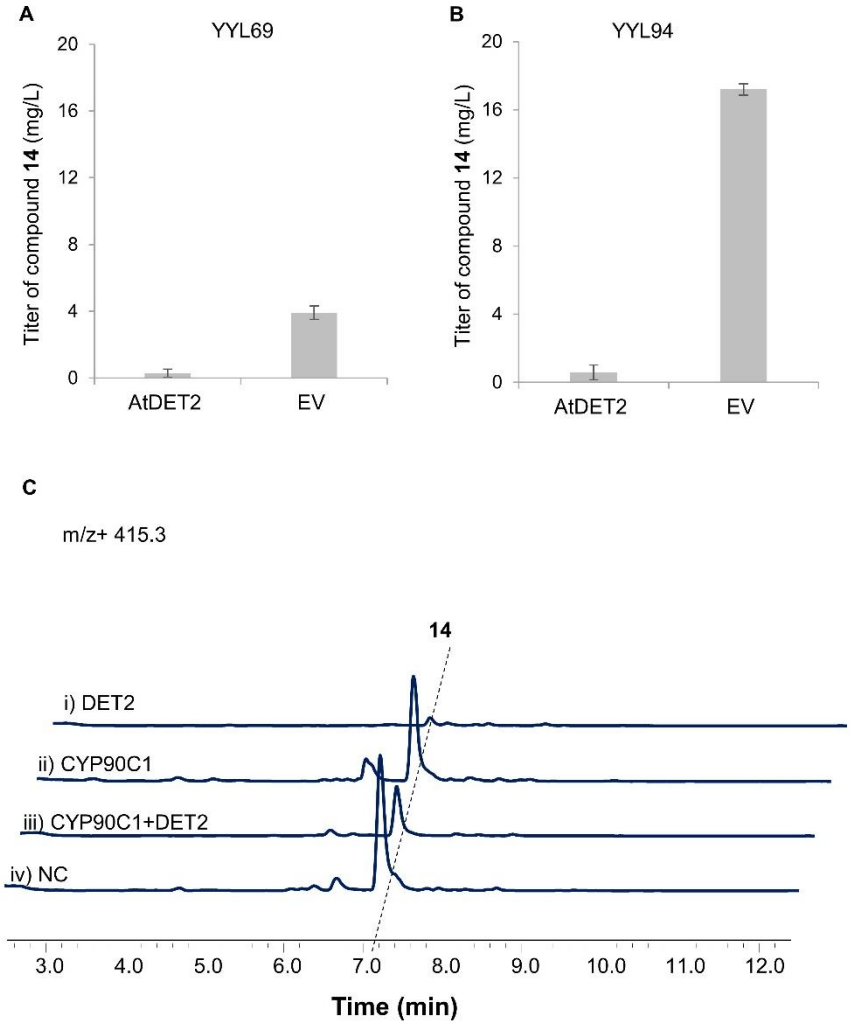


Figure S3.2 Effects of DET2 on 22-hydroxycampest-4-en-3-one, **14**. (A) Production of **14** in YYL69 expressing DET2 and the empty vector (EV). (B) Production of **14** in YYL94 expressing DET2 and the empty vector (EV). (C) SIM EIC using 22-hydroxycampest-4-en-3-one, **14**'s characteristic m/z^+ signal (MW=414.67 Da, $[C_{28}H_{47}O_3]^+$=415.3) of YYL94 expressing i) DET2, ii) CYP90C1, iii) DET2 and CYP90C1, iv) EV. The results are repeatable. Bars represent mean values of three biological replicates, and the error bars represent the standard deviation of the replicates.

Table S3.1 Strains used in Chapter III

Strain	Genotype
CENPK2.1D	*MATα, ura3-52, trp1-289, leu2-3,112 his3Δ1, MAL2-8^{C}, SUC2*
YYL67	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, P_{GPD}-dwf7-T_{CYC1}* *ybl059wΔ::P_{PGK1}-erg12-T_{PHO5}, P_{TEF1}-erg10-T_{CYC1}, P_{PYK1}-tHmg1-T_{MFA1}, P_{TPI1}-erg13-T_{STE2}* *are1Δ, are2Δ* *erg4Δ*
YYL69	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, P_{GPD}-dwf7-T_{CYC1}* *ybl059wΔ::P_{PGK1}-erg12-T_{PHO5}, P_{TEF1}-erg10-T_{CYC1}, P_{PYK1}-tHmg1-T_{MFA1}, P_{TPI1}-erg13-T_{STE2}* *ybr197cΔ::P_{TPI1}-CYP90A1-T_{PHO5}, P_{GPD}-CYP90B1-T_{CYC1}* *are1Δ, are2Δ* *erg4Δ*
YYL94	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, P_{GPD}-dwf7-T_{CYC1}* *ybl059wΔ::P_{PGK1}-erg12-T_{PHO5}, P_{TEF1}-erg10-T_{CYC1}, P_{PYK1}-tHmg1-T_{MFA1}, P_{TPI1}-erg13-T_{STE2}* *ybr197cΔ::P_{TPI1}-CYP90A1-T_{PHO5}, P_{GPD}-CYP90B1-T_{CYC1}* *ymr206wΔ::P_{TEF1}-erg20-A99G-T_{CYC1}, P_{GPD1}-msbp1 -T_{ADH1}, HIS3, P_{PGK1}-are2-T_{PHO5}* *are1Δ, are2Δ* *erg4Δ*
YYL102	*leu2Δ::P_{PYK1}-dwf1-T_{MFA1}, P_{TPI1}-atr1-T_{STE2}, P_{PGK1}-dwf5-T_{PHO5}, P_{GPD}-dwf7-T_{CYC1}* *ybl059wΔ::P_{PGK1}-erg12-T_{PHO5}, P_{TEF1}-erg10-T_{CYC1}, P_{PYK1}-thmg1-T_{MFA1}, P_{TPI1}-erg13-T_{STE2}* *ymr206wΔ::P_{TEF1}-erg20-A99G-T_{CYC1}, HIS, P_{PYK1}-are2-T_{MFA1}* *are1Δ, are2Δ* *erg4Δ*

Table S3.2 Plasmids used in Chapter III

Plasmids #	Genotype	Reference
pAG413GPD-ccdB	*Centromeric HIS3*, *attR1*-P_{GPD}-*ccdB-attR2*	1*
pAG414GPD-ccdB	*Centromeric TRP1*, *attR1*-P_{GPD}-*ccdB-attR2*	1*
pAG415GPD-ccdB	*Centromeric LEU2*, *attR1*-P_{GPD}-*ccdB-attR2*	1*
pAG416GPD-ccdB	*Centromeric URA3*, *attR1*-P_{GPD}-*ccdB-attR2*	1*
pYL1231	*attL1-*P_{GPD}*-MSBP1-*T_{ADH1}*-attL2*	This work
pYL661	*attL1- MSBP1-attL2*	This work
pYL1199	*attL1-MSBP2-attL2*	This work
pYL1707	*attL1- MSBP1-ABC-attL2*	This work
pYL1709	*attL1- MSBP1-ABD-attL2*	This work
pYL1706	*attL1- MSBP1-BCD-attL2*	This work
pYL1705	*attL1- MSBP1-BC-attL2*	This work
pYL1716	*attL1- MSBP1-CD-attL2*	This work
pYL1715	*attL1- MSBP1-C-attL2*	This work
pYL584	*attL1- CYP90A1-attL2*	This work
pYL184	*attL1-CYP90B1-attL2*	This work
pYL585	*attL1-CYP90C1-attL2*	This work
pYL237	*attL1-CYP90D1-attL2*	This work
pYL236	*attL1-CYP85A1-attL2*	This work
pYL181	*attL1- CYP85A2-attL2*	This work
pYL586	*attL1-DET2-attL2*	This work

1* Galanie, S., Thodey, K., Trenchard, I. J., Interrante, M. F., & Smolke, C. D. (2015). Complete biosynthesis of opioids in yeast. Science, 349(6252), 1095-1100.

2*Li, Y., & Smolke, C. D. (2016). Engineering biosynthesis of the anticancer alkaloid noscapine in yeast. Nature communications, 7, 12137.

Table S3.3 Amino acid sequences of proteins used in this chapter

Gene	Amino acid sequences
CYP90A1	MAFTAFLLLLSSIAAGFLLLLRRTRYRRMGLPPGSLGLPLIGETFQLIGAYKTENP EPFIDERVARYGSVFMTHLFGEPTIFSADPETNRFVLQNEGKLFECSYPASICNLL GKHSLLLMKGSLHKRMHSLTMSFANSSIIKDHLMLDIDRLVRFNLDSWSSRVLL MEEAKKITFELTVKQLMSFDPGEWSESLRKEYLLVIEGFFSLPLPLFSTTYRKAIQ ARRKVAEALTVVVMKRREEEEEGAERKKDMLAALLAADDGFSDEEIVDFLVAL LVAGYETTSTIMTLAVKFLTETPLALAQLKEEHEKIRAMKSDSYSLEWSDYKSM PFTQCVVNETLRVANIIGGVFRRAMTDVEIKGYKIPKGWKVFSSFRAVHLDPNH FKDARTFNPWRWQSNSVTTGPSNVFTPFGGGPRLCPGYELARVALSVFLHRLVT GFSWVPAEQDKLVFFPTTRTQKRYPIFVKRRDFAT*
CYP90B1	MFETEHHTLLPLLLLPSLLSLLLFLILLKRRNRRTRFNLPPGKSGWPFLGETIGYL KPYTATTLGDFMQQHVSKYGKIYRSNLFGEPTIVSADAGLNRFILQNEGRLFECS YPRSIGGILGKWSMLVLVGDMHRDMRSISLNFLSHARLRTILLKDVERHTLFVLD SWQQNSIFSAQDEAKKFTFNLMAKHIMSMDPGEEETEQLKKEYVTFMKGVVSA PLNLPGTAYHKALQSRATILKFIERKMEERKLDIKEEDQEEEEVKTEDEAEMSKS DHVRKQRTDDDLLGWVLKHSNLSTEQILDLILSLLFAGHETSSVAIALAIFFLQA CPKAVEELREEHLEIARAKKELGESELNWDDYKKMDFTQCVINETLRLGNVVRF LHRKALKDVRYKGYDIPSGWKVLPVISAVHLDNSRYDQPNLFNPWRWQQQNN GASSSGSGSFSTWGNNYMPFGGGPRLCAGSELAKLEMAVFIHHLVLKFNWELA EDDKPFAFPFVDFPNGLPIRVSRIL*
CYP90C1	MQPPASAGLFRSPENLPWPYNYMDYLVAGFLVLTAGILLRPWLWLRLRNSKTK DGDEEEDNEEKKKGMIPNGSLGWPVIGETLNFIACGYSSRPVTFMDKRKSLYGK VFKTNIIGTPIIISTDAEVNKVVLQNHGNTFVPAYPKSITELLGENSILSINGPHQKR LHTLIGAFLRSPHLKDRITRDIEASVVLTLASWAQLPLVHVQDEIKKMTFEILVKV LMSTSPGEDMNILKLEFEEFIKGLICIPIKFPGTRLYKSLKAKERLIKMVKKVVEE RQVAMTTTSPANDVVDVLLRDGGDSEKQSQPSDFVSGKIVEMMIPGEETMPTA MTLAVKFLSDNPVALAKLVEENMEMKRRKLELGEEYKWTDYMSLSFTQNVINE TLRMANIINGVWRKALKDVEIKGYLIPKGWCVLASFISVHMDEDIYDNPYQFDP WRWDRINGSANSSICFTPFGGGQRLCPGLELSKLEISIFLHHLVTRYSWTAEEDEI VSFPTVKMKRRLPIRVATVDDSASPISLEDH*
CYP90D1	MDTSSSLLFFSFFFFIIIVIFNKINGLRSSPASKKKLNDHHVTSQSHGPKFPHGSLG WPVIGETIEFVSSAYSDRPESFMDKRRLMYGRVFKSHIFGTATIVSTDAEVNRAV LQSDSTAFVPFYPKTVRELMGKSSILLINGSLHRRFHGLVGSFLKSPLLKAQIVRD MHKFLSESMDLWSEDQPVLLQDVSKTVAFKVLAKALISVEKGEDLEELKREFEN FISGLMSLPINFPGTQLHRSLQAKKNMVKQVERIIEGKIRKTKNKEEDDVIAKDV VDVLLKDSSEHLTHNLIANNMIDMMIPGHDSVPVLITLAVKFLSDSPAALNLLTE ENMKLKSLKELTGEPLYWNDYLSLPFTQKVITETLRMGNVIIGVMRKAMKDVEI KGYVIPKGWCFLAYLRSVHLDKLYYESPYKFNPWRWQERDMNTSSFSPFGGGQ RLCPGLDLARLETSVFLHHLVTRFRWIAEEDTIINFPTVHMKNKLPIWIKRI*
CYP85A1	MGAMMVMMGLLLIIVSLCSALLRWNQMRYTKNGLPPGTMGWPIFGETTEFLK QGPNFMRNQRLRYGSFFKSHLLGCPTLISMDSEVNRYILKNESKGLVPGYPQSM LDILGTCNMAAVHGSSHRLMRGSLLSLISSTMMRDHILPKVDHFMRSYLDQWN ELEVIDIQDKTKHMAFLSSLTQIAGNLRKPFVEEFKTAFFKLVVGTLSVPIDLPGT NYRCGIQARNNIDRLLRELMQERRDSGETFTDMLGYLMKKEGNRYPLTDEEIRD QVVTILYSGYETVSTTSMMALKYLHDHPKALQELRAEHLAFRERKRQDEPLGLE DVKSMKFTRAVIYETSRLATIVNGVLRKTTRDLEINGYLIPKGWRIYVYTREINY DANLYEDPLIFNPWRWMKKSLESQNSCFVFGGGTRLCPGKELGIVEISSFLHYFV TRYRWEEIGGDELMVFPRVFAPKGFHLRISPY*

CYP85A2	MGIMMMILGLLVIIVCLCTALLRWNQMRYSKKGLPPGTMGWPIFGETTEFLKQG PDFMKNQRLRYGSFFKSHILGCPTIVSMDAELNRYILMNESKGLVAGYPQSMLDI LGTCNIAAVHGPSHRLMRGSLLSLISPTMMKDHLLPKIDDFMRNYLCGWDDLET VDIQEKTKHMAFLSSLLQIAETLKKPEVEEYRTEFFKLVVGTLSVPIDIPGTNYRS GVQARNNIDRLLTELMQERKESGETFTDMLGYLMKKEDNRYLLTDKEIRDQVV TILYSGYETVSTTSMMALKYLHDHPKALEELRREHLAIRERKRPDEPLTLDDIKS MKFTRAVIFETSRLATIVNGVLRKTTHDLELNGYLIPKGWRIYVYTREINYDTSL YEDPMIFNPWRWMEKSLESKSYFLLFGGGVRLCPGKELGISEVSSFLHYFVTKY RWEENGEDKLMVFPRVSAPKGYHLKCSPY*
DET2	MEEIADKTFFRYCLLTLIFAGPPTAVLLKFLQAPYGKHNRTGWGPTVSPPIAWFV MESPTLWLTLLLFPFGRHALNPKSLLLFSPYLIHYFHRTIIYPLRLFRSSFPAGKNG FPITIAALAFTFNLLNGYIQARWVSHYKDDYEDGNWFWWRFVIGMVVFITGMYI NITSDRTLVRLKKENRGGYVIPRGGWFELVSCPNYFGEAIEWLGWAVMTWSWA GIGFFLYTCSNLFPRARASHKWYIAKFKEEYPKTRKAVIPFVY*
MSBP1	MALELWQTLKEAIHAYTGLSPVVFFTALALAFAIYQVISGWFASPFDDVNRHQR ARSLAQEEEPPIPQPVQVGEITEEELKQYDGSDPQKPLLMAIKHQIYDVTQSRMF YGPGGPYALFAGKDASRALAKMSFEEKDLTWDISGLGPFELDALQDWEYKFMS KYAKVGTVKVAGSEPETASVSEPTENVEQDAHVTTTPEKTVVDKSDDAPAETV LKKEE*
MSBP1-ABC	MALELWQTLKEAIHAYTGLSPVVFFTALALAFAIYQVISGWFASPFDDVNRHQR ARSLAQEEEPPIPQPVQVGEITEEELKQYDGSDPQKPLLMAIKHQIYDVTQSRMF YGPGGPYALFAGKDASRALAKMSFEEKDLTWDISGLGPFELDALQDWEYKFMS KYAKVGTVK*
MSBP1-ABD	MALELWQTLKEAIHAYTGLSPVVFFTALALAFAIYQVISGWFASPFDDVNRHQR ARSLAQEEEPPIPQPVQVGVAGSEPETASVSEPTENVEQDAHVTTTPEKTVVDKS DDAPAETVLKKEE*
MSBP1-BCD	MVVFFTALALAFAIYQVISGWFASPFDDVNRHQRARSLAQEEEPPIPQPVQVGEI TEEELKQYDGSDPQKPLLMAIKHQIYDVTQSRMFYGPGGPYALFAGKDASRAL AKMSFEEKDLTWDISGLGPFELDALQDWEYKFMSKYAKVGTVKVAGSEPETAS VSEPTENVEQDAHVTTTPEKTVVDKSDDAPAETVLKKEE*
MSBP1-BC	MVVFFTALALAFAIYQVISGWFASPFDDVNRHQRARSLAQEEEPPIPQPVQVGEI TEEELKQYDGSDPQKPLLMAIKHQIYDVTQSRMFYGPGGPYALFAGKDASRAL AKMSFEEKDLTWDISGLGPFELDALQDWEYKFMSKYAKVGTVK*
MSBP1-CD	MEITEEELKQYDGSDPQKPLLMAIKHQIYDVTQSRMFYGPGGPYALFAGKDASR ALAKMSFEEKDLTWDISGLGPFELDALQDWEYKFMSKYAKVGTVKVAGSEPET ASVSEPTENVEQDAHVTTTPEKTVVDKSDDAPAETVLKKEE*
MSBP1-C	MEITEEELKQYDGSDPQKPLLMAIKHQIYDVTQSRMFYGPGGPYALFAGKDASR ALAKMSFEEKDLTWDISGLGPFELDALQDWEYKFMSKYAKVGTVK*