

**Exploring the production of high-value compounds in plant *Catharanthus roseus* hairy roots and yeast *Yarrowia lipolytica***

by

**Le Zhao**

A dissertation submitted to the graduate faculty  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Chemical Engineering

Program of Study Committee:  
Zengyi Shao, Major Professor

Laura Jarboe  
Reuben Peters  
Levi Stanley  
Thomas Mansell

Iowa State University

Ames, Iowa

2017

Copyright © Le Zhao, 2017. All rights reserved.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS .....	v
ABSTRACT .....	vi
CHAPTER I. INTRODUCTION .....	1
Products of interest .....	1
Production of TIAs in <i>C. roseus</i> .....	2
Regulation of the TIA pathway.....	6
Transcriptional and metabolic profiling of transgenic <i>C. roseus</i> hairy root lines .....	9
Wax esters .....	10
<i>Y. lipolytica</i> .....	12
Production of wax esters in <i>Y. lipolytica</i> .....	13
Dissertation organization .....	14
Table .....	16
Figures.....	17
References.....	23
CHAPTER II. EXPRESSION OF TABERSONINE 16-HYDROXYLASE AND 16-HYDROXYTABERSONINE- <i>O</i> -METHYL TRANSFERASE IN <i>CATHARANTHUS ROSEUS</i> HAIRY ROOTS .....	28
Abstract .....	28
Introduction.....	29
Materials and Methods.....	31
Plasmid construction and clone generation.....	31
Hairy roots culture and induction .....	32
Alkaloids extraction .....	32
Metabolite standard from yeast fermentation .....	33
Metabolites identification using LC/MS or LC/MS/MS .....	34
Metabolites quantification with HPLC .....	35
RT-qPCR.....	35
Statistical analysis.....	36
Results.....	36
Genetic engineering of T16H and 16OMT in <i>C. roseus</i> hairy root .....	36
Identification of new metabolites in <i>C. roseus</i> hairy roots overexpressing T16H and 16OMT.....	37

Quantification of alkaloids in the induced and uninduced hairy roots expressing T16H and 16OMT.....	39
Transcriptional changes of the TIA pathway genes and regulators after inducing T16H and 16OMT in <i>C. roseus</i> hairy roots .....	41
Discussion.....	42
Conclusion .....	45
Figures.....	47
References.....	56
<b>CHAPTER III. METABOLIC PROFILING OF <i>CATHARANTHUS ROSEUS</i> HAIRY ROOT LINES OVEREXPRESSING TRANSCRIPTION ACTIVATORS FOR TERPENOID INDOLE ALKALOID PATHWAY .....</b>	<b>59</b>
Introduction.....	59
Materials and Methods.....	61
Alkaloid extraction.....	61
Alkaloid analysis.....	62
Results and Discussions.....	63
Effects of ORCA2 overexpression on the metabolite levels of the TIA pathway .....	63
Effects of CrBPF1 overexpression on the metabolite levels of the TIA pathway .....	66
Figures.....	68
References.....	75
<b>CHAPTER IV. PRODUCTION OF WAX ESTERS IN YEAST <i>YARROWIA LIPOLYTICA</i> .....</b>	<b>76</b>
Introduction.....	76
Materials and methods .....	77
Strains, media, and cultivation methods .....	77
Gene deletion in <i>Yarrowia lipolytica</i> .....	77
Construction and transformation of plasmid and linear DNA .....	78
Random integration of genes into <i>Y. lipolytica</i> .....	79
Nile Red staining and lipid separation via thin layer chromatography (TLC) .....	80
Total lipid extraction and analysis .....	80
Lipid saponification and analysis.....	81
Copy number analysis.....	82
Results.....	82
Knocking out the substrate competitive pathways to increase acyl-CoA pool.....	82
Heterologous expression of the wax ester synthetic pathway in the four <i>Y. lipolytica</i> strains.....	84

Distributions of free fatty alcohol types and wax ester types produced in engineered strains.....	85
Random genome integration to avoid plasmid instability issue .....	86
Nitrogen limited fermentation to optimize culture medium for wax ester production ...	88
Discussions .....	89
Conclusion .....	90
Table .....	91
Figures.....	92
References.....	106
CHAPTER V. SUMMARY .....	109
References.....	110

## ACKNOWLEDGEMENTS

I would like to thank my two advisors, Professor Jacqueline Shanks and Professor Zengyi Shao, for guiding and helping me on those interesting research projects, as well as their support and encouragement during my PhD years.

I would like to thank ...

Reuben Peters, Laura Jarboe, Thomas Mansell, Levi Stanley for serving on my Program of Study committee and for all the insightful questions, comments and advice.

Chunyao Li, Sue Gibson, Jiayi Sun, Christie Peebles for great collaboration in the *Catharanthus roseus* projects.

Jiachen Zi for helping me on structural elucidation with NMR spectroscopy when I was preparing in-house standards of TIA metabolites.

Sibongile Mafu, Reuben Peters for providing yeast host to express plant genes.

Zhihong Song, Lucas Showman, Ann Perera for helping on the metabolite analysis in several projects.

James Yu, Suzanne Sandmeyer, Yu Jiang for providing plasmids, strains, and experiences in *Yarrowia lipolytica* project.

All the members in Zengyi Shao's and Jacqueline Shanks' groups, especially Mingfeng Cao, Jong Moon Yoon, and Meirong Gao, for discussing projects and helping me in my PhD years.

My friends, Xiaofei Hu, Yingxi Chen, Tao Jin, Yingfei Chen, for giving me great help and fun in my life.

My husband, Fuyuan Jing, very much for advising me on my projects and supporting me in my life.

## ABSTRACT

This dissertation focuses on studying the production of two categories of high-value compounds in bio organisms. The first group is terpenoid indole alkaloids (TIAs) in plant *Catharanthus roseus*, and the second group is wax esters, one of fatty acid derivatives.

TIAs belong to secondary metabolites in *C. roseus* and some of them have wide pharmaceutical applications. In particular, vinblastine and vincristine are two TIAs with anticancer properties and have been marked and used in chemotherapeutic reagents. The biggest issue for TIA production is that the content of those secondary metabolites in plant is extremely low. To improve the TIA production, we studied the regulation mechanism of the TIA pathway and explored the feasibility of valuable TIA production in hairy root culture.

Compared with the whole plant, plant tissue culture, such as hairy root culture, has many advantages, like fast growth, large-scale cultivation, and ease of genetic engineering. But the biggest issue for hairy root is that the vinblastine and vincristine synthetic pathway is blocked, mainly one of their precursors, vindoline, can not be synthesized in hairy root. To explore whether *C. roseus* hairy root could produce the intermediates in the vindoline pathway by overexpressing the pathway enzymes, we co-expressed the first two genes, tabersonine 16-hydrolase (T16H) and 16-*O*-methyltransferase (16OMT) in the vindoline pathway into hairy root. Transcriptional analysis and metabolic profiling were done to compare the difference between the parent hairy root lines and the engineered hairy root lines.

For the metabolic profiling, since the standards for those intermediates were not available, we prepared in-house standards by expressing the plant genes in *Saccharomyces cerevisiae*, fed substrate, and purified TIA compounds from yeast cell culture. Liquid

chromatography (LC) coupled with either photodiode array detector (PDA) or mass spectrometry (MS) were applied to isolate and analyze the TIA compounds by their UV-Visible absorption spectra and molecular weights.

In addition, fundamental research was done in *C. roseus* hairy root to study the effects of transcription regulators on the transcript levels and metabolite levels of the TIA pathway. Two of the seven reported transcription activators of the TIA pathway, octadecanoid-responsive *Catharanthus* AP2-domain 2(ORCA3) and MYB-like DNA-binding protein (BPF1), were overexpressed in hairy root separately, Two of the transcription repressors, G-box binding factors (GBF1 and GBF2), were knocked down by RNA interference in hairy root. And the transcription analysis and metabolic profiling of the transcription regulator-engineered hairy root lines were done to see what were the effects caused by those regulators.

Wax esters have a lot of applications in lubricant, skin care products, cosmetics, inking, and coating industries. Currently the main bio source for high-performed wax ester is from the seeds of jojoba. The tight supply makes wax esters high-value compounds. To reduce the production cost, we introduced the wax ester biosynthetic pathway into an oleaginous yeast, *Yarrowia lipolytica*. The free fatty alcohol, and wax ester were quantified in the engineered *Y. lipolytica*.

To provide more substrate for wax ester synthesis, we knocked out some genes in the substrate competitive pathways, and constructed four strains with different combination of knockout genes. To utilize the most abundant fatty acid types in *Y. lipolytica*, we compared three fatty acyl-CoA reductases (FAR), the first enzyme in the wax ester pathway, from different species. It was found that those three FARs had different substrate specificity and the wax ester production varied a lot in those strains.

To solve the plasmid instability issue, we randomly integrated FAR gene and WS gene into *Y. lipolytica* genome, and studied the effect of nitrogen limited fermentation on the wax ester production in one of our best strains.

## CHAPTER I. INTRODUCTION

### **Products of interest**

There are two main categories of compounds in my research projects, one is terpenoid indole alkaloids, belonging to the secondary metabolites in a medicinal plant, and the other is fatty acid derivatives, which are primary metabolites in a few organisms.

Terpenoid indole alkaloids (TIAs) contain a structural moiety of indole from amino acid tryptophan and a structural moiety of terpene from isoprene groups (Figure 1.1). Some of the TIAs have significant physiological activity. For examples, ajmaline found in most species of the *Rauwolfia* genus and *Catharanthus roseus* (*C. roseus*) is an antiarrhythmic and antihypertension agent, vinblastine and vincristine isolated from *C. roseus* are widely applied in chemotherapeutical reagents to treat Hodgkin's lymphoma, lung cancers and other types of diseases. The quantity of the TIA in plants is extremely low, which makes those medicinal compounds valuable. To improve the TIA production and reduce production costs, some of my collaborated research projects focus on transcription analysis and metabolite profiling of the TIA pathway in engineered *C. roseus* hairy root lines.

Fatty acids consist of a long hydrocarbonate chain and one or more carboxylic groups. Most of the natural fatty acids have the carbon chain lengths from 4 to 22, with 18 most common (Figure 1.2). Fatty acids are the main components of oil and fats. Fatty acids and its derivatives have overwhelming commercial uses, such as soap, waxes, polishes, leather treating, emulsifiers and so on. Wax esters are one type of high valued fatty acid derivatives, and the chemical structure is an ester of a fatty acid and a fatty alcohol. My studies focused on producing wax esters in a non-native host, oleaginous yeast *Yarrowia lipolytica* (*Y. lipolytica*).

### **Production of TIAs in *C. roseus***

Vinblastine and vincristine are the two most important TIAs in *C. roseus*. When researchers began to analyze *C. roseus* in the late 1950's, vincristine and vinblastine were found to be able to lower the number of white cells in blood. Through decades of research, the medicinal mechanism of vinblastine and vincristine has been well studied and these two compounds are currently widely applied in chemotherapies to treat many types of cancers. To decrease their toxicity and explore broader pharmacological applications, several semi-synthetic derivatives were synthesized, among which, vindesine and vinorelbine have been marketed. Together with vinblastine and vincristine, these four compounds have different pharmaceutical profiles and are used to treat different types of cancers [1]. Figure 1.3 shows the structure of these four vinca alkaloids.

*C. roseus* produces more than 130 different TIAs, and is the only source for obtaining vinblastine and vincristine. The biosynthetic pathway of TIAs in *C. roseus* is complex and highly branched (Figure 1.4). The isolation of vinblastine and vincristine from *C. roseus* is laborious and costly due to their relatively low levels in the plant. Approximately 500 kg of dried leaves are needed in order to isolate 1 gram of vinblastine [2]. Research has been focusing on production of vinca alkaloids either by fermentation of endophytic fungi in *C. roseus* or by large-scale culturing of plant suspension cells, hairy roots, and other tissues.

Plant endophytic fungi play important roles in plant secondary metabolite biosynthesis. Some endophytic fungi produce the same metabolites as their hosts. However no fermentation based on plant endophytic fungi has been commercialized due to the low and sometimes unstable yields. More than 183 endophytic fungi were isolated from different tissues of *C. roseus* and some of them were reported to possess the ability to produce either vinblastine or vincristine.

Among the fungi isolated from Chinese *C. roseus*, it was reported that one endophytic fungus, *Fusarium oxysporum*, produced vincristine, and another endophytic fungi species, *Alternaria sp.*, produced vinblastine. The *Fusarium oxysporum* isolated from Indian species produced 76  $\mu\text{g}$  vinblastine and 67  $\mu\text{g}$  vincristine in one liter of fermentation broth, and the dimers were characterized via HPLC, MS and  $^1\text{H}$  NMR [3]. For *Fusarium*, the biosynthetic pathways for the production of vinblastine and vincristine from tryptophan and geraniol have not been elucidated.

Biotechnology progress in plant *in vitro* cell and tissue culture provides alternatives to produce valuable secondary metabolites [4]. Suspension cells and hairy roots, being amenable to large-scale industrial application, are the two most in-depth studied cultures [5]. Plant suspension cells are biosynthetically totipotent with the potential to produce all the compounds in parent plants. Rao *et al.* listed the plant species whose cell cultures produce more secondary metabolites than the intact plants do [6]. The bottlenecks of cell culture development are low product yields, genetic instability, and metabolite variability [7]. Hairy roots in most cases grow with the similar rates as cell culture but with higher biochemical and genetic stability. In addition, since hairy roots are differentiated, compartmentation of portions of cellular pathways may assist in the production of toxic intermediates and metabolites. They keep most of the properties of the natural roots and contain metabolites synthesized in the roots. This distinction is important since transport of metabolites to and from other plant tissues is not present in a root culture [8].

*C. roseus* cell suspension and hairy root cultures do not produce the two bisindole alkaloids, vinblastine and vincristine, as a result of the inability to synthesize one of the monoterpene indole alkaloids (MIA), vindoline (Figure 1.4). Vindoline condenses with another MIA, catharanthine, forming the bisindole alkaloids. In suspension cell cultures, transcripts of *N*-methyltransferase (NMT), desacetoxyvindoline 4-hydroxylase (D4H) and deacetylvindoline

acetyltransferase (DAT) are absent and consequently the biosynthesis of vindoline is blocked [9-11]. Root cultures lack the idioblast and laticifer, which are specialized cell types for the late steps of vindoline synthesis, so vindoline cannot be produced in hairy root cultures as well. A comprehensive RNA-seq data from different tissues of *C. roseus*, including leaf, flower and root, also indicated that enzymes involved in the vindoline and vinblastine pathways are restricted to aerial tissues [12]. The intermediate metabolite tabersonine in hairy root culture accumulates and is converted via the lochnericine and hörhammericine branch, instead of the vindoline branch [13]. The high flux through tabersonine in hairy roots indicates the potential of this culture type for production of vindoline if an active tabersonine to vindoline pathway could be engineered into the roots.

Elicitation is a common strategy to increase secondary metabolite levels in intact plants, plant tissues and cell cultures. Among various elicitors, jasmonic acid, its conjugates and precursors, collectively known as jasmonates (JAs), are effective in activating the biosynthesis of many plant secondary metabolites through initiating plant defense mechanisms against pathogens and herbivores [14]. Various strategies to increase alkaloids levels, including adjusting medium, temperature, light, and aeration, have been reviewed [15,16]. Many plant phytohormones and elicitors, both biotic and abiotic, have been applied to increase alkaloids accumulation [16,17].

In microbial production systems, various metabolic engineering strategies are often applied to improve production. In contrast, plant metabolic engineering is still in its infancy and many toolkits are needed. Particle bombardment and *Agrobacterium*-mediated transformation are the two frequently used methods to introduce foreign DNA molecules, and protocols have been established for transforming *Catharanthus roseus* (vinca alkaloids producer) [18,19]. Metabolic engineering of *C. roseus* cell suspensions, hairy root cultures and even intact plants has been

studied with the aim of increasing TIA accumulation. Overexpression of transcription factors, overexpression of a TIA enzyme, or a combination of enzymes has been the main strategies to date. Precursor feeding strategies are often combined with these experiments in order to determine rate-limiting steps. These efforts often only result in modest improvement of TIA accumulation (Table 1.1) since transcriptional regulation is hierarchical and a large portion of the regulatory network is still unclear.

Besides production of valuable secondary metabolites in native organisms, heterologous synthesis, mainly in microbes, brings new opportunities and technical challenges. The advantages of heterologous synthesis involve: (1) ease of genetic manipulation and culture processing in heterologous hosts, (2) lack of complicated regulation belonging to the native organisms, and (3) simplified purification process without contamination by structurally similar metabolites that exist in the native organisms. Meanwhile, the challenges are also apparent. First of all, enabling a heterologous host for producing plant secondary metabolites usually requires the expression of multiple genes, and expressing plant-sourced proteins in microbial systems is sometimes challenging. Secondly, the accumulation of some intermediates caused by unbalanced gene expression may result in a toxicity issue. Furthermore, some metabolites require plant organelles for their synthesis, so that an alternative scaffold in the microbial host will have to be created. Regardless of these challenges, the rapid development in synthetic biology points out additional directions to tackle the challenges in this area, for example, Martin *et al.* [20] developed *E. coli* strains as platform hosts for the production of terpenoid compounds.

Heterologous expression of the TIA synthetic pathway in non-native plant hosts and microbial hosts is under development and will potentially pave the way to produce interesting TIAs from hosts that grow faster and are easier to be cultured. Most of the heterologous

expression efforts focused on tryptophan decarboxylase (TDC), strictosidine synthase (STR), and strictosidine  $\beta$ -d-glucosidase (SGD), the three enzymes in the upstream of TIA pathway (Figure 1.4). These three genes were expressed in tobacco suspension cells, tobacco plants, *Cinchona officinalis* hairy root culture, other plant species tissue cultures, yeast and *E. coli*, individually or in combination, with the precursor fed to the culture [17]. Table 1.1 also listed some examples of genetic engineering in non-native host. In 2015, as the last two unknown genes in vindoline biosynthetic pathway branch, tabersonine 3-oxygenase (T3O) and tabersonine 3-reductase (T3R), catalyzing the “hydration” reaction from 16-methoxytabersonine to 16-methoxy-2,3-dihydro-3-hydroxytabersonine were isolated, De Luca *et al.* assembled the seven-step pathway from tabersonine to vindoline in yeast [21]. By feeding with tabersonine or one of the intermediates, 16-methoxytabersonine, the engineered yeast strains were able to accumulate vindoline and other intermediates, and secreted 95% of the MIA to the medium. In another example, as the genes involved in the upstream secologanin synthetic pathway were all characterized in succession, O’Connor’s group introduced the whole pathways in *S. cerevisiae* together with other genetic manipulations to produce the central intermediate, strictosidine, without any precursor feeding [22].

### **Regulation of the TIA pathway**

The TIA pathways involve at least 35 intermediates, 30 biosynthetic enzymes, eight regulatory genes, and several cellular compartments [23-25]. Perhaps not surprisingly, the genetic manipulation of the TIA pathway in *C. roseus* is rarely successful in increasing the final levels of TIA metabolites. Because the TIA pathway is highly regulated, a complete understanding of the regulatory mechanisms of the TIA pathway may be a requirement for

successful metabolic engineering. Despite the complicated regulation, recent research has shed light on this area.

Extensive research has studied the influence of signaling molecules on TIA biosynthesis. Although the elucidation of the entire regulatory mechanism is not complete, those signaling molecules, such as jasmonate (JA), ethylene, nitrous oxide (NO), and salicylic acid, not only demonstrate involvement in TIA biosynthesis, but also mediate TIA production by either synergistic or antagonistic effects [26].

The most direct regulation of the TIA pathway happens at the transcriptional level. Transcription factors, binding to a specific element and regulating the expression of genes, are a major mechanism controlling the biosynthetic genes of the TIA pathway. Transcription factors are usually regulated by signaling molecules or other elements. During the past few years, much effort was made to identify the transcription factors regulating the TIA pathway and to elucidate which genes are regulated by specific transcription factors.

The most well-known transcription factors are the Octadecanoid-responsive *Catharanthus* AP2/ERF domain (ORCAs), including ORCA1, ORCA2, and ORCA3, which are members of the AP2/ERF transcription factor family. ORCA3, for which expression is inducible with JA and fungal elicitation, is the most widely studied. The overexpression of ORCA3 in *C. roseus* hairy roots enhanced the transcript levels of several biosynthetic genes (*AS*, *TDC*, *DXS*, *CPR*, *G10H*, *SLS*, *STR*, *SGD*, and *D4H*) involved in TIA pathway and consequently increased the accumulation of some TIAs [27]. Previously, overexpression of ORCA3 in *C. roseus* suspension cells also enhanced the expression of *AS*, *TDC*, *DXS*, *CPR*, *STR*, and *D4H* genes [28].

Yeast one-hybrid screening with a *Str* promoter indicated that ORCA2 activates the *Str* promoter and its expression is rapidly inducible with JA and fungal elicitation; in contrast,

ORCA1 is expressed constitutively and is not involved in JA- and fungal elicitor-induced *Str* gene expression [29]. Recently, the overexpression of ORCA2 in *C. roseus* hairy roots enhanced the transcript levels of *AS*, *TDC*, *G10H*, *STR*, *D4H*, *T16H*, and *PRX1* genes [30]. Besides the three ORCAs, another transcription factor, the *C. roseus* box P-binding factor (CrBPF-1) can also bind the *Str* promoter but at a different position. Research suggests that CrBPF-1 may enhance the expression of the *Str* gene when the ORCAs are already bound to the promoter [31]. CrMYC1 [32] and CrMYC2 [33] are basic helix-loop-helix transcription factors. Both of these two transcription factors are induced by JA and fungal elicitation. The mRNA levels of *CrMYC1* and *Str* genes were enhanced after induction by JA and fungal elicitation, which suggests that CrMYC1 activates the expression of the *Str* gene [32]. CrMYC2 is postulated to act upstream of ORCA2 and ORCA3, activating their transcription [33].

ORCA2, ORCA3, CrBPF-1, CrMYC1, and CrMYC2 act as transcriptional activators for some TIA biosynthetic genes as well as for several transcription repressors for TIA biosynthesis. Three members of the Cys2/His2-type (transcription factor IIIA-type) zinc finger protein family from *C. roseus*, ZCT1, ZCT2, and ZCT3, are activated by ORCA2 and ORCA3, and repress the activity of the promoters of *TDC* and *STR*. In addition, the ZCT proteins can repress the activating activity of the AP2/ERF domain of the ORCAs [34]. In addition to the ZCT proteins, G-box binding factors 1 and 2 (GBF-1 and GBF-2) also act as repressors of *STR* gene expression [35].

CrWRKY1 and CrWRKY2 are jasmonate responsive WRKY transcription factors that activate several genes involved in TIA biosynthesis [36,37]. Overexpression of *CrWRKY1* in *C. roseus* hairy roots up-regulated *TDC*, as well as the transcriptional repressors *ZCT1*, *ZCT2*, and *ZCT3*, and down-regulated the transcriptional activators *ORCA2*, *ORCA3*, and *CrMYC2* [36]. In

contrast, overexpression of *CrWRKY2* in *C. roseus* hairy roots led to an increase in mRNA transcripts of *TDC*, *NMT*, *DAT*, *MAT*, as well as both specific TIA transcriptional activators (*ORCA2*, *ORCA3* and *CrWRKY1*) and repressors (*ZCT1* and *ZCT3*).

The complex interactions of the regulation of TIA biosynthesis is summarized in Figure 1.5. Although the isolation of transcription factors regulating the TIA biosynthetic pathway is progressing, further characterization of the regulatory pathways is needed.

### **Transcriptional and metabolic profiling of transgenic *C. roseus* hairy root lines**

In order to have a better understanding of the transcriptional regulators in TIA pathway, we generated and characterized several transgenic *C. roseus* hairy root lines with different genotypes. The first one was overexpression of a transcriptional activator *ORCA2*, the second one was overexpression of another transcriptional activator *CrBPF1*, and the third one carried RNA interference (RNAi) constructs of two transcriptional repressors, *CrGBF1* and *CrGBF2*. The time course analysis of transcript levels in those transgenic hairy root lines was characterized to determine the effects of those regulators on TIA biosynthetic and regulatory gene expression levels on time. The results from real-time PCR provided critical information about the TIA regulatory network, including which genes are regulated by which regulators and the timing with which changes in transcript levels occur. Metabolite analysis was also conducted on those hairy root lines. In order to profile more TIAs, heterologous expression of plant genes in yeast and analytical instruments, like LC-MS and NMR, were applied to prepare more in-house TIA standards. The results of the metabolic profiling revealed how those transcriptional regulators affect TIA metabolite levels.

In addition, to study the expression status of vindoline biosynthetic pathway in hairy root, we co-overexpressed the first two genes, tabersonine 16-hydroxylase (*T16H*) and 16-

hydroxytabersonine 16-*O*-methyltransferase (16OMT), catalyzing tabersonine to the first two intermediates in vindoline biosynthetic pathway. Again, transcriptional and metabolic profiling of the TIA biosynthetic pathway was done to characterize the engineered hairy root lines to have an intense study of the changes in the transcript and metabolite levels.

### **Wax esters**

Wax esters include all the esters of fatty acids with long chain fatty alcohols (Figure 1.6). Based on the different melting temperatures, they are divided into either liquid waxes or solid waxes. The liquid waxes are used in high pressure lubricants, pharmaceutical, cosmetics, linoleum and printing inks industries, while the solid waxes are used in candles, polish products and so on. The natural sources of wax esters are extremely broad, mineral, microorganisms, plants and animals could contain various types of wax esters. Currently, most of wax esters are purified primarily from mineral oils, while the biorenewable sources of wax esters increasingly attract interest.

A previous most accessible natural source for wax esters was sperm whale. In the late 18<sup>th</sup> and early 19<sup>th</sup> centuries, sperm whale oil was used as an outstanding lubricant because of its low viscosity and stability, and it does not freeze down to -30°C. The compositions of sperm oil are mostly wax esters and a small proportion of triglyceride. The most abundant wax esters (mainly cetyl palmitate) is in adult males (71~94%). As the development of new lubricant being able to substitute the sperm oil lubricant with same performance and sperm whales being an endangered species because of overhunting, the US government and several other countries banned the whaling and the imports of whale products. From then on, the sperm oil gradually steps down from the stage of lubricant industry.

The current natural source of most wax esters is the seeds of Jojoba (*Simmondsia chinensis*). Jojoba seeds contain 50% in dry weight of wax esters which have similar structure as sperm whale oil and the wax esters human skin produced. Jojoba oil is mainly used in cosmetics, pharmaceuticals, automotives and food products. And the applications are still limited by the tight supply and high price of the jojoba oil. There are many other species contain relatively high amount of wax ester, but the yields are low so that they only have narrow applications. For examples, beeswax isolated from honeycomb contains 35 to 80% of wax esters, and is used in skin care, cosmetics and candle-making. The leaves of the carnauba palm (*Copernicia cerifera*) have a thick coating of wax esters, which is mainly used in polishing. Another type of natural wax esters used in cosmetics is lanolin extracted from the wool of sheep.

Besides animal and plant species producing wax ester, some types of bacteria also produce wax esters naturally. They belong to *Acinetobacter* sp., *Marinobacter* sp., *Psychrobacter* sp., and *Rhodococcus* sp.. The most studied wax ester-producing bacterium is *Acinetobacter baylyi* (*A. baylyi*). *A. baylyi* accumulates plenty of wax esters in the existence of alkanes or alkanols under nitrogen-limited conditions [38]. The pathway from acyl-CoA to wax esters has been well characterized except that an aldehyde reductase is unknown, and the gene encoding bifunctional enzyme, wax ester synthase/acyl coenzyme A: diacylglycerol acyl transferase (WS/DGAT), has been cloned into many hosts to produce biofuels, such as fatty acid methyl ester or fatty acid ethyl ester. In addition, some metabolic engineering works involved in wax ester production were done in *A. baylyi*. By substituting the original acyl-CoA reductase to another fatty acid reductase complex LuxCDE, Santala's group shifted the fatty alkanol chain length of wax esters toward shorter and more saturated ones [39].

## **Y. lipolytica**

*Y. lipolytica* is one of the most extensively studied oleaginous yeasts, which is capable of producing large amount of various metabolites and enzymes for industrial applications. The advantages of using *Y. lipolytica* as an expression and production host is that (1) it has an intense secretory activity, (2) it was classified as generally recognized as safe (GRAS) by FDA, (3) many genetic tools have been developed in this system.

The most important product coming from *Y. lipolytica* is a protein, lipase. Lipases (E.C. 3.1.1.3) can catalyze several reactions including hydrolysis, esterification, interesterification and transesterification. Thus lipases have wide applications, such as detergents, substituting chemical surfactants [40], for the treatment of oil contaminant wastewaters [41], for the organic synthesis of optically active pharmaceuticals or intermediates [42], for flavor development in food industry [43], and in biodiesel production [44]. And the metabolic engineering and bioreactor fermentation works of producing lipases from *Y. lipolytica* were well reviewed by Fickers *et al.* [45]. Besides lipases, *Y. lipolytica* can also produce and secrete other endogenous enzymes, such as extracellular protease, alkaline protease, RNase, phosphatase, and esterase. Moreover, it is also an ideal expression host for heterologous protein production like laccase and epoxidase hydrolase.

For the production of some organic acids in the glycolysis and TCA cycle, such as citric acid, isocitric acid,  $\alpha$ -ketoglutaric acid, and pyruvic acid, *Y. lipolytica* presents the ability to produce and secrete large amount of those primary metabolites, which is comparable to most of the conventional commercial production fungi. The convenient genetic engineering in *Y. lipolytica* brings this fungus close to the commercial host for those organic acid productions. For example, *Y. lipolytica* has been engineered to produce comparable amount of citric acid as its

conventional producer, *Aspergillus niger*, and its fermentation process has been and is being used to produce citric acid by Archer Daniels Midland [46-48].

*Y. lipolytica*, as one of the oleaginous microorganisms, is able to produce high amount of microbial oils, also known as single cell oils. These oils have similar compositions as plant oils, and have the potential to be used in the same industrial fields as plant oils. The naturally isolated *Y. lipolytica* strains can only accumulate lipids less than 20% of cell dry weight. But now the lipid content in *Y. lipolytica* has been improved to 87% of cell dry weight, which makes it the rising star in oleaginous microorganisms [49].

In addition, *Y. lipolytica* have many other industrial applications. It is used to produce  $\gamma$ -decalactone, a peach-like aroma compound used in food and beverages. It can also been employed in the bioremediation of oil-contaminated environments. The applications of *Y. lipolytica* have been discussed in several recent published reviews [50,51].

### **Production of wax esters in *Y. lipolytica***

Since *Y. lipolytica* can accumulate large amount of lipids, together with the capability for high-level production of heterologous proteins and GRAS status, it becomes the ideal host of producing various fatty acid-based high-value bioproducts. “LipoYeasts” project in Europe targeted on studying and engineering *Y. lipolytica* to produce wax esters, hydroxylated fatty acids, polyhydroxyalkanoates (PHAs), carotenoids and carotenoid esters [52]. Though lots of efforts have been achieved in this project, so far, there is no publication on the wax esters production in *Y. lipolytica*. In one of my research projects, I introduced the pathway from acyl-CoA to wax ester into *Y. lipolytica*, knocked out the substrate competitive pathway for wax esters, compared the effect of different FARs on wax ester production, integrated heterologous

genes into *Y. lipolytica* genome, and studied the effect of medium composition on wax ester production.

### **Dissertation organization**

This dissertation is organized in the following manner. Chapter 1 is the introduction of my two studied subjects, terpenoid indole alkaloids (TIA) and wax esters. It includes the significances of those two types of compounds and all the reviews on TIA studies and wax ester productions.

Chapter 2 describes a collaborated study on overexpressing the first two genes in the vindoline pathway in *C. roseus* hairy roots. In this chapter, we identified new metabolites in the transgenic hairy root lines, quantified most of the TIAs, and analyzed the transcriptional changes of the TIA pathway genes and regulators. The results showed that the transgenic hairy root lines could produce the first two metabolites in vindoline pathway, as well as some new byproducts. It also illustrated how the overexpression of the first two genes triggered the complex transcriptional responses.

Chapter 3 is a summary of my contributions in two published papers. In this chapter, we studied the effects of overexpressing two transcriptional activators, ORCA2 and CrBPF1, on the metabolic profiles in TIA pathway. In total, time course analysis of four metabolites in the upstream pathway and twelve TIAs in the downstream pathway were done. And we found that by overexpressing those two transcriptional activators, especially ORCA2, some of the metabolites in TIA pathway did increase their levels to some extent at specific time points, while most of the metabolites kept the same level compared with controls, indicating the regulation mechanism is extremely complex and the TIA pathway is tightly regulated.

Chapter 4 details a manuscript that producing wax esters in yeast *Yarrowia lipolytica*. In this study, we knocked out some substrate competitive pathway, aiming to increase the acyl-CoA pool for wax ester production, and studied the effects of three different fatty acyl-CoA reductases (FAR), nitrogen limited fermentation on the titers and contents of wax esters. In addition, to avoid the plasmid instability issue, we randomly integrated FAR gene singly and combining WS gene into the chromosome. To date, the best titer of wax ester we got was around 550 mg/L, which was the highest wax ester production in microbes reported so far.

This dissertation was under the guidance, supervision and support of my major professor, Zengyi Shao.

## Table

Table 1.1 Examples of recent metabolic engineering efforts in TIA pathway

hosts	engineering strategies	role in TIA pathway	expression system	main observations
native host	overexpression of WRKY1 [36]	transcription activator	hairy root	3-fold increase in serpentine; 10-fold increase in ajmalicine; 2-fold decrease in catharanthine
	overexpression of ORCA3 [53]	transcription activator	hairy root	3-fold increase in hörhammericine; 3-fold increase in lochnericine
	overexpression of DAT [13]	pathway enzyme	hairy root	4-fold increase in hörhammericine
	overexpression of As $\alpha$ and As $\beta$ , and feeding of 10-deoxy-D-xylulose, loganin and secologanin [54]	pathway enzyme	hairy root	2.3-fold increase in hörhammericine; 1.5-fold increase in cathenamine; 1.3-fold increase in catharanthine; 1.8-fold increase in ajmalicine; 2.1-fold increase in lochnericine
	overexpression of ORCA3 and feeding of loganin [55]	transcription activator	suspension cell	3.2-fold increase in TIA; detectable strictosidine and ajmalicine
	overexpression of DAT [56]	pathway enzyme	plant	2-fold increase in vindoline
	overexpression of ORCA2 [57]	transcription activator	hairy root	detectable 19-hydroxytabersonine after induction
	co-overexpression of ORCA3 and SGD [58]	transcription activator and pathway enzyme	hairy root	0.5-fold increase in TIA including serpentine, ajmalicine, catharanthine, tabersonine, lochnericine and hörhammericine
non-native hosts	co-expression of TDC and STR, and feeding of secologanin [59]	pathway enzyme	tobacco suspension cell	strictosidine can be detected
	co-expression of TDC and STR [60]	pathway enzyme	<i>Cinchona officinalis</i> hairy root	1.2 mg g <sup>-1</sup> DW of tryptamine; 2.0 mg g <sup>-1</sup> DW of strictosidine; not a stable trait
	co-expression of STR and SGD, and feeding of tryptamine and secologanin [61]	pathway enzyme	<i>Saccharomyces cerevisiae</i>	Strictosidine (major product, 2g L <sup>-1</sup> ) and cathenamine can be detected
	expression of most genes in strictosidine biosynthetic pathway and other genetic modifications in expression host [22]	pathway enzyme	<i>Saccharomyces cerevisiae</i>	Strictosidine can be detected without any feeding.
	expression of the pathway from tabersonine to vindoline and feeding of tabersonine [21]	pathway enzyme	<i>Saccharomyces cerevisiae</i>	Vindoline can be detected.

## Figures

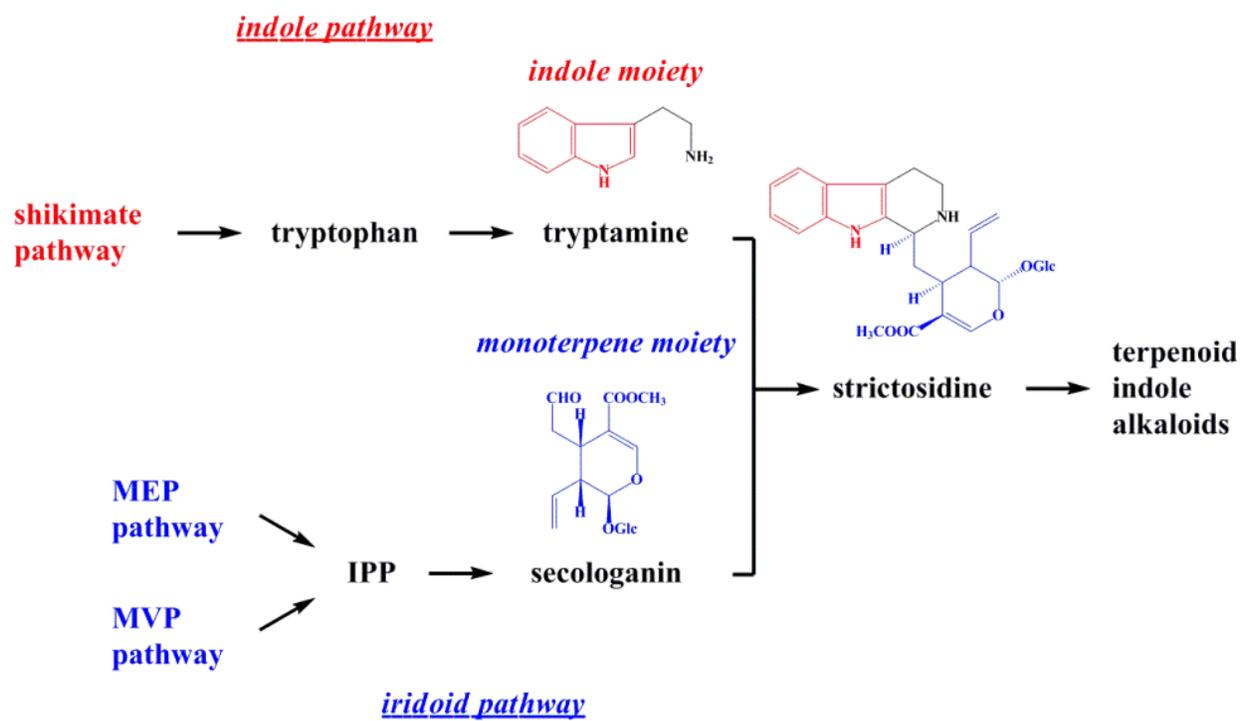
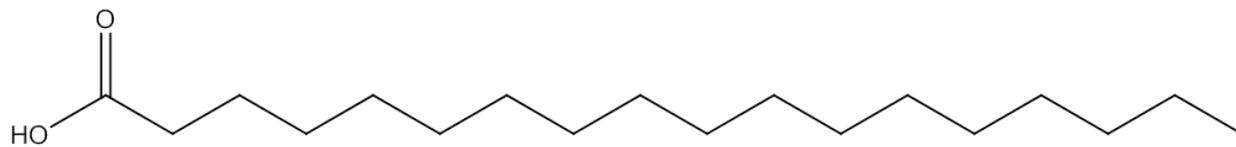
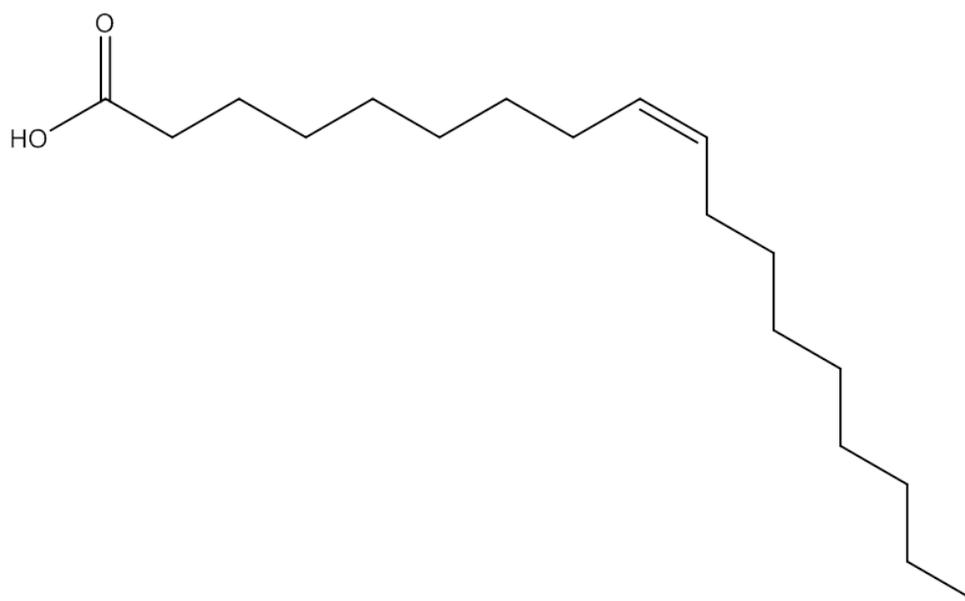


Figure 1.1 An overview of the pathways leading to TIA biosynthesis



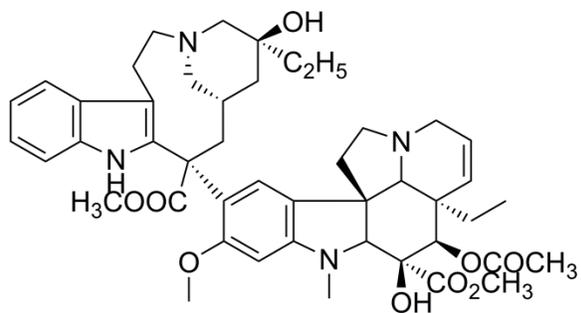
Stearic acid

Saturated fatty acid

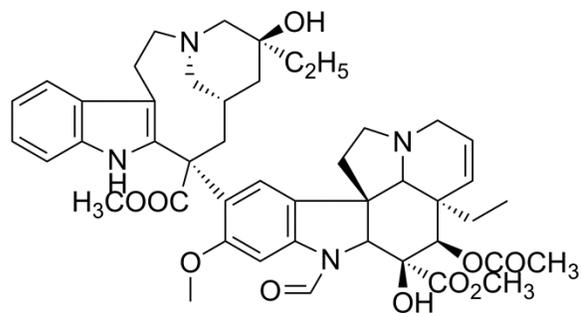
*cis*-Oleic acid

Unsaturated fatty acid

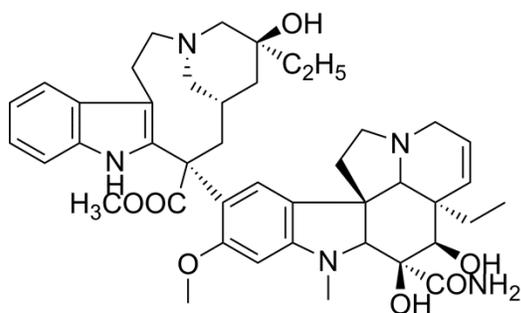
Figure 1.2 Examples of fatty acids.



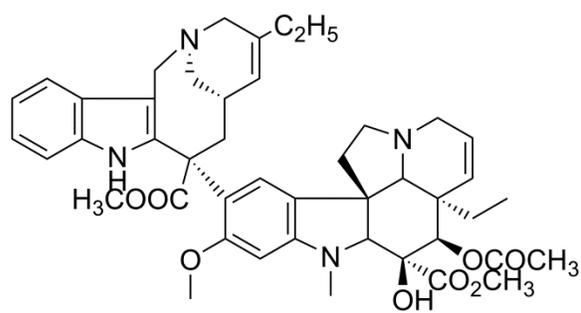
vinblastine



vincristine



vindesine



vinorelbine

Figure 1.3 Chemical structures of vinblastine, vincristine, vindesine, and vinorelbine.



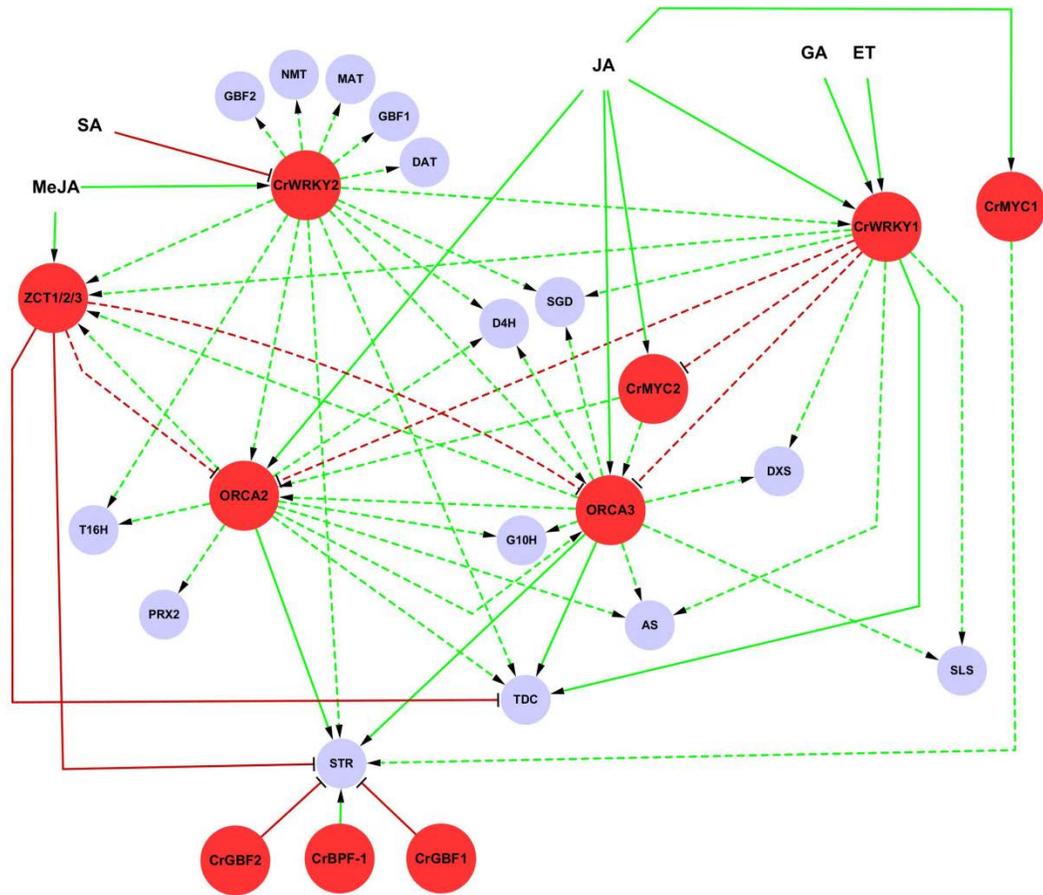
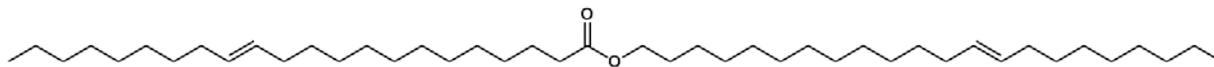
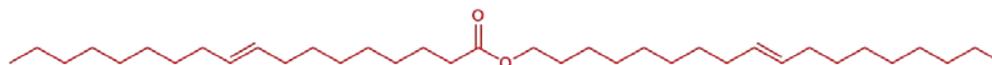


Figure 1.5 Regulation map of the TIA biosynthesis pathway. MeJA: methyl jasmonic acid; SA: salicylic acid; JA: jasmonic acid; ET: ethylene; GA: gibberel-lic acid; the nodes with empty background are signaling molecules; the nodes with red background are transcriptional factors; the nodes with grey background are genes in TIA biosynthesis pathway; dashed lines represent interactions that may be direct or indirect; solid lines indicate potentially direct interactions. Green lines with arrows represent transactivation and red lines with bars represent transrepression. Note: CrWRKY2 overexpressed line has great positive effect on ZCT1, ZCT3 but only slight positive effect on ZCT2, however, in order to make this figure clear, ZCT1, ZCT2, and ZCT3 are lumped.

Jojoba wax ester



Human skin wax ester



Sperm whale oil wax ester

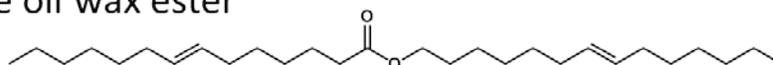


Figure 1.6 Examples of wax esters.

## References

- 1 Zhou, X.J., Rahmani, R. (1992) Preclinical and clinical pharmacology of vinca alkaloids. *Drugs*, **44 Suppl 4**, 1-16; discussion 66-9.
- 2 Noble, R.L. (1990) The discovery of the vinca alkaloids--chemotherapeutic agents against cancer. *Biochem Cell Biol*, **68**, 1344-51.
- 3 Kumar, A., Patil, D., Rajamohanam, P.R., Ahmad, A. (2013) Isolation, purification and characterization of vinblastine and vincristine from endophytic fungus *Fusarium oxysporum* isolated from *Catharanthus roseus*. *Plos One*, **8**.
- 4 DiCosmo, F., Misawa, M. (1995) Plant cell and tissue culture: alternatives for metabolite production. *Biotechnol Adv*, **13**, 425-53.
- 5 Schlatmann, J.E., Hoopen, H.J.G.t., Heijnen, J.J. (1996) Large-Scale Production of Secondary Metabolites by Plant Cell Cultures. in *Plant Cell Culture Secondary Metabolism Toward Industrial Application*, eds DiCosmo F, Misawa M), CRC Press.
- 6 Rao, S.R., Ravishankar, G.A. (2002) Plant cell cultures: chemical factories of secondary metabolites. *Biotechnol Adv*, **20**, 101-53.
- 7 Wilson, S.A., Roberts, S.C. (2012) Recent advances towards development and commercialization of plant cell culture processes for the synthesis of biomolecules. *Plant Biotechnol J*, **10**, 249-68.
- 8 Sharma, P., Padh, H., Shrivastava, N. (2013) Hairy root cultures: a suitable biological system for studying secondary metabolic pathways in plants. *Eng Life Sci*, **13**, 62-75.
- 9 Shukla, A.K., Shasany, A.K., Verma, R.K., Gupta, M.M., et al. (2010) Influence of cellular differentiation and elicitation on intermediate and late steps of terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *Protoplasma*, **242**, 35-47.
- 10 Vazquez-Flota, F., De Luca, V., Carrillo-Pech, M., Canto-Flick, A., et al. (2002) Vindoline biosynthesis is transcriptionally blocked in *Catharanthus roseus* cell suspension cultures. *Mol Biotechnol*, **22**, 1-8.
- 11 Deluca, V., Balsevich, J., Tyler, R.T., Kurz, W.G. (1987) Characterization of a novel *N*-methyltransferase (NMT) from *Catharanthus roseus* plants : Detection of NMT and other enzymes of the indole alkaloid biosynthetic pathway in different cell suspension culture systems. *Plant Cell Rep*, **6**, 458-61.
- 12 Verma, M., Ghangal, R., Sharma, R., Sinha, A.K., et al. (2014) Transcriptome analysis of *Catharanthus roseus* for gene discovery and expression profiling. *Plos One*, **9**, e103583.
- 13 Magnotta, M., Murata, J., Chen, J., De Luca, V. (2007) Expression of deacetylvindoline-4-O-acetyltransferase in *Catharanthus roseus* hairy roots. *Phytochemistry*, **68**, 1922-31.

- 14 Howe, G.A. (2004) Jasmonates as signals in the wound response. *J Plant Growth Regul*, **23**, 223-37.
- 15 Verpoorte, R., van der Heijden, R., van Gulik, W.M., ten Hoopen, H.J.G. (1991) Plant biotechnology for the production of alkaloids: present status and prospects. in *The Alkaloids*, (ed Brossi A), Academic Press, San Diego.
- 16 Mujib, A., Ilah, A., Aslam, J., Fatima, S., et al. (2012) *Catharanthus roseus* alkaloids: application of biotechnology for improving yield. *Plant Growth Regul*, **68**, 111-27.
- 17 van der Heijden, R., Jacobs, D.I., Snoeijs, W., Hallared, D., et al. (2004) The *Catharanthus* alkaloids: pharmacognosy and biotechnology. *Curr Med Chem*, **11**, 607-28.
- 18 van der Fits, L., Memelink, J. (1997) Comparison of the activities of CaMV 35S and FMV 34S promoter derivatives in *Catharanthus roseus* cells transiently and stably transformed by particle bombardment. *Plant Mol Biol*, **33**, 943-6.
- 19 Goddijn, O.J., Pennings, E.J., van der Helm, P., Schilperoort, R.A., et al. (1995) Overexpression of a tryptophan decarboxylase cDNA in *Catharanthus roseus* crown gall calluses results in increased tryptamine levels but not in increased terpenoid indole alkaloid production. *Transgenic Res*, **4**, 315-23.
- 20 Martin, V.J.J., Pitera, D.J., Withers, S.T., Newman, J.D., et al. (2003) Engineering a mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat Biotechnol*, **21**, 796-802.
- 21 Qu, Y., Easson, M.L., Froese, J., Simionescu, R., et al. (2015) Completion of the seven-step pathway from tabersonine to the anticancer drug precursor vindoline and its assembly in yeast. *Proc Natl Acad Sci U S A*, **112**, 6224-9.
- 22 Brown, S., Clastre, M., Courdavault, V., O'Connor, S.E. (2015) De novo production of the plant-derived alkaloid strictosidine in yeast. *Proc Natl Acad Sci U S A*, **112**, 3205-10.
- 23 van der Heijden, R., Jacobs, D.I., Snoeijs, W., Hallard, D., et al. (2004) The *Catharanthus* alkaloids: pharmacognosy and biotechnology. *Curr Med Chem*, **11**, 607-28.
- 24 Facchini, P.J., De Luca, V. (2008) Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. *Plant J*, **54**, 763-84.
- 25 Burlat, V., Oudin, A., Courtois, M., Rideau, M., et al. (2004) Co-expression of three MEP pathway genes and geraniol 10-hydroxylase in internal phloem parenchyma of *Catharanthus roseus* implicates multicellular translocation of intermediates during the biosynthesis of monoterpene indole alkaloids and isoprenoid-derived primary metabolites. *Plant J*, **38**, 131-41.
- 26 El-Sayed, M., Verpoorte, R. (2004) Growth, metabolic profiling and enzymes activities of *Catharanthus roseus* seedlings treated with plant growth regulators. *Plant Growth Regul*, **44**, 53-8.

- 27 Peebles, C.A., Hughes, E.H., Shanks, J.V., San, K.Y. (2009) Transcriptional response of the terpenoid indole alkaloid pathway to the overexpression of ORCA3 along with jasmonic acid elicitation of *Catharanthus roseus* hairy roots over time. *Metab Eng*, **11**, 76-86.
- 28 van der Fits, L., Memelink, J., van der Fits, L., Memelink, J. (2000) ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Sci Signal*, **289**, 295.
- 29 Menke, F.L., Champion, A., Kijne, J.W., Memelink, J. (1999) A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene *Str* interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2. *EMBO J*, **18**, 4455-63.
- 30 Li, C.Y., Leopold, A.L., Sander, G.W., Shanks, J.V., et al. (2012) The ORCA2 transcription factor plays a key role in regulation of the terpenoid indole alkaloid pathway. *Metab Eng*.
- 31 van der Fits, L., Zhang, H., Menke, F.L., Deneka, M., et al. (2000) A *Catharanthus roseus* BPF-1 homologue interacts with an elicitor-responsive region of the secondary metabolite biosynthetic gene *Str* and is induced by elicitor via a JA-independent signal transduction pathway. *Plant Mol Biol*, **44**, 675-85.
- 32 Chatel, G., Montiel, G., Pre, M., Memelink, J., et al. (2003) *CrMYC1*, a *Catharanthus roseus* elicitor- and jasmonate-responsive bHLH transcription factor that binds the G-box element of the strictosidine synthase gene promoter. *J Exp Bot*, **54**, 2587-8.
- 33 Zhang, H.T., Hedhili, S., Montiel, G., Zhang, Y.X., et al. (2011) The basic helix-loop-helix transcription factor CrMYC2 controls the jasmonate-responsive expression of the ORCA genes that regulate alkaloid biosynthesis in *Catharanthus roseus*. *Plant J*, **67**, 61-71.
- 34 Pauw, B., Hilliou, F.A., Martin, V.S., Chatel, G., et al. (2004) Zinc finger proteins act as transcriptional repressors of alkaloid biosynthesis genes in *Catharanthus roseus*. *J Biol Chem*, **279**, 52940-8.
- 35 Sibénil, Y., Benhamron, S., Memelink, J., Giglioli-Guivarc'h, N., et al. (2001) *Catharanthus roseus* G-box binding factors 1 and 2 act as repressors of strictosidine synthase gene expression in cell cultures. *Plant Mol Biol*, **45**, 477-88.
- 36 Suttipanta, N., Pattanaik, S., Kulshrestha, M., Patra, B., et al. (2011) The transcription factor CrWRKY1 positively regulates the terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *Plant Physiol*, **157**, 2081-93.
- 37 Suttipanta, N. (2011) Characterization of *G10H* promoter and isolation of WRKY transcription factors involved in *Catharanthus* terpenoid indole alkaloid biosynthesis pathway. University of Kentucky.

- 38 Ishige, T., Tani, A., Sakai, Y.R., Kato, N. (2003) Wax ester production by bacteria. *Curr Opin Microbiol*, **6**, 244-50.
- 39 Santala, S., Efimova, E., Koskinen, P., Karp, M.T., et al. (2014) Rewiring the wax ester production pathway of *Acinetobacter baylyi* ADP1. *Acs Synth Biol*, **3**, 145-51.
- 40 Hemachander, C., Puvanakrishnan, R. (2000) Lipase from *Ralstonia pickettii* as an additive in laundry detergent formulations. *Process Biochem*, **35**, 809-14.
- 41 Cammarota, M.C., Freire, D.M.G. (2006) A review on hydrolytic enzymes in the treatment of wastewater with high oil and grease content. *Bioresource Technol*, **97**, 2195-210.
- 42 Akai, S., Hanada, R., Fujiwara, N., Kita, Y., et al. (2010) One-Pot Synthesis of Optically Active Allyl Esters via Lipase-Vanadium Combo Catalysis. *Organic Letters*, **12**, 4900-3.
- 43 Akin, N., Aydemir, S., Kocak, C., Yildiz, M.A. (2003) Changes of free fatty acid contents and sensory properties of white pickled cheese during ripening. *Food Chem*, **80**, 77-83.
- 44 Fan, X.H., Niehus, X., Sandoval, G. (2012) Lipases as Biocatalyst for Biodiesel Production. *Lipases and Phospholipases: Methods and Protocols*, **861**, 471-83.
- 45 Fickers, P., Marty, A., Nicaud, J.M. (2011) The lipases from *Yarrowia lipolytica*: Genetics, production, regulation, biochemical characterization and biotechnological applications. *Biotechnology Advances*, **29**, 632-44.
- 46 Forster, A., Aurich, A., Mauersberger, S., Barth, G. (2007) Citric acid production from sucrose using a recombinant strain of the yeast *Yarrowia lipolytica*. *Appl Microbiol Biot*, **75**, 1409-17.
- 47 Ikram-ul, H., Ali, S., Qadeer, M.A., Iqbal, J. (2004) Citric acid production by selected mutants of *Aspergillus niger* from cane molasses. *Bioresource Technol*, **93**, 125-30.
- 48 Fickers, P., Benetti, P.H., Wache, Y., Marty, A., et al. (2005) Hydrophobic substrate utilisation by the yeast *Yarrowia lipolytica*, and its potential applications. *Fems Yeast Res*, **5**, 527-43.
- 49 Liu, L., Pan, A., Spofford, C., Zhou, N., et al. (2015) An evolutionary metabolic engineering approach for enhancing lipogenesis in *Yarrowia lipolytica*. *Metabolic Engineering*, **29**, 36-45.
- 50 Liu, H.H., Ji, X.J., Huang, H. (2015) Biotechnological applications of *Yarrowia lipolytica*: Past, present and future. *Biotechnology Advances*, **33**, 1522-46.
- 51 Goncalves, F.A.G., Colen, G., Takahashi, J.A. (2014) *Yarrowia lipolytica* and Its Multiple Applications in the Biotechnological Industry. *Sci World J*.
- 52 Sabirova, J.S., Haddouche, R., Van Bogaert, I.N., Mulaa, F., et al. (2011) The 'LipoYeasts' project: using the oleaginous yeast *Yarrowia lipolytica* in combination with

- specific bacterial genes for the bioconversion of lipids, fats and oils into high-value products. *Microb Biotechnol*, **4**, 47-54.
- 53 Peebles, C.A.M., Hughes, E.H., Shanks, J.V., San, K.Y. (2009) Transcriptional response of the terpenoid indole alkaloid pathway to the overexpression of ORCA3 along with jasmonic acid elicitation of *Catharanthus roseus* hairy roots over time. *Metab Eng*, **11**, 76-86.
- 54 Peebles, C.A.M., Hong, S.B., Gibson, S.I., Shanks, J.V., et al. (2006) Effects of terpenoid precursor feeding on *Catharanthus roseus* hairy roots over-expressing the alpha or the alpha and beta subunits of anthranilate synthase. *Biotechnol Bioeng*, **93**, 534-40.
- 55 van der Fits, L., Memelink, J. (2000) ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Science*, **289**, 295-7.
- 56 Wang, Q., Xing, S.H., Pan, Q.F., Yuan, F., et al. (2012) Development of efficient *Catharanthus roseus* regeneration and transformation system using *Agrobacterium tumefaciens* and hypocotyls as explants. *Bmc Biotechnol*, **12**.
- 57 Li, C.Y., Leopold, A.L., Sander, G.W., Shanks, J.V., et al. (2013) The ORCA2 transcription factor plays a key role in regulation of the terpenoid indole alkaloid pathway. *Bmc Plant Biol*, **13**.
- 58 Sun, J., Peebles, C.A. (2015) Engineering overexpression of ORCA3 and strictosidine glucosidase in *Catharanthus roseus* hairy roots increases alkaloid production. *Protoplasma*.
- 59 Hallard, D., vanderHeijden, R., Verpoorte, R., Cardoso, M.I.L., et al. (1997) Suspension cultured transgenic cells of *Nicotiana tabacum* expressing tryptophan decarboxylase and strictosidine synthase cDNAs from *Catharanthus roseus* produce strictosidine upon secologanin feeding. *Plant Cell Rep*, **17**, 50-4.
- 60 Geerlings, A., Hallard, D., Caballero, A.M., Cardoso, I.L., et al. (1999) Alkaloid production by a *Cinchona officinalis* 'Ledgeriana' hairy root culture containing constitutive expression constructs of tryptophan decarboxylase and strictosidine synthase cDNAs from *Catharanthus roseus*. *Plant Cell Rep*, **19**, 191-6.
- 61 Geerlings, A., Redondo, F.J., Contin, A., Memelink, J., et al. (2001) Biotransformation of tryptamine and secologanin into plant terpenoid indole alkaloids by transgenic yeast. *Appl Microbiol Biot*, **56**, 420-4.

## CHAPTER II. EXPRESSION OF TABERSONINE 16-HYDROXYLASE AND 16-HYDROXYTABERSONINE-*O*-METHYL TRANSFERASE IN *CATHARANTHUS ROSEUS* HAIRY ROOTS

This chapter is a manuscript submitted to Biotechnology and Bioengineering. I contributed in the identification of new TIAs and quantification of all the TIAs in this work.

Jiayi Sun<sup>1,#</sup>, Le Zhao<sup>2,#</sup>, Zengyi Shao<sup>2</sup>, Jacqueline Shanks<sup>2</sup>, and Christie A. M. Peebles<sup>1</sup>

<sup>1</sup> Dept. of Chemical and Biological Engineering, Colorado State University

<sup>2</sup> Dept. of Chemical and Biological Engineering, Iowa State University

<sup>#</sup>These authors contributed equally to the paper

### Abstract

The monoterpene indole alkaloids vindoline and catharanthine, which are exclusively synthesized in the medicinal plant *Catharanthus roseus*, are the two important precursors for the production of pharmaceutically important anti-cancer medicines vinblastine and vincristine. Hairy root culture is an ideal platform for alkaloids production due to its industry scalability, genetic and chemical stability, and availability of genetic engineering tools. However, *C. roseus* hairy roots do not produce vindoline due to the lack of expression of the seven-step pathway from tabersonine to vindoline [1]. The present study describes the genetic engineering of the first two genes tabersonine 16-hydroxylase (*T16H*) and 16-*O*-methyl transferase (*16OMT*) in the missing vindoline pathway under the control of a glucocorticoid-inducible promoter to direct tabersonine toward vindoline biosynthesis in *C. roseus* hairy roots.

In two transgenic hairy roots, the induced overexpression of *T16H* and *16OMT* resulted in the accumulation of vindoline pathway metabolites 16-hydroxytabersonine and 16-methoxytabersonine. The levels of root-specific alkaloids, including lochnericine, 19-hydroxytabersonine and hörhammericine, significantly decreased in the induced hairy roots in comparison to the uninduced control lines. This suggests tabersonine was successfully channeled

to the vindoline pathway away from the roots competing pathway based on the overexpression. Interestingly, another two new metabolites were detected in the induced hairy roots and proposed to be the epoxidized 16-hydroxytabersonine and lochnerinine. It indicates the introduction of vindoline pathway genes in hairy roots can cause unexpected terpenoid indole alkaloids (TIA) profile alterations. Furthermore, the complex transcriptional changes in TIA genes and regulators detected by RT-qPCR point out the tight regulation of the TIA pathway in response to *T16H* and *16OMT* engineering in *C. roseus* hairy roots.

**Key words:** Terpenoid indole alkaloid, vindoline pathway, metabolic engineering, Madagascar periwinkle, plant secondary metabolites

## Introduction

Vinca alkaloids vinblastine and vincristine are powerful antineoplastic drugs that are widely used in many chemotherapy regimens for a variety of cancers [2]. They act by binding to intracellular tubulin which results in the inhibition of cell division by blocking mitosis [3]. The chirality and complexity of these chemicals make the synthesis of these drugs using chemical methods economically infeasible [4]. *Catharanthus roseus* (Madagascar periwinkle) continues to be the exclusive source for the industrial production of vinblastine and vincristine. The low yields of these dimeric monoterpene indole alkaloids (TIAs) have motivated extensive effort to genetically engineer *C. roseus* for overproduction. Currently, the lack of tools for whole plant engineering, the genetic instability of the cell cultures, and the high accumulation of alkaloids in highly differentiated hairy roots have made hairy roots a promising system to engineer [5, 6].

The biosynthesis of TIAs in *C. roseus* requires two precursors, secologanin from terpenoid pathway and tryptamine from indole pathway (Figure 2.1). The first alkaloid strictosidine synthesized by strictosidine synthase (STR) is converted to strictosidine aglycone

via strictosidine glucosidase (SGD) [7]. The highly active aglycone is catalyzed to corynanthe-type alkaloids (serpentine, ajmalicine), iboga-type alkaloids (catharanthine), and aspidosperma-type alkaloids (tabersonine, vindoline, hörhammericine) via several branched pathways [8]. However, the genes involved in these branched pathways are still largely unknown [9]. The coupling of vindoline and catharanthine produces the two pharmaceutically important bisindole alkaloids vinblastine and vincristine (Figure 2.1) [10]. In addition, the complex compartmentalization of the TIA pathway [11] spatially separates catharanthine and vindoline in different locations, which is one reason for the low yield of bisindole alkaloids within *C. roseus*. Catharanthine is synthesized in leaf epidermis and stored in upper cuticle, and vindoline is produced in laticifers and idioblast [12]. Moreover, these downstream alkaloid pathways are tissue specific. In *C. roseus* cell cultures and hairy roots cultures, the absence of vindoline accumulation is the major challenge to produce the desired bisindole alkaloids in these biological systems, but catharanthine and the vindoline precursor tabersonine do accumulate [13]. Tabersonine in un-differentiated cells and hairy roots is directed to 19-O-acetyhorhammericine instead of vindoline [14]. Recently the full seven steps catalyzing the conversions from tabersonine to vindoline was elucidated [15]. Engineering of this missing vindoline pathway in hairy roots or cell culture is now possible.

To demonstrate the potential for the vindoline pathway to be expressed in hairy roots, we introduced the first two enzymes in the vindoline pathway, tabersonine 16-hydroxylase (T16H) and 16-O-methyl transferase (16OMT), into *C. roseus* hairy roots under the control of an inducible promoter system. The metabolic profile changes were measured after overexpressing *T16H* and *16OMT*. Although no vindoline accumulation was detected due to the lack of expressing the later steps of the vindoline pathway in hairy roots, the catalytic products of T16H

and 16OMT were identified as expected in the engineered hairy roots. This is the first report of accumulating vindoline pathway intermediates in hairy roots. Furthermore, previous studies have pointed out the tight regulation of the TIA pathway based on the genetic engineering works in *C. roseus* hairy roots [16, 17]. In order to better understand how the transcription of alkaloid pathway genes and regulators respond to the metabolic alterations at this branch point, the relative mRNA levels of TIA genes including tabersonine 19-hydroxylase (T19H), minovincinine 19-hydroxy-O-acetyltransferase (MAT), strictosidine synthase (STR), strictosidine beta-glucosidase (SGD), and TIA regulators including AP2-domain DNA-binding proteins (ORCAs) and zinc finger proteins (ZCTs) were examined by RT-qPCR analysis in this study.

## Materials and Methods

### Plasmid construction and clone generation

Plasmid *pTA7002/T16H* and *pTA7002/16OMT* were obtained from Dr. Ka-Yiu San, and *T16H* and *16OMT* sequences were verified by sequencing. The sequencing results of these two genes matched the published sequence for *T16H* (GenBank: FJ647194.1) and *16OMT* (GenBank: EF444544.1). *16OMT* was cut from *pTA7002/16OMT* and moved to the intermediate plasmid *pUCGALA* [18] at the *XhoI/SpeI* site to construct *pUCGALA/16OMT*. *16OMT* along with the promoter sequence *GAL4-UAS* was cut from *pUCGALA/16OMT* with restriction enzyme *SbfI*, and was constructed next to the right border in the T-DNA region of *pTA7002/T16H* (Figure 2.2). The *cis* orientation and sequence of *16OMT* in *pTA7002/T16H/16OMT* was further verified by sequencing. Both *T16H* and *16OMT* genes are under the control of a glucocorticoid-inducible promoter [19].

Plasmids *pTA7002/T16H/16OMT* were electro-transferred into *Agrobacterium rhizogenes* ATCC15834. The presence of the plasmids was confirmed by sequencing. The generation of transgenic *C. roseus* hairy roots was previously described [20]. Briefly, *A. rhizogenes* containing the plasmid was cultured in 6ml YEM media at 28 °C and 225 rpm for 36 h. Forceps dipped into the agrobacteria were used to pinch the stem of the *C. roseus* seedlings. After 6 weeks, hairy roots protruding from the wounding sites were harvested and grown on selection plates (hairy roots media supplemented with 350 mg/L cefotaximine and 30 mg/L hygromycin). Hairy roots growing on selection media were transferred to new hairy roots plates and then adapted to liquid culture.

### **Hairy roots culture and induction**

Hairy roots were sub-cultured every three weeks in 50 ml hairy roots media consisting of 30 g/L sucrose, half strength Gamborg salt (Sigma-Aldrich), full strength Gamborg vitamin (Sigma-Aldrich) as previously described [21]. Triplicates of the transgenic hairy roots at late exponential growth stage (18 days after sub-culture) were treated with 3 µM inducer dexamethasone to induce overexpression of the genes or with an equal volume of ethanol as a negative control. After 72 h of induction, 300 mg fresh hairy roots were grinded with liquid nitrogen in a mortar and were frozen at -80 °C for RNA extraction. The remaining hairy roots were stored at -80 °C for further alkaloid extraction.

### **Alkaloids extraction**

The cultured hairy root was lyophilized and ground with mortar and pestle. About 100 mg of the powdered hairy root was weighed. Each of the samples was spiked with 150 µg vincamine as an internal standard. Metabolites were extracted in 4 ml of methanol with vortexing

for 2 min and sonicating for 10 min. The extracts were centrifuged at 3000 rpm for 3 min. The supernatant was removed and the biomass was extracted twice with 2 ml of methanol in each time. The combined supernatants were dried on a nitrogen evaporator. To further isolate TIAs from the methanol extracted mixture, the dried extract was partitioned between 2 ml of ethyl acetate and 2 ml of 1% HCl solution. TIAs stayed in the water phase. The ethyl acetate upper phase was discarded, and the pH of water phase was adjusted to 10 with ammonium hydroxide. When the water phase became alkaline, the solubility of TIAs in water decreased. So 2 ml of ethyl acetate was used to extract TIAs from the water phase. The ethyl acetate extract was dried on a nitrogen evaporator and re-dissolved in 500  $\mu$ L of methanol.

### **Metabolite standard from yeast fermentation**

A *T16H* gene encoding tabersonine 16-hydroxylase and a *16OMT* gene encoding 16-hydroxytabersonine-O-methyltransferase were cloned into a yeast plasmid *pESC-Leu*. The promoter and terminator for *T16H* ORF were *PGAL1* and *TCYCI*. And the promoter and terminator for *16OMT* ORF were *PGAL10* and *TADH1*. The *pESC-Leu-T16H-16OMT* construct was transformed into *Saccharomyces cerevisiae* *WAT11*. *S. cerevisiae* harboring the desired plasmid were firstly cultured in Synthetic Defined Medium without leucine (SD-Leu) to grow biomass. After 48 hours, the collected *S. cerevisiae* *WAT11* cells were inoculated into 1 L SD-Leu with the glucose replaced by raffinose and galactose to induce the *GAL1* and *GAL10* promoters, keeping an initial OD as 0.2. And the medium was supplemented with 150  $\mu$ M tabersonine. The medium was then extracted after another 48 hours with 1 L of ethyl acetate three times and concentrated with a rotary evaporator. The TIAs in the extract were isolated using same water and ethyl acetate partition method described in previous section.

### **Metabolites identification using LC/MS or LC/MS/MS**

The extract from *S. cerevisiae* WAT11 and hairy roots was analyzed with the Agilent Technologies 1100 series liquid chromatography (LC) system with a binary pump, a temperature-controlled autosampler, a photo diode array detector (PDA) coupled to an Agilent Technologies Mass Selective Trap SL detector equipped with a nanoflow electrospray. The column used to separate alkaloids was a Phenomenex Kinetex 5  $\mu\text{m}$  C18 (150  $\times$  2.1 mm). The flow rate was kept at 0.2 ml/min. The mobile phase was a 30:70 mixture of acetonitrile: 10mM ammonium acetate (pH=5) during the first 6 min. The mobile phase was linearly ramped to 64:36 from the 6th min to the 18th min and maintained at that ratio till the 25th min. The ratio was then further increased to 85:15 within 5min and maintained for 20 min. In the next 5 min, the ratio was returned to 30:70 and the column was allowed to re-equilibrate for another 5 min.

The mass spectrometer was operated in positive mode. Nebulizer pressure was set to 25 psi and the dry gas (nitrogen) was heated up to 350  $^{\circ}\text{C}$  with a flow rate at 10 l/min. The m/z values were obtained with a full scan from 50 to 600 to detect the molecular weight of different TIA compounds. Most of the peaks in the spectra were identified with the comparison of calculated molecular weight and the comparison of retention time from available standard compounds. In the MS/MS mode, helium was used as the collision gas, and manual MS/MS programs were set up for specific time segments, which were the retention times of some TIAs. The m/z of the target compound was chosen in the time segment with 4 m/z window width, and 1 voltage was applied as the amplitude of the excitation, with a 27% cutoff of the m/z of the precursor ion.

## Metabolites quantification with HPLC

Five microliters of the alkaloid extract was injected into the Waters HPLC system consisted of two 510 pumps, a 717plus Autosampler, and a 996 Photo Diode Array (PDA) detector. And we used a Phenomenex Luna 3  $\mu\text{m}$  C18(2) LC column (250  $\times$  4.6 mm) to separate the alkaloids. The HPLC protocol was almost the same as the LC-MS protocol, except the flow rate was kept at 0.5 ml/min. Ajmalicine, serpentine, catharanthine were identified and quantified at 254 nm with their own standard compounds (Sigma-Aldrich). Hörhammericine, lochnericine, tabersonine, 16-hydroxytabersonine, 16-methoxytabersonine were identified at 329 nm with their own standard compounds (Sigma-Aldrich or house-made), and quantified with tabersonine standard compound [22]. The two putative compounds, epoxidized 16-hydroxytabersonine and lochnerinine, were also quantified with tabersonine standard compound.

## RT-qPCR

Total RNA was isolated from the frozen hairy root powder using TRIzol reagent according to manufacturer's instructions (Ambion RNA by Life Technologies). DNA was removed from the sample with TURBO DNA-free according to the manufacturer's instructions (Ambion RNA by Life Technologies). cDNA was synthesized from 500 ng RNA using random primers and GoScript reverse transcriptase according to manufacturer's instructions (Promega). A 'no-amplification control' (without reverse transcriptase) was performed for each sample. cDNA was diluted 10 times to 200  $\mu\text{L}$  with nuclease-free water. Each q-PCR reaction (20  $\mu\text{L}$ ) contained 1  $\mu\text{L}$  diluted cDNA, 1.25 pmol/mL mixed primers, 10  $\mu\text{L}$  SsoAdvanced SYBR green super mix (BIO-RAD) and nuclease-free water. The primers used for qPCR of *T16H* are 5'-*GCGGAACCTAACATTGCAGA*-3' and 5'-*GCACATCAACAAGGTCCTCC*-3'. The qPCR primer pairs for *16OMT* are 5'-*CTTGTTTGAGGGCTTGGCTT*-3' and 5'-

*TCAAACATGTACCTGCAACA-3'*. The primers used for the other genes were previously described [16]. The q-PCR amplifications were carried out in BIO-RAD CFX Connect™ Real-Time PCR Detection System with the program: 10 min at 95 °C, 40 cycles of 15 s at 95 °C, 1 min at 60 °C. The relative gene expression was quantified by using the comparative threshold cycle CT method as previously described [23]. The *40S ribosomal protein S9 (RPS9)* was used for the control gene [24].

### Statistical analysis

Data were analyzed using the Student's t-test.

## Results

### Genetic engineering of T16H and 16OMT in *C. roseus* hairy root

Stems of *C. roseus* seedling were infected with *Agrobacterium rhizogenes* ATCC 15834 carrying the plasmid *pTA7002/T16H/16OMT*. After selection on solid media containing hygromycin followed by six sub-culture cycles in liquid hairy root media, two stable transgenic hairy root lines #S3 and #S8 were obtained and used for further analysis. The expression of *T16H* and *16OMT* were controlled by the same glucocorticoid-inducible promoter [19]. Both genes showed increased expression in the two induced hairy roots compared to the uninduced ones by RT-qPCR. In the induced #S3 hairy roots, *T16H* mRNA was increased  $452.75 \pm 64.07$  times and *16OMT* mRNA was increased  $27.14 \pm 6.39$  times compared to the uninduced control. #S8 hairy roots demonstrated a  $107.45 \pm 19.16$  fold increase in *T16H* mRNA level and  $149.61 \pm 30.12$  fold increase in *16OMT* after 72 h induction. Noticeably, #S3 hairy roots showed a much higher overexpression fold change in *T16H* than in *16OMT*, while #S8 hairy roots revealed

a lower overexpression fold change in *T16H* than in *16OMT*. These differences based on mRNA analysis by qPCR could be due to the differences in the background level of transcripts in the two different hairy roots lines. The significant clonal variation between these two transgenic hairy roots may also result from the random T-DNA insertion into the nuclear chromosome mediated by *Agrobacterium* transformation. Nevertheless, the use of the glucocorticoid inducible promoter system in this study allows us to investigate the effects of increased expression to that of the control within the same background.

### **Identification of new metabolites in *C. roseus* hairy roots overexpressing T16H and 16OMT**

Tabersonine is an important branch point in vindoline production. In order to divert tabersonine into the vindoline pathway, the first two enzymes in vindoline pathway *T16H* and *16OMT* (Figure 2.3) were engineered into *C. roseus* hairy roots under the control of inducible promoter system. Due to the inaccessibility of TIAs intermediates 16-hydroxytabersonine and 16-methoxytabersonine on market, we chose to use *S. cerevisiae* as a production host to prepare the standard molecules. *S. cerevisiae* does not naturally produce TIAs due to the lack of the terpenoid pathway derived from geranyl pyrophosphate (GPP) [25], therefore the separation of TIAs from other yeast native metabolites is much easier since the chemical structures of these contaminants are very different from TIAs. Functional expression of microsomal plant P450s has precedence in yeast [26]. In order to obtain 16-hydroxytabersonine and 16-methoxytabersonine standards for LC/MS, a *S. cerevisiae* strain *WAT11* transformed with the plasmid *pESC-Leu-T16H-16OMT* expressing T16H and 16OMT was fed with tabersonine. The two alkaloids, 16-hydroxytabersonine and 16-methoxytabersonine, together with the substrate, tabersonine sharing the  $\alpha$ -methyleneindoline structure, were identified in yeast extracts by LC/MS at 329 nm

according to their molecular weights (Figure 2.4). The yeast extracted alkaloids were subsequently used for analyze hairy root samples.

By comparing the chromatograms of the induced and uninduced hairy roots extracts, four new peaks were detected in the induced hairy roots compared to the uninduced control (Figure 2.4). Peak 1 and 2 in the induced roots showed the same retention time as 16-hydroxytabersonine and 16-methoxytabersonine from the yeast extracts. Mass to charge ratio ( $m/z$ ) of peak 1 and 2 further confirmed that peak 1 is the T16H-catalyzed product 16-hydroxytabersonine, and peak 2 is the 16OMT-catalyzed product 16-methoxytabersonine. Interestingly, two unexpected peaks (Figure 2.4, peak 3 and 4) showed up in the induced hairy roots. Based on the  $m/z$  value of each peak, the molecular weight is 16 higher than that of 16-hydroxytabersonine and 16-methoxytabersonine respectively, suggesting the new metabolites could likely be the oxidized forms of 16-hydroxytabersonine and 16-methoxytabersonine.

The extremely low abundance of the two new TIA metabolites in the hairy roots made purification infeasible. In order to collect more evidences to confirm their chemical structures, MS/MS was performed for the two new TIAs, as well as other five related TIAs, including tabersonine, lochnericine, 16-hydroxytabersonine, 16-methoxytabersonine, and 19-hydroxytabersonine (Figure 2.4c). Comparing the fragment patterns of tabersonine, 16-hydroxytabersonine, and 16-methoxytabersonine, we could see that the modifications at the 16 position in the indole moiety did not change the fragment patterns. There was only single abundant fragment ion  $[M+H-32]^+$  in each of their fragment spectra. On the other hand, the modifications in the terpene moiety changed the fragment patterns. For example, by adding a hydroxyl group at the 19 position (comparing the chromatograms of tabersonine and 19-hydroxytabersonine in Figure 2.4c), besides the most abundant  $[M+H-32]^+$  peak, a second

abundant  $[M+H-18]^+$  appeared; by forming an epoxide ring between the 6 and 7 positions (comparing the chromatograms of tabersonine and lochnericine in Figure 2.4c), the fragment pattern changed dramatically, with four abundant peaks appeared ( $[M+H-18]^+$ ,  $[M+H-32]^+$ ,  $[M+H-60]^+$ ,  $[M+H-125]^+$ ). The fragment patterns of the two putative compounds were similar to the one of lochnericine, they all contained abundant  $[M+H-18]^+$ ,  $[M+H-32]^+$ ,  $[M+H-60]^+$ ,  $[M+H-125]^+$  peaks, so the two unknown TIAs may have an epoxide ring occurring at the 6 and 7 positions. Meanwhile, since the modifications at the 16 position in the indole moiety kept the most abundant fragment ion  $[M+H-32]^+$ , the putative 6,7-epoxidized-16-hydroxytabersonine and 6,7-epoxidized-16-methoxytabersonine contained more  $[M+H-32]^+$  fragment ion than lochnericine. Altogether, those pieces of evidences provided strong supporting information to indicate that the two unknown metabolites could be 6,7-epoxidized-16-hydroxytabersonine (peak 3 in Figure 2.4) and 6,7-epoxidized-16-methoxytabersonine (peak 4 in Figure 2.4). In addition, T6,7-epoxidase (T6,7E) activity has been demonstrated in *C. roseus* hairy roots and this enzyme is responsible for generating the epoxidized derivatives of tabersonine (Figure 2.3) [27], we hypothesize that T6,7E can convert 16-hydroxytabersonine and 16-methoxytabersonine to epoxidized 16-hydroxytabersonine and epoxidized 16-methoxytabersonine (also named as lochnerinine).

### **Quantification of alkaloids in the induced and uninduced hairy roots expressing T16H and 16OMT**

After identifying the four new metabolites in the transgenic hairy roots expressing *T16H* and *16OMT*, the yields of all the detected alkaloids in the induced and uninduced hairy roots were quantified by HPLC (Figure 2.5). 16-hydroxytabersonine was increased from the undetectable level in the uninduced hairy roots to  $0.15 \pm 0.88$  mg/g dry weight (DW) in the 72 h

induced #S3 hairy roots and to  $0.05 \pm 0.00$  mg/g DW in the induced #S8 hairy roots. The concentration of 16-methoxytabersonine was also increased to  $0.15 \pm 0.06$  and  $0.42 \pm 0.04$  mg/g DW after 72 h induction of *T16H* and *16OMT* in #S3 and #S8 hairy roots respectively, while the uninduced control lines barely showed accumulation of 16-methoxytabersonine (Figure 2.5a). The induced #S3 hairy roots produced more 16-hydroxytabersonine but less 16-methoxytabersonine than the induced #S8 hairy roots. In #S3 hairy roots, the fold increase of *T16H* transcripts was much higher than that of *16OMT* after 72 h induction. Therefore the intermediate 16-hydroxytabersonine was accumulated possibly due to the relatively low expression of *16OMT*. In contrast, 16-hydroxytabersonine could be rapidly converted to 16-methoxytabersonine in the induced #S8 possibly due to the high-level expression of *16OMT* in this hairy root line. The proposed epoxidized 16-hydroxytabersonine and lochnerinine showed very similar accumulation trends with 16-hydroxytabersonine and 16-methoxytabersonine, which further demonstrated that these two metabolites are most likely the derivatives of 16-hydroxytabersonine and 16-methoxytabersonine.

The metabolite levels in the 19-O-acetyhörhammericine pathway including lochnericine, hörhammericine and 19-hydroxytabersonine (Figure 2.3) were significantly decreased in the two induced hairy roots compared to the uninduced ones (Figure 2.5b). Tabersonine level was not changed in #S3 roots after induction, while a 26% decrease in tabersonine was observed in the induced #S8 hairy roots. Figure 2.6 shows the obvious accumulation of the total measured vindoline pathway metabolites in the induced roots and the great decrease in the total 19-O-acetyhörhammericine pathway metabolites after overexpressing *T16H* and *16OMT*. Such changes were expected because it indicates tabersonine has been converted into the desired intermediates in vindoline pathway and less flow was channeled to the root-specific 19-O-

acetyhörhammericine pathway. However, the aspidosperma alkaloid pool consisting of tabersonine and its derivatives (including vindoline pathway intermediates) did not show significant changes after 72 h induction in both root lines (Figure 2.6). This indicates that no additional carbon was pulled into tabersonine biosynthesis.

Interestingly, catharanthine and ajmalicine were increased by 23% and 21% in #S8 hairy roots after induction, while serpentine showed 43% increase after induction in #S3 hairy roots (Figure 2.5c). Small but significant increases in total alkaloid pool (all measured alkaloids), corynanthe alkaloid pool (serpentine and ajmalicine), and iboga alkaloid (catharanthine) were noted in #S8 hairy roots after 72 h induction, while in #S3 lines the total alkaloid level was slightly reduced, and corynanthe alkaloid and iboga alkaloid levels remained unchanged after induction (Figure 2.6). These changes outside the tabersonine metabolic pathway may result from the complex regulations involved in the TIA pathway. The transcriptional changes of the TIA genes and regulators were therefore further analyzed by RT-qPCR.

### **Transcriptional changes of the TIA pathway genes and regulators after inducing T16H and 16OMT in *C. roseus* hairy roots**

Transcriptional alteration of TIA pathway genes and transcriptional factors was assessed in the 72 hr induced and uninduced hairy root lines using RT-qPCR (Figure 2.7). Interestingly, the 19-O-acetyhörhammericine pathway genes *T19H* and *MAT* were significantly down regulated in both #S3 and #S8 induced hairy roots compared to the uninduced control. Similarly, it was reported that the overexpression of the last gene *DAT* from vindoline pathway in *C. roseus* hairy roots can inhibit the activity of roots native *MAT* [28].

From the previous studies, genetic perturbation of the TIA genes or transcription factors in *C. roseus* hairy roots caused complex transcriptional changes of the TIA pathway genes and

the associated regulators [16, 29, 30]. The mRNA levels of the two alkaloid pathway genes *STR* and *SGD* upstream of tabersonine did not show any changes after overexpressing *T16H* and *I6OMT* in #S8 hairy roots, while *SGD* were up regulated in #S3 hairy roots after induction. The transcription of two positive regulators *ORCA2* and *ORCA3* were not changed in #S3 hairy roots after induction, while *ORCA2* was up regulated and *ORCA3* was down regulated in the induced #S8 hairy roots compare to the uninduced control. In addition, overexpression of *T16H* and *I6OMT* also triggered the negative regulation response of TIA pathway. Notably, the negative TIA regulators *ZCTs* were greatly up regulated in both root lines after 72 h induction, with the only exception that *ZCT3* was not changed in the induced #S3 hairy roots compared to the control.

## Discussion

In this study, the first two genes *T16H* and *I6OMT* were successfully introduced into *C. roseus* hairy roots with the aim of initiating production of metabolite intermediates of the vindoline pathway. The induced overexpression of *T16H* and *I6OMT* resulted in the synthesis of 16-hydroxytabersonine and 16-methoxytabersonine that are the first two intermediates in the vindoline pathway. Previous efforts have not identified any metabolite along the pathway from tabersonine to vindoline in other hairy root lines [31]. The production of these alkaloids in this study is the first report of these metabolites in hairy roots. Interestingly, two unknown alkaloids were detected in the two *T16H* and *I6OMT* induced hairy roots. The presence of unexpected metabolites after introducing the leaf-specific genes into hairy roots might be caused by two reasons: the naturally occurring metabolites were catalyzed by the newly introduced enzymes due to the promiscuous substrate specificity, or the catalytic products of the new enzymes were converted by the native enzymes. The molecular weight of the two new compounds is 16 higher

than that of 16-hydroxytabersonine and 16-methoxytabersonine individually. Moreover, these two previously unknown metabolites showed similar accumulation trends with 16-hydroxytabersonine and 16-methoxytabersonine. Thus they are likely the oxidized form of 16-hydroxytabersonine and 16-methoxytabersonine. In hairy roots, the tabersonine metabolic pathway involves four enzymes including T19H, MAT, T6,7E, and T6,7R, among which, T19H [14] and T6,7E [27] oxidize tabersonine and its derivatives at different positions (Figure 2.3). The two unknown metabolites showed similar fragmentation patterns with the tabersonine derivative with an epoxide ring on the 6,7 position based on MS/MS analysis (Figure 2.4c). In addition, feeding 16-hydroxytabersonine and 16-methoxytabersonine to *S. cerevisiae* WAT11 overexpressing T19H did not result in the conversion of these two metabolites (data not shown here). The function of T6,7E in crude protein from hairy roots was characterized even though the sequence of T6,7E remains unknown. The metabolites lochnericine, hörhammericine, and 19-O-acetyl hörhammericine are all oxygenated derivatives of tabersonine at position 6,7 which results from T6,7E enzyme activity [27]. Thus this root-dominant enzyme T6,7E is proposed to have a broader substrate specificity and contributes to the production of the two possible epoxidized 16-hydroxytabersonine and 16-methoxytabersonine in the induced hairy roots.

Lochnericine and hörhammericine both had significant lower levels in the induced hairy roots compared to the uninduced controls. This indicates that the induced *T16H* is competing for the precursor tabersonine with the root-specific enzymes T6,7E or T19H. Moreover, the down regulation of *T19H* and *MAT* mRNA levels after induction is also responsible for the decrease in the metabolites in this pathway branch. However, it is unclear whether these down regulations resulted from transcription-level control. Noticeably, the vindoline pathway and the 19-O-acetylhörhammericine pathway not only share the same precursor tabersonine, some enzymes in

these two pathways have certain homology in sequence or similar function. For example, MAT in the 19-O-acetylhorhammericine pathway shares 63% nucleic acid and 78% amino acid identities with DAT in the vindoline pathway [32]. They catalyze acetyl transfer reaction at different hydroxyl positions (Figure 2.3). T16H and T19H can hydrolyze the substrate tabersonine at 16 and 19 positions, respectively. T3O in vindoline pathway and T6,7E from 19-O-acetylhorhammericine pathway both catalyze oxygenate reaction. T6,7E epoxidizes the double bond at 6,7 position of tabersonine and tabersonine derivatives. T3O oxygenates the double bond at position 3 of tabersonine and its derivatives. The engineering of the terminal step of vindoline biosynthesis *DAT* inhibited the activity of roots MAT in *C. roseus*. These results showed a sign of competitive relationship between the vindoline pathway and 19-O-acetylhorhammericine pathway. The total aspidosperma alkaloids accumulation did not change much after overexpressing *T16H* and *16OMT*, because the increase in metabolites in the vindoline pathway is compensated for the decrease in lochnericine and hörhammericine. Nevertheless, it indicates tabersonine flux was successfully diverted to the desired vindoline pathway from the root native 19-O-acetylhorhammericine pathway.

Moreover, engineering of *T16H* and *16OMT* also caused metabolite level changes beyond tabersonine metabolic pathway. The level of corynanthe and iboga alkaloids, which are alkaloids that diverge from a common precursor (strictosidine aglycone), increased in the induced #S8 hairy roots, while the induced hairy roots of #S3 did not show significant changes in these two alkaloid pools. In addition, the total measured alkaloids showed opposite accumulating trends in two induced hairy roots #S3 and #S8. It is hypothesized that a much higher overexpression level of *T16H* in the induced #S3 hairy roots than that in the induced #S8 hairy roots caused a higher metabolic burden or a negative effect on total alkaloids accumulation.

Many previous *C. roseus* engineering studies showed that the modification of TIA genes usually led to complex transcriptional changes in other pathway genes and transcription factors [16, 17, 33]. Similarly, the complicated mRNA changes in the TIA pathway regulators were triggered after inducing two vindoline pathway genes *T16H* and *16OMT*. The relative expression of the positive regulators, ORCAs, and the negative regulators, ZCTs, have been correlated with the jasmonate-dependent alkaloid biosynthesis [34]. In this study, ZCTs showed a great downregulation in the induced hairy roots compared to the uninduced control, while ORCAs revealed complex changes. The tight control of TIA biosynthesis by transcription factors is a strategy used by plants to balance growth and defense. In the future, a better understanding of how to bypass negative regulation is necessary to successfully engineer the TIA pathway for increased alkaloid accumulation.

### **Conclusion**

The vindoline pathway derived from the precursor tabersonine is spatially regulated in *C. roseus*. The inactivity of this pathway in hairy roots and cell suspension cultures is the biggest barrier for the production of vindoline, vinblastine and vincristine. In the present study, we describe the expression of the first two genes, *T16H* and *16OMT*, in the vindoline pathway that initiates the channeling of tabersonine to vindoline pathway intermediates in *C. roseus* hairy root cultures. It is not surprising that expressing only the first two genes of the vindoline pathway is not sufficient to make vindoline in hairy roots. Besides seeing the accumulation of the expected products of *T16H* and *16OMT*, two additional alkaloids were produced, and the concentrations of other root specific metabolites were significantly changed. This study also illustrates how introduction of *T16H* and *16OMT* triggered complex transcriptional responses, especially the up-regulation of the negative transcription factors. This study provides insights regarding how to

express the full vindoline biosynthesis pathway to generate vindoline accumulating hairy roots. In the future, eliminating reactions competing for tabersonine, removing root dominant enzymes such as T6,7E, and silencing the negative regulators in combination with overexpressing the vindoline pathway will be necessary to reach industrial relevant yields of vindoline production in hairy roots.

## Figures

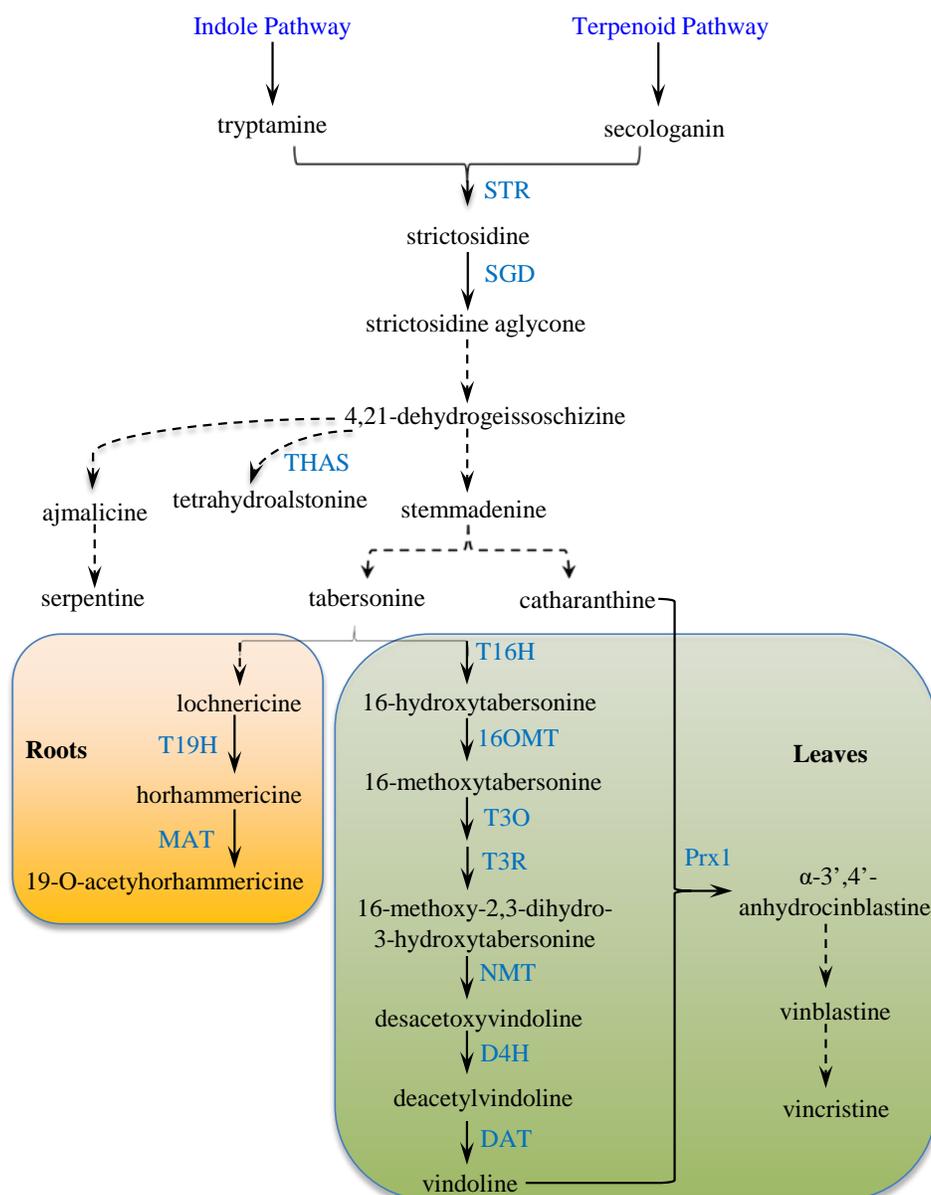


Figure 2.1 Terpene indole alkaloid pathway in *C. roseus*. STR, strictosidine synthase; SGD, strictosidine beta-glucosidase; THAS, tetrahydroalstonine synthase; T19H, tabersonine 19-hydroxylase; MAT, minovincininine 19-hydroxy-O-acetyltransferase; T16H, tabersonine 16-hydroxylase; 16OMT, 16-O-methyltransferase; T3O, tabersonine 3-oxygenase; T3R, tabersonine 3-reductase; NMT, N-methyltransferase; D4H, desacetoxyvindoline 4-hydroxylase; DAT, deacetylvindoline-4-O-acetyltransferase; Prx1, Peroxidase 1.

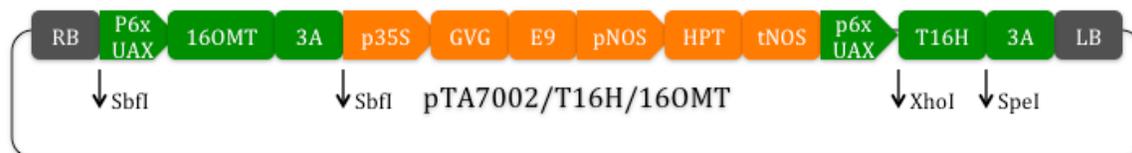


Figure 2.2 Plasmid construct. RB, right border of T-DNA; LB, left border of T-DNA; p35S, cauliflower mosaic virus 35S promoter; GVG, chimeric transcription factor containing GAL4 DNA-binding domain and VP16 transactivating domain and rat-GR HBD; E9, pea rbcS-E9 polyadenylation terminator sequence; pNOS, nopaline synthase promoter; HPT, hygromycin phosphotransferase gene; tNOS, nopaline synthetase polyadenylation sequence terminator; p6xUAX, GVG-regulated promoter; 3A, pea rbcS-3A polyadenylation terminator sequence.

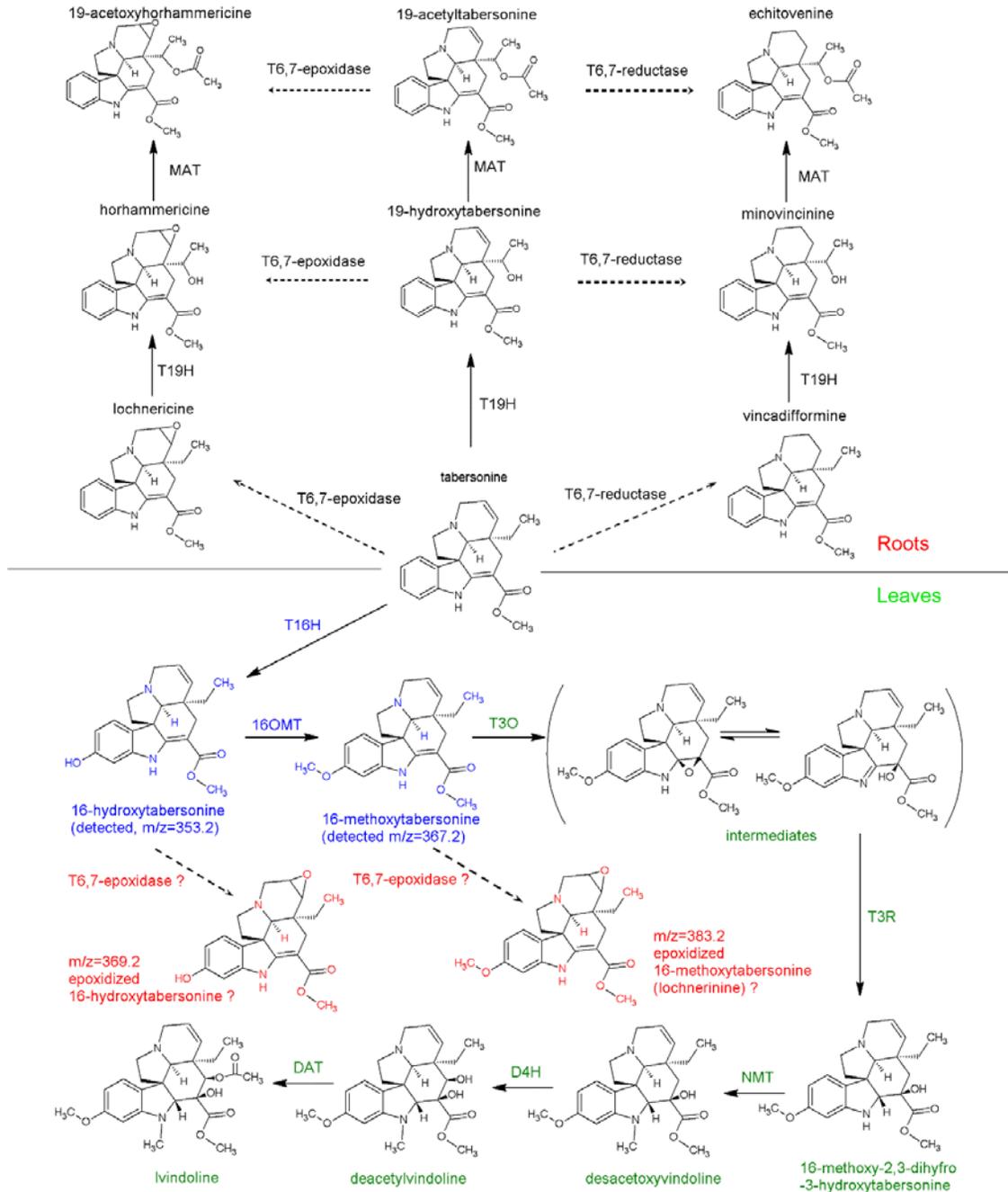
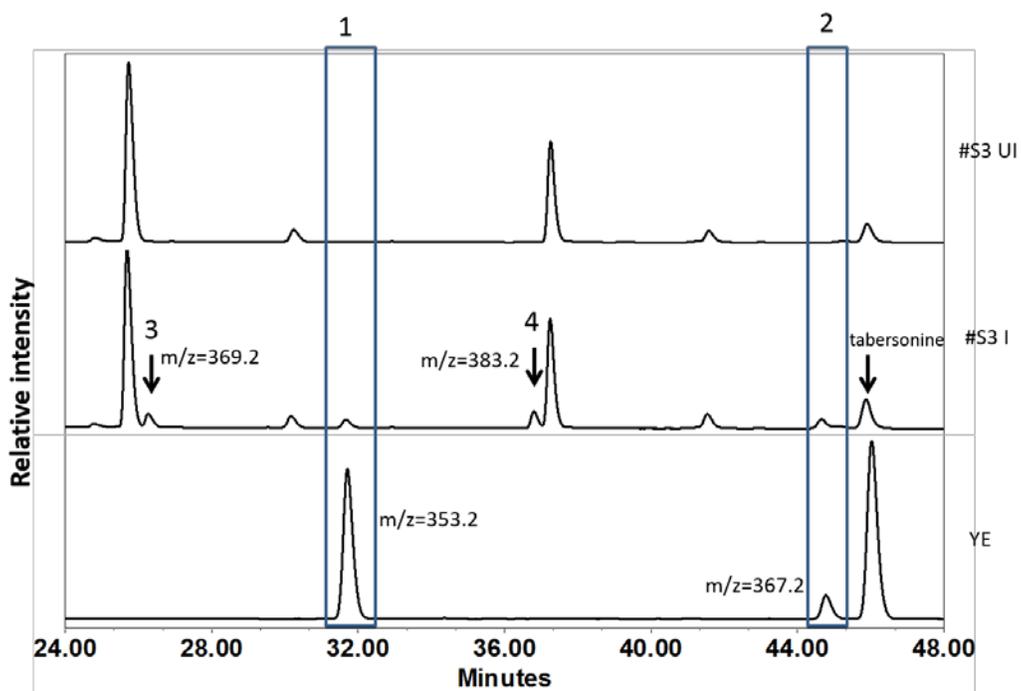
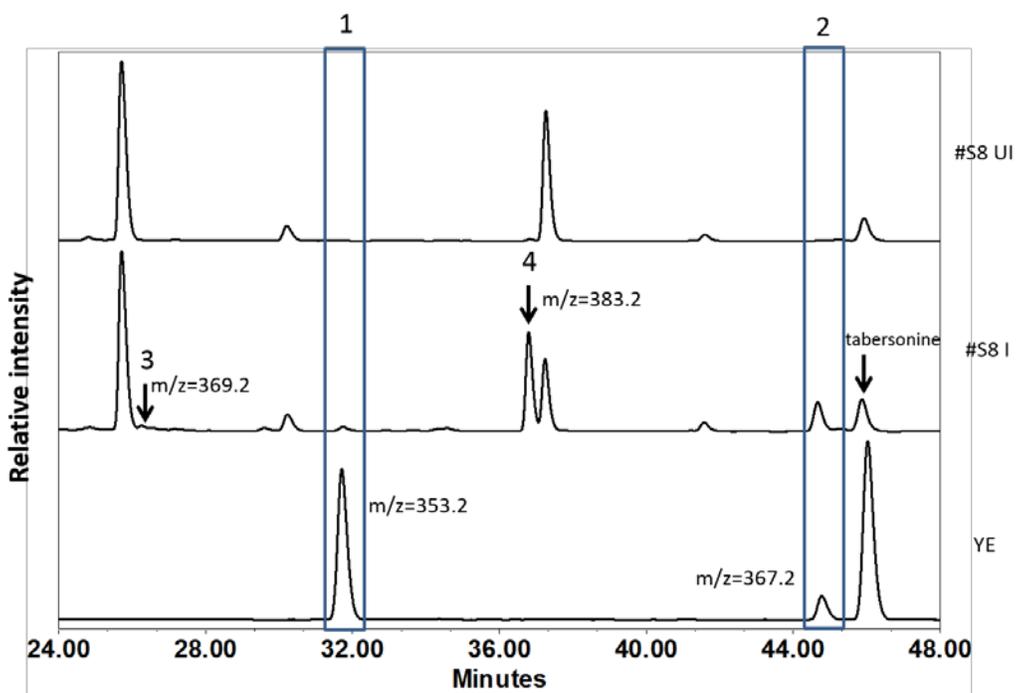


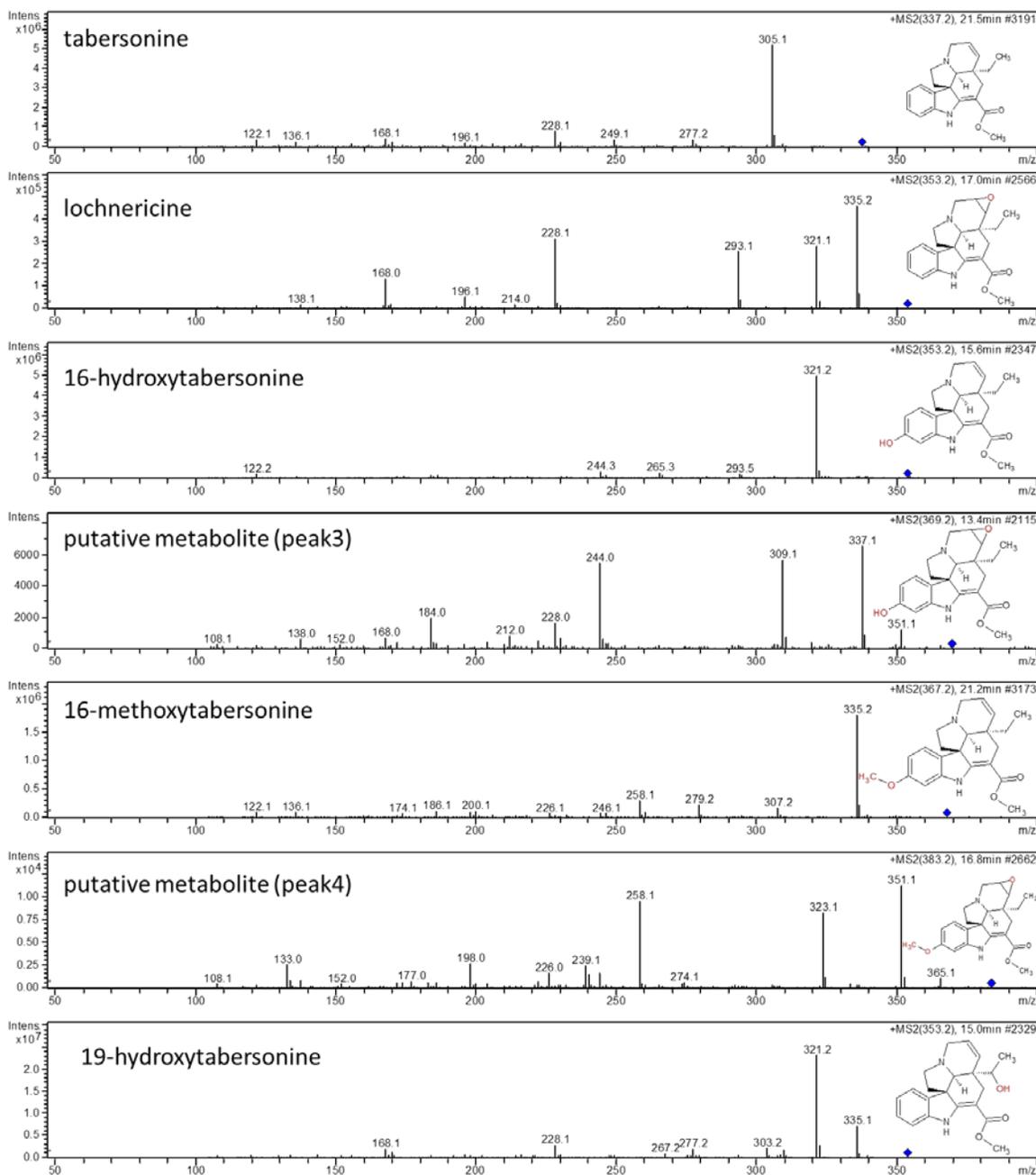
Figure 2.3 Tabersonine metabolism in roots and leaves. Metabolite and enzyme names showed in black indicate that they are naturally present in hairy roots. Blue ones represent the engineered enzymes and the detected metabolites after overexpressing the engineered genes in hairy roots. Red represents the hypothesized enzyme and metabolites after overexpressing *T16H* and *16OMT*. Green represents the enzymes or metabolites which are not present in the transgenic hairy roots.



(a)

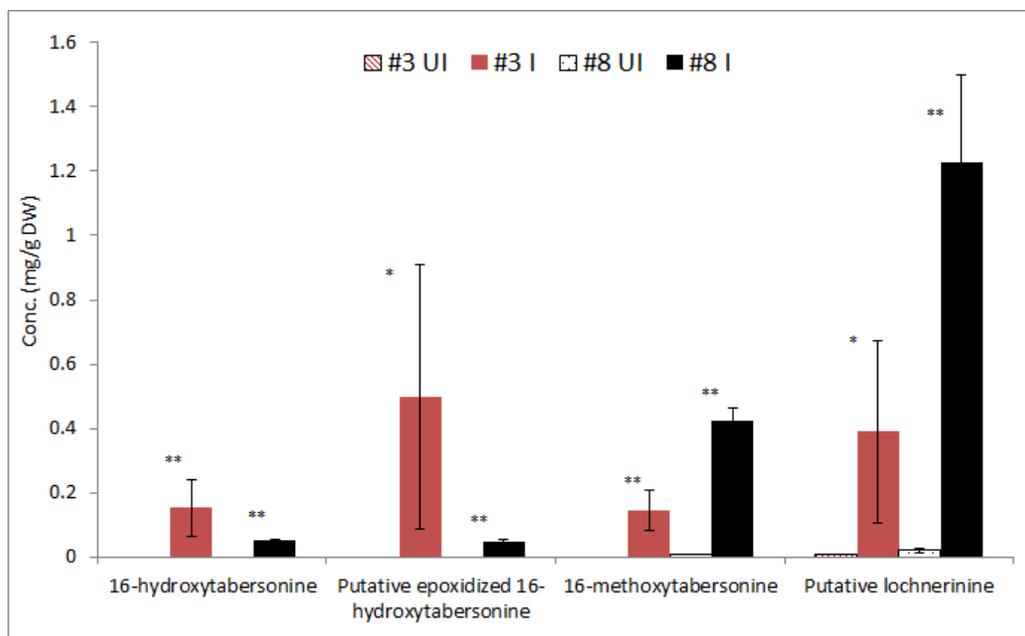


(b)

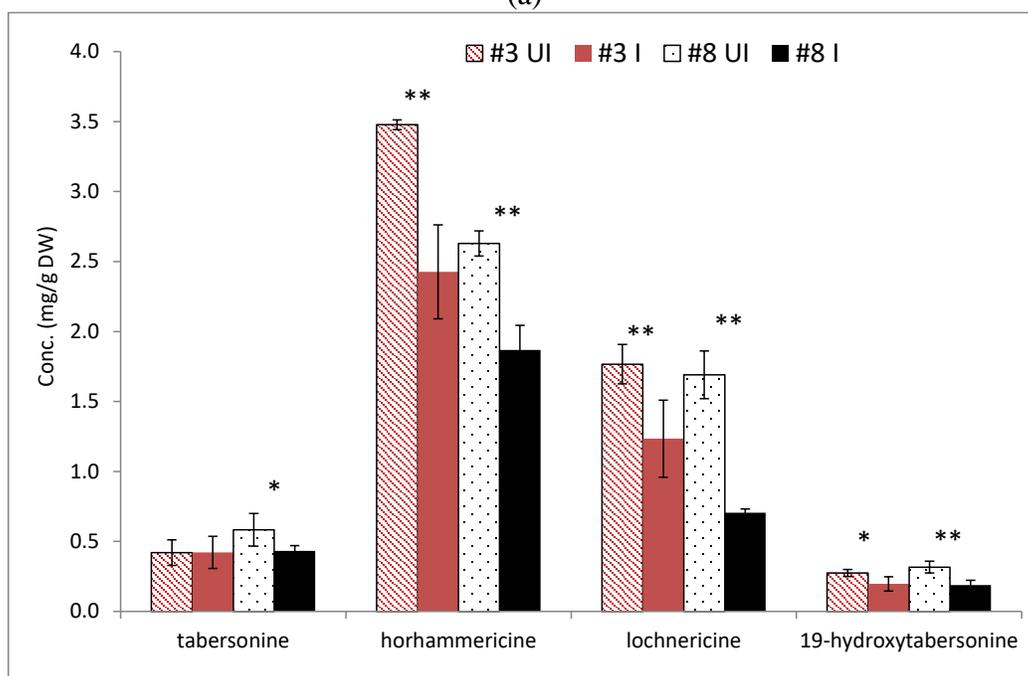


(c)

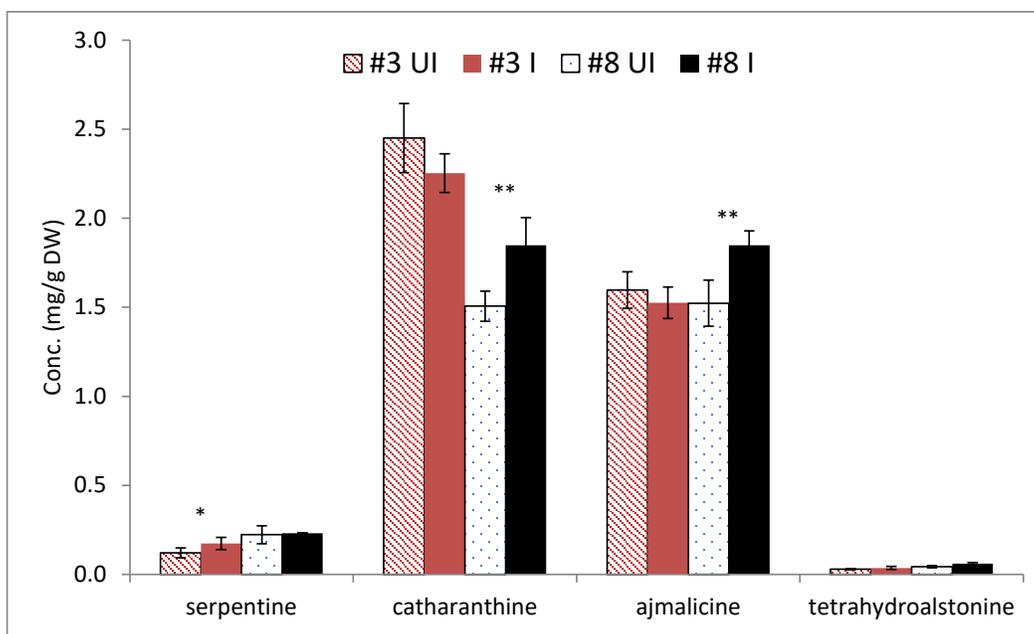
Figure 2.4 The selected LC chromatograms of alkaloid extracted from the hairy roots cell lines #S3 (a) and #S8 (b) induced (I) or un-induced (UI) for 72 hr, and the selected LC-MS/MS ion chromatograms of the extracted alkaloids (c). YE represents alkaloids extracts from the medium of tabersonine-supplemented yeast cell cultures expressing *pESC-Leu-T16H-16OMT*. Peak 1, 16-hydroxytabersonine. Peak 2, 16-methoxytabersonine. Peak 3, putative epoxidized 16-hydroxytabersonine. Peak 4, putative epoxidized 16-methoxytabersonine.



(a)



(b)



(c)

Figure 2.5 The alkaloid levels in the transgenic *Catharanthus roseus* hairy roots expressing T16H and 16OMT. The two hairy roots overexpressing T16H and 16OMT (#S3-I and #S8-I) were induced with 3  $\mu$ M dexamethasone during exponential growth phase and harvested 72 h later. The controls (#S3-UI and #S8-UI) were the same root line fed with an equivalent volume of ethanol. HPLC analysis was used to determine the concentrations of the new metabolites: (a) metabolites appearing after overexpressing T16H and 16OMT (16-hydroxytabersonine, 16-methoxytabersonine, proposed epoxidized 16-hydroxytabersonine and lochnerinine), (b) root aspidosperma metabolites (tabersonine, lochnericine, hörhammericine, and 19-hydroxytabersonine) and, (c) other TIAs (serpentine, ajmalicine, catharanthine, and tetrahydroalstonine). “\*” ( $p < 0.1$ ) and “\*\*” ( $p < 0.05$ ) indicates significant results between the UI and I cultures. Error bars represent the standard deviation of triplicate cultures.

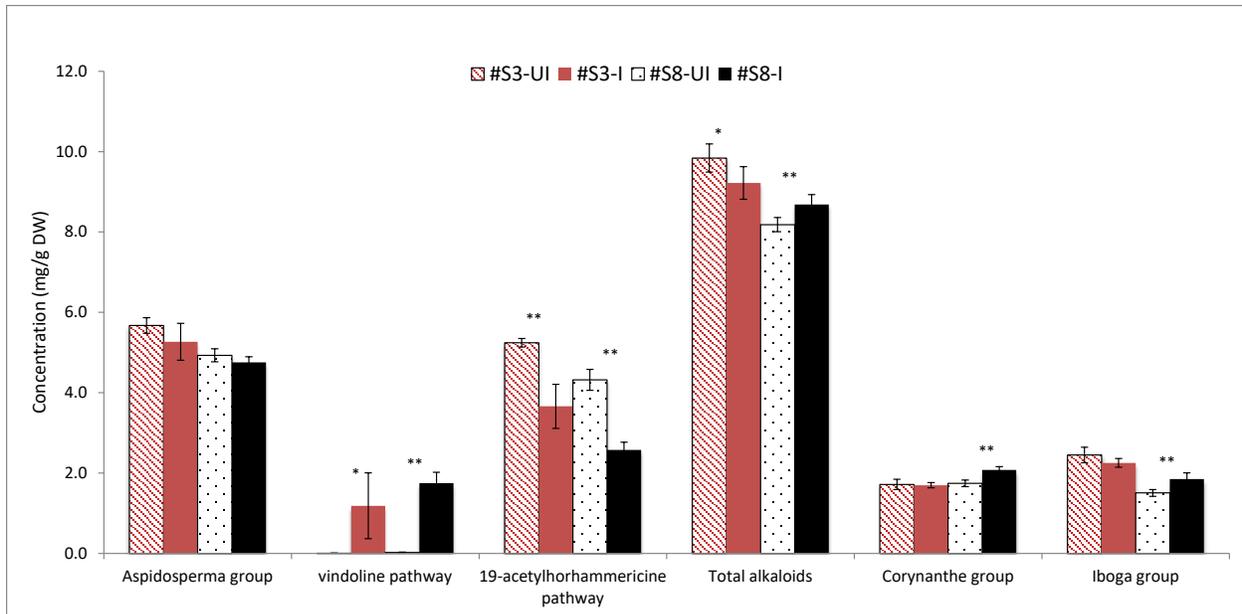


Figure 2.6 The terpenoid indole alkaloid pools in the transgenic *Catharanthus roseus* hairy roots expressing T16H and 16OMT. The hairy roots overexpressing T16H and 16OMT (#S3-I and #S8-I) were induced with 3 $\mu$ M dexamethasone during exponential growth phase and harvested 72 h later. The controls (#S3-UI and #S8-UI) were the same root line fed with an equivalent volume of ethanol. HPLC analysis was used to determine the TIA pools belong to aspidosperma family (tabersonine, lochnericine, hörhammericine, 16-hydroxytabersonine, 16-methoxytabersonine, putative epoxidized 16-hydroxytabersonine and putative lochnericine), vindoline pathway (16-hydroxytabersonine, 16-methoxytabersonine, putative epoxidized 16-hydroxytabersonine and putative lochnericine), 19-acetylhorhammericine pathway (lochnericine and hörhammericine), the total alkaloids (all alkaloids measured), corynanthe family (serpentine and ajmalicine), and iboga family (catharanthine). Error bars represent the standard deviation of triplicate cultures. “\*” ( $p < 0.1$ ) and “\*\*” ( $p < 0.05$ ) indicates significant results between the UI and I cultures. Error bars represent the standard deviation of triplicate cultures.

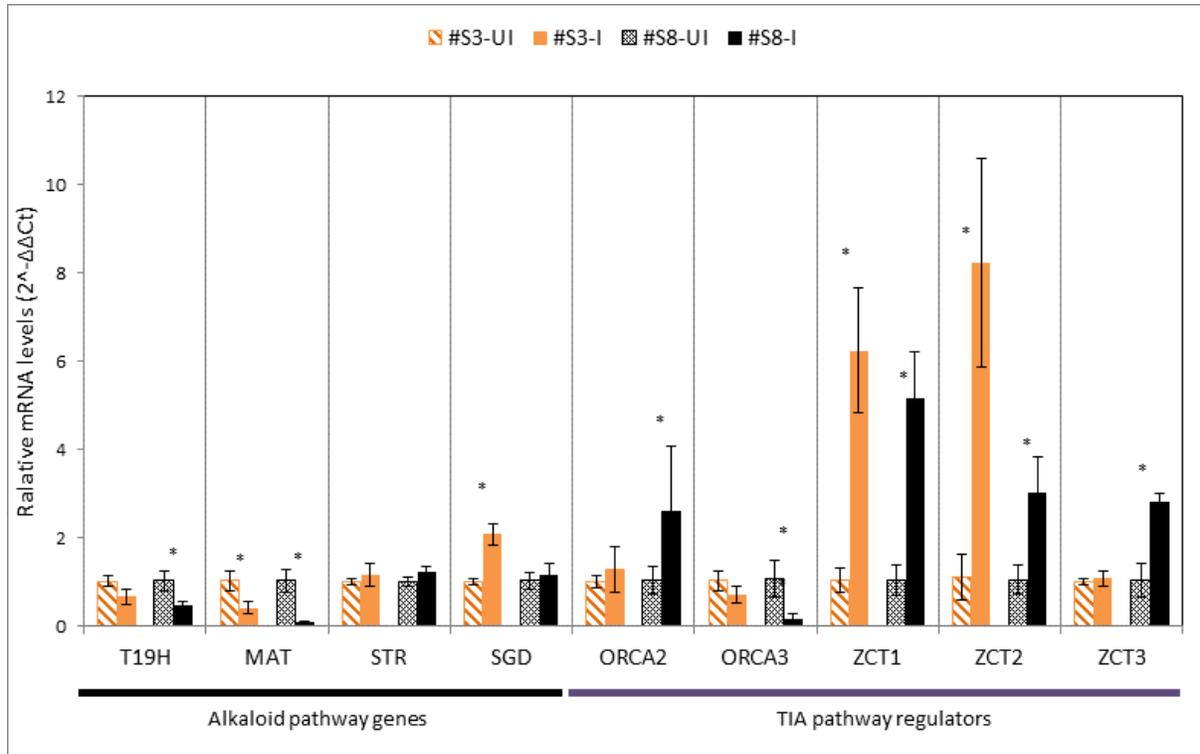


Figure 2.7 RT-qPCR of the TIA pathway genes and transcription factors mRNA transcripts fold change in the 72 h induced root line compared to the uninduced root line. Data represents the mean of triplicate  $\pm$  standard deviation. “\*” indicates significant results ( $p < 0.1$  and two fold change) between the UI and I cultures. Error bars represent the standard deviation of triplicate cultures.

## References

1. Murata, J. and V. De Luca, Localization of tabersonine 16-hydroxylase and 16-OH tabersonine-16-O-methyltransferase to leaf epidermal cells defines them as a major site of precursor biosynthesis in the vindoline pathway in *Catharanthus roseus*. *Plant J*, 2005. **44**(4): p. 581-94.
2. Nobili, S., et al., Natural compounds for cancer treatment and prevention. *Pharmacological Research*, 2009. **59**(6): p. 365-378.
3. Chi, S.M., et al., Theoretical insight into the structural mechanism for the binding of vinblastine with tubulin. *Journal of Biomolecular Structure & Dynamics*, 2015. **33**(10): p. 2234-2254.
4. Sears, J.E. and D.L. Boger, Total Synthesis of Vinblastine, Related Natural Products, and Key Analogues and Development of Inspired Methodology Suitable for the Systematic Study of Their Structure-Function Properties. *Accounts of Chemical Research*, 2015. **48**(3): p. 653-662.
5. Georgiev, M.I., A.I. Pavlov, and T. Bley, Hairy root type plant in vitro systems as sources of bioactive substances. *Applied Microbiology and Biotechnology*, 2007. **74**(6): p. 1175-1185.
6. Pasquali, G., D.D. Porto, and A.G. Fett-Neto, Metabolic engineering of cell cultures versus whole plant complexity in production of bioactive monoterpene indole alkaloids: Recent progress related to old dilemma. *Journal of Bioscience and Bioengineering*, 2006. **101**(4): p. 287-296.
7. McKnight, T.D., et al., Nucleotide sequence of a cDNA encoding the vacuolar protein strictosidine synthase from *Catharanthus roseus*. *Nucleic Acids Res*, 1990. **18**(16): p. 4939.
8. O'Connor, S.E. and J.J. Maresh, Chemistry and biology of monoterpene indole alkaloid biosynthesis. *Nat Prod Rep*, 2006. **23**(4): p. 532-47.
9. Miettinen, K., et al., The seco-iridoid pathway from *Catharanthus roseus*. *Nature Communications*, 2014. **5**.
10. Costa, M.M., et al., Molecular cloning and characterization of a vacuolar class III peroxidase involved in the metabolism of anticancer alkaloids in *Catharanthus roseus*. *Plant Physiol*, 2008. **146**(2): p. 403-17.
11. Courdavault, V., et al., A look inside an alkaloid multisite plant: the *Catharanthus* logistics. *Current Opinion in Plant Biology*, 2014. **19**: p. 43-50.
12. Roepke, J., et al., Vinca drug components accumulate exclusively in leaf exudates of Madagascar periwinkle. *Proceedings of the National Academy of Sciences of the United States of America*, 2010. **107**(34): p. 15287-15292.

13. Shanks, J.V., et al., Quantification of metabolites in the indole alkaloid pathways of *Catharanthus roseus*: Implications for metabolic engineering. *Biotechnology and Bioengineering*, 1998. **58**(2-3): p. 333-338.
14. Giddings, L.A., et al., A Stereoselective Hydroxylation Step of Alkaloid Biosynthesis by a Unique Cytochrome P450 in *Catharanthus roseus*. *Journal of Biological Chemistry*, 2011. **286**(19): p. 16751-16757.
15. Qua, Y., et al., Completion of the seven-step pathway from tabersonine to the anticancer drug precursor vindoline and its assembly in yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 2015. **112**(19): p. 6224-6229.
16. Sun, J. and C.A. Peebles, Engineering overexpression of ORCA3 and strictosidine glucosidase in *Catharanthus roseus* hairy roots increases alkaloid production. *Protoplasma*, 2015.
17. Peebles, C.A., et al., Transcriptional response of the terpenoid indole alkaloid pathway to the overexpression of ORCA3 along with jasmonic acid elicitation of *Catharanthus roseus* hairy roots over time. *Metab Eng*, 2009. **11**(2): p. 76-86.
18. Hughes, E.H., et al., Metabolic engineering of the indole pathway in *Catharanthus roseus* hairy roots and increased accumulation of tryptamine and serpentine. *Metab Eng*, 2004. **6**(4): p. 268-76.
19. Hughes, E.H., et al., Characterization of an inducible promoter system in *Catharanthus roseus* hairy roots. *Biotechnol Prog*, 2002. **18**(6): p. 1183-6.
20. Bhadra, R., S. Vani, and J.V. Shanks, Production of indole alkaloids by selected hairy root lines of *Catharanthus roseus*. *Biotechnol Bioeng*, 1993. **41**(5): p. 581-92.
21. Peebles, C.A.M., et al., Transient effects of overexpressing anthranilate synthase alpha and beta subunits in *Catharanthus roseus* hairy roots. *Biotechnology Progress*, 2005. **21**(5): p. 1572-1576.
22. Morgan, J.A., et al., Effects of buffered media upon growth and alkaloid production of *Catharanthus roseus* hairy roots. *Applied Microbiology and Biotechnology*, 2000. **53**(3): p. 262-265.
23. Shalel-Levanon, S., K.Y. San, and G.N. Bennett, Effect of oxygen, and ArcA and FNR regulators on the expression of genes related to the electron transfer chain and the TCA cycle in *Escherichia coli*. *Metab Eng*, 2005. **7**(5-6): p. 364-74.
24. Menke, F.L., et al., Involvement of the octadecanoid pathway and protein phosphorylation in fungal elicitor-induced expression of terpenoid indole alkaloid biosynthetic genes in *catharanthus roseus*. *Plant Physiol*, 1999. **119**(4): p. 1289-96.
25. Brown, S., et al., De novo production of the plant-derived alkaloid strictosidine in yeast. *Proc Natl Acad Sci U S A*, 2015. **112**(11): p. 3205-10.

26. Pompon, D., et al., Yeast expression of animal and plant P450s in optimized redox environments. *Methods Enzymol*, 1996. **272**: p. 51-64.
27. Rodriguez, S., et al., Jasmonate-induced epoxidation of tabersonine by a cytochrome P-450 in hairy root cultures of *Catharanthus roseus*. *Phytochemistry*, 2003. **64**(2): p. 401-9.
28. Magnotta, M., et al., Expression of deacetylvindoline-4-O-acetyltransferase in *Catharanthus roseus* hairy roots. *Phytochemistry*, 2007. **68**(14): p. 1922-1931.
29. Li, C.Y., et al., The ORCA2 transcription factor plays a key role in regulation of the terpenoid indole alkaloid pathway. *BMC Plant Biol*, 2013. **13**: p. 155.
30. Sun, J., et al., Examining the transcriptional response of overexpressing anthranilate synthase in the hairy roots of an important medicinal plant *Catharanthus roseus* by RNA-seq. *BMC Plant Biol*, 2016. **16**(1): p. 108.
31. Shanks, J.V., et al., Quantification of metabolites in the indole alkaloid pathways of *catharanthus roseus*: implications for metabolic engineering. *Biotechnol Bioeng*, 1998. **58**(2-3): p. 333-8.
32. Laflamme, P., B. St-Pierre, and V. De Luca, Molecular and biochemical analysis of a Madagascar periwinkle root-specific minovincinine-19-hydroxy-O-acetyltransferase. *Plant Physiol*, 2001. **125**(1): p. 189-98.
33. Li, C.Y., et al., CrBPF1 overexpression alters transcript levels of terpenoid indole alkaloid biosynthetic and regulatory genes. *Front Plant Sci*, 2015. **6**: p. 818.
34. Goklany, S., et al., Jasmonate-Dependent Alkaloid Biosynthesis in *Catharanthus roseus* Hairy Root Cultures Is Correlated with the Relative Expression of Orca and Zct Transcription Factors. *Biotechnology Progress*, 2013. **29**(6): p. 1367-1376.

### CHAPTER III. METABOLIC PROFILING OF *CATHARANTHUS ROSEUS* HAIRY ROOT LINES OVEREXPRESSING TRANSCRIPTION ACTIVATORS FOR TERPENOID INDOLE ALKALOID PATHWAY

This chapter is revised from two published papers “The ORCA2 transcription factor plays a key role in regulation of terpenoid indole alkaloid pathway” and “CrBPF1 overexpression alters transcript levels of terpenoid indole alkaloid biosynthetic and regulatory genes”. The metabolite quantification work done by Shanks’ group was picked out, and I contributed in identification of some TIAs and their quantification.

Le Zhao<sup>1</sup>, Guy Sander<sup>2</sup>, Jacqueline Shanks<sup>1</sup>

<sup>1</sup>Dept. of Chemical and Biological Engineering, Iowa State University

<sup>2</sup>Dept. of Chemical Engineering, University of Minnesota Duluth

#### Introduction

Madagascar periwinkle *Catharanthus roseus* produces over 130 terpenoid indole alkaloids (TIAs), and is a main source for several high value TIAs as pharmaceutical compounds. Vinblastine and vincristine are the two TIAs which were found to have anti-cancer functions, and have been marketed and widely used in chemotherapeutic regimens in the treatment of lymphoma and leukemia [1]. And so far, *C. roseus* is the sole host to produce vincristine and vinblastine. Besides, some other TIAs accumulated in *C. roseus* are used to treat hypertension, such as ajmalicine and serpentine. The current commercial production of vinblastine and vincristine is a conventional and low-efficient method, extraction from plant leaves. The low contents of the target compounds make them expensive. Though the total chemical synthesis of vincristine and vinblastine has been studied [2, 3], it is still not economically feasible to replace the plant extraction. Another alternative way to increase TIA production is through plant tissue and cell cultures, or even bacterial cultures. For plant tissue and cell cultures, lots of efforts have been done on optimizing the culture conditions and feeding all kinds of elicitors, and some simple metabolic engineering works have been done in hairy root culture, suspension cell culture, etc. Those manipulations have been reviewed in the introduction of the thesis. While for

heterologous production of those valuable TIAs, the studies are in their infancy and limited by the elucidation of the whole pathway and some deficiencies on heterologous expression, such as protein misfolding. As secondary metabolites, TIA pathway is highly regulated and large portions of the networks on metabolites, transcription factors, and genes in *C. roseus* are still unknown. To better engineer the production systems, we have to learn the pathway and the regulation mechanism more deeply.

The TIA biosynthetic pathway is highly branched (Figure 3.1). The first TIA in the pathway, strictosidine, contains an indole moiety from tryptamine and terpene moiety from secologanin. The condensation reaction is catalyzed by strictosidine synthase (STR), then strictosidine  $\beta$ -D-glucosidase (SGD) deglycosylates strictosidine to strictosidine aglycone. Further enzymatic steps result in the formation of numerous TIAs via several branches. In specific, one branch of the TIA pathway leads to the production of ajmalicine and serpentine, a second branch produces catharanthine, and a third branch leads to an intermediate, tabersonine, which is a key metabolite in the TIA pathway. The metabolites derived from tabersonine are differentiated in different organs. The pathway from tabersonine to vindoline activates in plant leaves, vindoline and catharanthus are condensed by a major class III peroxidase (PRX1), and this leads to the production of anticancer drugs, vinblastine and vincristine. While in root, tabersonine is derived into many other metabolites, such as lochnericine and horhammericine.

Transcriptional activators and repressors play importance roles in regulating TIA synthesis. To date, seven putative activators (ORCA2, ORCA3, CrBPF1, CrMYC1, CrMYC2, CrWRKY1 and CrWRKY2) and five putative repressors (ZCT1, ZCT2, ZCT3, GBF1 and GBF2) have been implicated as regulators of the TIA pathway. The studies on those transcriptional factors have been reviewed in the introduction of this thesis. ORCA2

(Octadecanoid-Responsive Catharanthus AP2-domain protein 2) is AP2-domain transcription factor that are proposed to activate *STR* expression by binding to the jasmonate and elicitor-responsive element (JERE) in the *STR* promoter [4]. CrBPF1 was identified as a protein that binds to a separate element in the *STR* promoter through a signal transduction pathway that is responsive to elicitors but not jasmonate and acts downstream of protein phosphorylation and calcium influx [5].

To date, there are only a few studies about ORCA2 and CrBPF1 characterization. Here we studied systematically on the transcriptional and metabolic effects lead by overexpression the two transcriptional activators separately. Our collaborated lab, Sue Gibson's group, focused on the generation of the transgenic *C. roseus* hairy root lines and did transcription analysis of several transcriptional factors and genes involved in the TIA pathway on those transgenic lines. While in our lab, we did metabolic profiling of the TIAs on the hairy root lines. In this chapter, we summarized all the works have been done on the metabolic profiling part.

## **Materials and Methods**

### **Alkaloid extraction**

Frozen hairy root tissue samples, collected as described above, were lyophilized and ground to a powder. Approximately 50 mg of the freeze-dried and ground hairy roots were added to a 50-mL centrifuge tube and extracted with 10 mL of methanol using a sonicating probe (Model VC 130PB, Sonics & Materials, Inc.) for 10 min while held on ice. The extracts were then centrifuged at 4000 rpm for 12 min at 15 °C. The supernatant was removed and the biomass was extracted one more time in the same manner. The combined supernatants were passed through a 0.45 µm nylon filter (25 mm, PJ Cobert), dried using a nitrogen evaporator

(Organomation Associates, Inc.), reconstituted in 2 mL of methanol, and passed through a 0.22  $\mu\text{m}$  nylon filter (13 mm, PJ Cobert) for HPLC analysis. Extracts were stored at  $-25\text{ }^{\circ}\text{C}$ .

### **Alkaloid analysis**

20  $\mu\text{L}$  of the alkaloid extract was injected into a Phenomenex Luna C18(2) column (250 x 4.6 mm) using three solvent systems. The Waters high performance liquid chromatography system consisted of a 1525 binary pump, a 717plus Autosampler, and a 996 Photo Diode Array (PDA) detector. For the measurement of tryptophan and tryptamine, a previously described method was used with UV detection at 218 nm [6]. Quantification was performed by comparison to standard curves. For detection of iridoid glycosides a second solvent system was used. Data extracted at 239 nm were used to quantify loganin and secologanin. For 30 min, at a flow rate of  $1\text{ mL min}^{-1}$ , the mobile phase was linearly ramped from a 90:10 to a 75:25 mixture of 1% formic acid (v/v)/0.25% trichloroacetic acid (w/v):acetonitrile. The ratio was then returned to 90:10 and the column was allowed to re-equilibrate.

For the detection of TIAs, a third solvent system was used. Data extracted at 254 nm were used to quantify strictosidine, ajmalicine, serpentine, catharanthine, vindoline, vinblastine and vincristine using standard curves. Data extracted at 329 nm were used to quantify tabersonine, hörhammericine, lochnericine, 19-hydroxytabersonine (19OHTab) and 16-hydroxytabersonine (16OHTab) using retention time standards, photodiode array detection, and a comparison to a tabersonine standard curve. The retention time standards of 19OHTab and 16OHTab were in-house prepared and the procedures were described in Chapter II. Metabolite levels were determined based on peak area, with the exception of loganin, which was quantified based on peak height. The protocol was adapted from a previously published protocol [6]. For the first 5 min, at a flow rate of  $1\text{ mL min}^{-1}$ , the mobile phase was a 30:70 mixture of acetonitrile:

100 mM ammonium acetate (pH 7.3). The mobile phase was linearly ramped to 64:36 over the next 10 min and maintained at that ratio for 15 min. The flow rate was increased to  $1.4 \text{ mL min}^{-1}$  over 5 min. The ratio was then increased to 80:20 during the next 5 min and maintained for 15 min. The flow rate and mobile phase ratio were then returned to  $1 \text{ mL min}^{-1}$  and 30:70 and the column was allowed to re-equilibrate.

For LC-MS analysis, the Agilent Technologies 1100 series liquid chromatography (LC) system with a binary pump, a temperature-controlled autosampler and a diode-array detector mated with an Agilent MSD model SL ion trap mass spectrometer (MS).  $10 \mu\text{L}$  of the alkaloid extract was injected into the LC-MS system with a Phenomenex Luna C18(2) 100A column ( $150 \times 2.0 \text{ mm}$ ). The LC-MS protocol is adapted from the third solvent system of HPLC methods for a smaller diameter of column. The mass spectrometer was operated in positive ion mode and full scan data was collected between 50 and 700  $m/z$ . Operation and analyses were performed with ChemStation software (Agilent Technologies). The LC-MS analysis was performed at the W. M. Keck Metabolomics Research Laboratory.

## Results and Discussions

### Effects of ORCA2 overexpression on the metabolite levels of the TIA pathway

ORCA2-OE line was generated expressing ORCA2 with an ethanol-inducible promoter. After sequencing confirmation and adaption, RT-PCR results of the engineered lines (induced and uninduced) and the control lines (induced and uninduced) further confirmed that *ORCA2* transcript level had been brought up in the induced ORCA2-OE line (Figure 3.2). *ORCA2* transcript levels increase dramatically after adding ethanol to ORCA2-OE root cultures, with *ORCA2* transcript levels increasing approximately eight fold within 24 h. In comparison, *ORCA2*

transcript levels in the uninduced ORCA2-OE line and the ethanol-induced or uninduced control line are relatively low.

To analyze the effects of *ORCA2* overexpression, the levels of 17 TIA and related metabolites were investigated, with 13 of those metabolites found to be present at detectable levels in at least some of the samples analyzed. The time course analysis of those metabolites was tracked over a 72 hour period following 0.02% ethanol induction, and comparisons were made among the four lines as the ones in *ORCA2* transcript level detection.

In one of the upstream pathways for TIA synthesis, the indole pathway, tryptophan and tryptamine were analyzed. Tryptamine levels increase significantly within 24 h of induction and remain increased through 72 h. Tryptamine levels remain unchanged in the uninduced ORCA2-OE cultures and the induced and uninduced cultures of the control line (Figure 3.3). Levels of tryptophan were below detection limits. In the other upstream pathway, terpenoid pathway, the two metabolites, loganin and secologanin, which are very close to the formation of the first TIA, strictosidine, were quantified. Loganin levels in induced and uninduced ORCA2-OE cultures are similar (Figure 3.3). In contrast, secologanin levels are somewhat variable in the ORCA2-OE line. Secologanin levels are significantly decreased in induced ORCA2-OE cultures relative to uninduced ORCA2-OE cultures at 24 and 72 h after induction, but are not significantly different at the other time points assayed (Figure 3.3).

For the downstream TIA pathway, the levels of 12 TIAs from different branches were investigated in aliquots of the same tissue samples used to analyze transcript levels (Figure 3.4). Nine of these TIAs were present at detectable levels, whereas the levels of the other three TIAs (vindoline, vinblastine and vincristine) were below detection thresholds. In addition, a tabersonine-like compound, designated Unk54, was detected in the samples. The levels of

strictosidine, the first TIA to be formed from secologanin and tryptamine, are significantly reduced in the induced versus uninduced *ORCA2*-OE cultures within 24 h after the start of induction and remain reduced throughout the remainder of the 72-h time course. Ajmalicine and serpentine are formed via the same branch of the TIA pathway (Figure 3.1). Both ajmalicine and serpentine levels exhibit modestly significant increases in the induced versus uninduced *ORCA2*-OE cultures 48 h after the start of induction, but are not significantly different at the other time points assayed. Catharanthine, one of the precursors for vinblastine production, also exhibits modestly significantly higher levels 48 h after induction, but not at the other time points assayed. The levels of tabersonine, which is the starting material for multiple branches of the TIA pathway, are significantly reduced in induced versus uninduced *ORCA*-OE cultures within 6 h of the start of *ORCA2* induction. Tabersonine is converted via one branch of the TIA pathway to lochnericine and hörhammericine. Lochnericine levels are not significantly altered by *ORCA2* overexpression, but hörhammericine levels are significantly reduced in the induced versus uninduced *ORCA2*-OE cultures at 24 and 72 h after the start of induction. Tabersonine is converted via a different branch of the TIA pathway to 16-hydroxytabersonine (16OHTab). 16OHTab levels exhibit modestly significant increases in the induced versus uninduced *ORCA2*-OE cultures at 48 and 72 h after the start of *ORCA2* induction. Tabersonine is also converted to 19-hydroxytabersonine (19OHTab). 19OHTab levels increase within 12 h of the start of *ORCA2* induction and remain elevated for at least 72 h. Levels of vindoline, the end metabolite of this branch of the TIA pathway and precursor to vinblastine production, are below detection limits. An un-identified metabolite, designated Unk54, was also detected in the tissue extracts. Based on its UV absorption spectra between 200 and 400 nm (data not shown), Unk54 is a tabersonine-like compound. Unk54 levels exhibit modestly significant increases in the induced versus uninduced

ORCA2-OE cultures at 48 h after the start of induction, but not at the other time points assayed. The levels of the pharmaceutically important TIAs, vinblastine and vincristine, were below detection limits.

### **Effects of CrBPF1 overexpression on the metabolite levels of the TIA pathway**

CrBPF1-OE line was generated expressing CrBPF1 with a  $\beta$ -estradiol-inducible promoter. After sequencing confirmation and adaption, RT-PCR results of the CrBPF1 endogenous gene and the CrBPF1 transgene from four lines (CrBPF1-OE induced line, CrBPF1-OE uninduced line, control induced line, control uninduced line) confirmed that *CrBPF1* transcript level had been brought up in the induced CrBPF1-OE line. In specific, *CrBPF1* transgene mRNA levels increased rapidly in the CrBPF1-OE line after addition of 20  $\mu$ M  $\beta$ -estradiol, rising approximately 50 fold within 6 h, and remained high for at least 72 h. Uninduced cultures of the CrBPF1-OE line exhibited much lower transcript levels for the *CrBPF1* transgene than the induced cultures, indicating that  $\beta$ -estradiol is necessary for high-level expression of the *CrBPF1* transgene (Figure 3.5A). As expected, qRT-PCR reactions using a primer pair specific for the *CrBPF1* transgene produced only a very low signal from RNA isolated from the control line (data not shown), which lacks the *CrBPF1* transgene. Transcript levels for the *CrBPF1* endogenous gene were not significantly affected by treatment with  $\beta$ -estradiol and were similar in the CrBPF1-OE and control lines (Figure 3.5B).

For the indole pathway, an attempt was also made to analyze tryptophan and tryptamine levels in aliquots of the same tissue samples used to analyze gene expression. However, tryptophan and tryptamine levels were below the detection threshold in many of the samples. For the terpenoid pathway, the levels of loganin and secologanin were also determined in aliquots of the same tissue samples used for gene expression analyses. The addition of 20- $\mu$ M  $\beta$ -estradiol to

the media caused no substantial alterations in the levels of either of these metabolites over the time period analyzed (Figures 3.6).

Towards that end, the levels of ten TIA metabolites were analyzed over a 72-h period in the CrBPF1-OE line grown in the presence or absence of 20- $\mu$ M  $\beta$ -estradiol, with seven of those metabolites being present at detectable levels (Figure 3.7). The metabolites analyzed were tabersonine, lochnericine, hörhammericine, catharanthine, serpentine, ajmalicine, strictosidine, vindoline, vincristine and vinblastine, with the levels of the last three being below the detection threshold. Overexpression of *CrBPF1* had only modest effects on the levels of the other seven metabolites, with the largest statistically significant effect being ~40% lower serpentine levels in the induced versus un-induced CrBPF1-OE cultures at the 12-h time point. The levels of the same metabolites were also analyzed in the control line, at 0 and 24 h after transfer to fresh media with 0 or 20  $\mu$ M  $\beta$ -estradiol. Addition of 20- $\mu$ M  $\beta$ -estradiol to the media had little effect on the levels of any of the TIA metabolites analyzed in the control line (data not shown).

## Figures

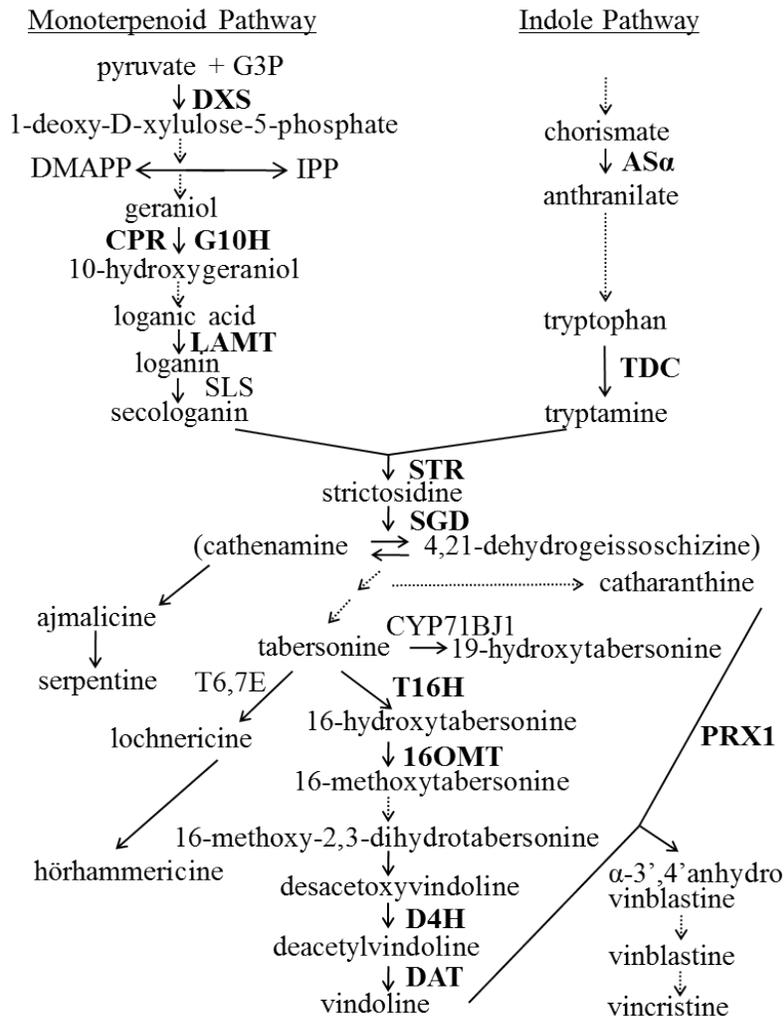


Figure 3.1 TIA biosynthesis in *C. roseus*. Enzyme abbreviations are written in capital letters next to the arrow indicating the reaction catalyzed by each enzyme. The metabolites resulting from different enzymatic conversions are indicated by the appropriate arrows. Solid arrows represent single enzymatic conversions, whereas dashed arrows indicate multiple enzymatic reactions. The genes that were included for analysis of expression are marked in bold. 16OMT, 16-hydroxytabersonine-*O*-methyl-transferase; AS, anthranilate synthase; CPR, cytochrome P450 reductase; CYP71BJ1, CYP71 cytochrome P450 hydroxylase; D4H, desacetoxyvindoline 4-hydroxylase; DAT, deacetylvindoline 4-*O*-acetyltransferase; DMAPP, dimethylallyl pyrophosphate; DXS, 1-deoxy-D-xylulose-5-phosphate synthase; G3P, glyceraldehyde 3-phosphate; G10H, geraniol 10-hydroxylase; IPP, isopentenyl diphosphate; LAMT, loganic acid *O*-methyltransferase; PRX1, vacuolar class III peroxidase; SGD, strictosidine glucosidase; SLS, secologanin synthase; STR, strictosidine synthase; T6,7E, tabersonine 6,7-epoxidase; T16H, tabersonine 16-hydroxylase; TDC, tryptophan decarboxylase.

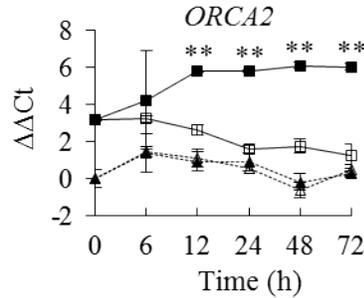


Figure 3.2 Time course analysis of *ORCA2* expression. Results shown are the: uninduced *ORCA2*-OE line (□, solid line), 0.02% ethanol-induced *ORCA2*-OE line (■, solid line), uninduced control line (Δ, dashed line) and 0.02% ethanol-induced control line (▲, dashed line). The relative mRNA levels are presented as  $\Delta\Delta CT$ .  $\Delta\Delta CT = \Delta CT_{\text{uninduced control roots at 0 h}} - \Delta CT_{\text{other}}$ , where  $\Delta CT = CT_{ORCA2} - CT_{EF1/UBQ11}$ . Positive  $\Delta\Delta CT$  values indicate that *ORCA2* transcript levels are higher in the indicated hairy root line grown for the indicated time under the indicated conditions than in the uninduced control line at 0 h. Conversely, negative  $\Delta\Delta CT$  values indicate that *ORCA2* transcript levels are lower in the indicated hairy root line grown for the indicated time under the indicated conditions than in the uninduced control line at 0 h. Results are the average  $\Delta\Delta CT$  value of three biological replicates, with two technical replicates for each biological replicate. Error bars indicate standard deviations. \*\* indicates that *ORCA2* transcript levels in the induced versus uninduced *ORCA2*-OE cultures differed at the same time point by  $p \leq 0.01$ , according to Student's t-test. Student's t-test results for the induced versus uninduced cultures of the control line are not depicted.

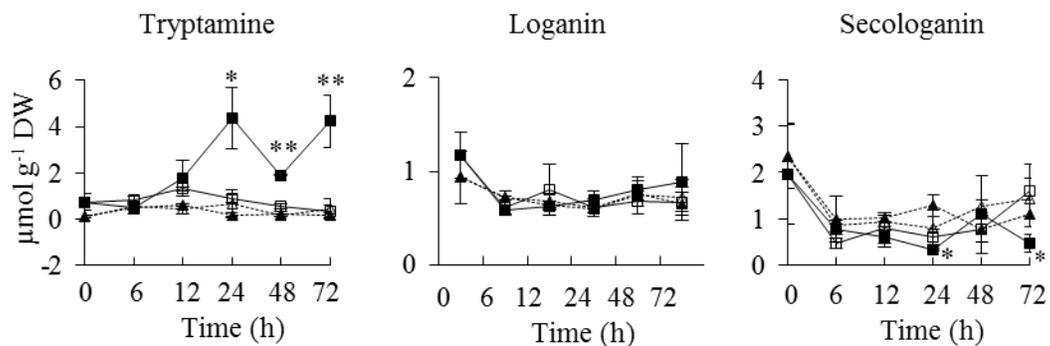


Figure 3.3 Time course analysis of metabolite levels from the indole and terpenoid pathways. Results shown are the: uninduced ORCA2-OE line (□, solid line), 0.02% ethanol-induced ORCA2-OE line (■, solid line), uninduced control line (Δ, dashed line) and 0.02% ethanol-induced control line (▲, dashed line). Results for metabolite levels are the averages of three biological replicates. Error bars indicate standard deviations. Metabolite levels differed with: \* =  $p \leq 0.1$ , \*\* =  $p \leq 0.05$  according to Student's t-test. Student's t-test results for the induced versus uninduced cultures of the control line are not depicted.

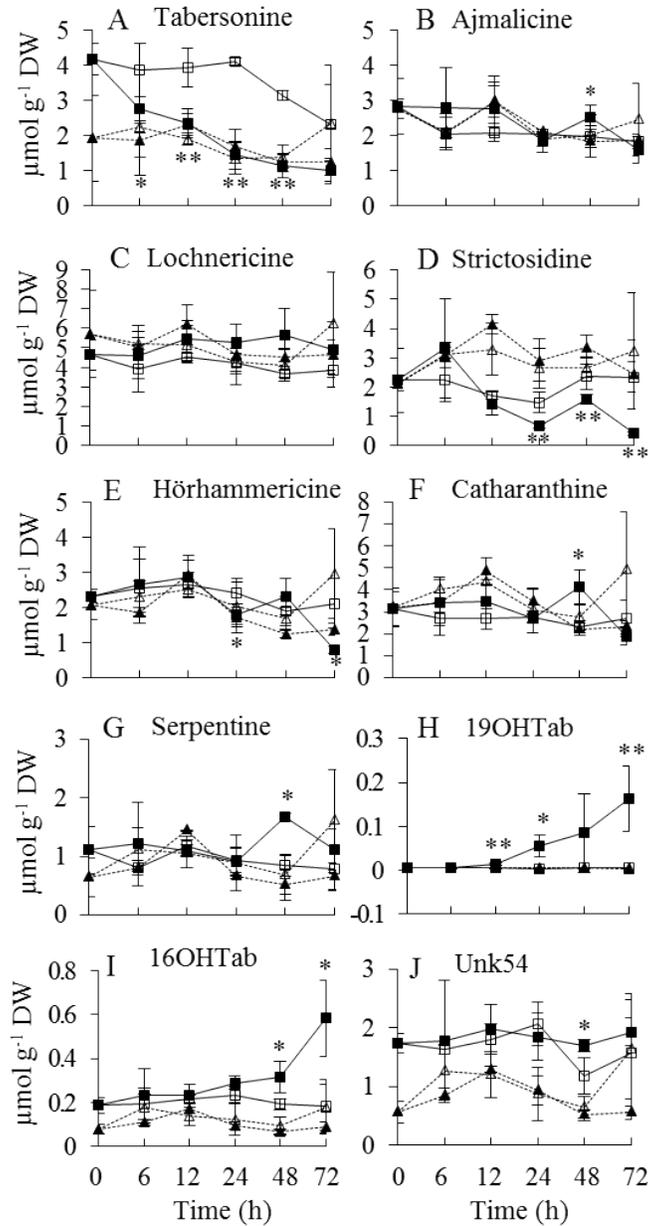


Figure 3.4 Time course analysis of TIA metabolite levels. Levels of vindoline, vinblastine and vincristine were below detection thresholds. Unk54 represents a tabersonine-like compound. Results shown are the: uninduced ORCA2-OE line (□, solid line), 0.02% ethanol-induced ORCA2-OE line (■, solid line), uninduced control line (Δ, dashed line) and 0.02% ethanol-induced control line (▲, dashed line). Results are the average values of three biological replicates. Error bars indicate standard deviations. Levels of the indicated metabolite in the induced versus uninduced ORCA2-OE cultures differed at the same time point with: \* =  $p \leq 0.1$ , \*\* =  $p \leq 0.05$  according to Student's t-test. Student's t-test results for the induced versus uninduced cultures of the control line are not depicted.

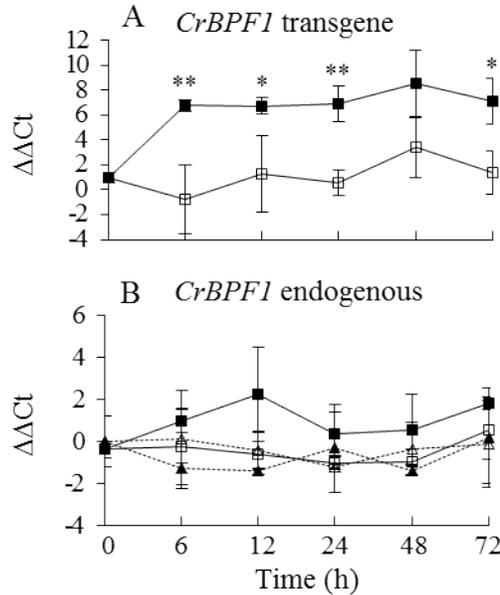


Figure 3.5 Time course analysis of *CrBPF1* endogenous and transgene mRNA levels. *CrBPF1* transcripts produced from the endogenous *CrBPF1* gene and the *CrBPF1* transgene were quantified independently by qRT-PCR using primers specific for each gene (Supplemental Table 4.1). As expected, the primer pair specific for transcripts produced by the *CrBPF1* transgene yields only a very low background signal for the control line (data not shown), which lacks the *CrBPF1* transgene. Results depicted are the following: un-induced *CrBPF1*-OE line (□, solid line),  $\beta$ -estradiol induced *CrBPF1*-OE line (■, solid line), un-induced control line ( $\Delta$ , dashed line) and  $\beta$ -estradiol induced control line ( $\blacktriangle$ , dashed line). **(A)** Relative *CrBPF1* transgene mRNA levels are indicated as  $\Delta\Delta\text{Ct}$ . Note that *CrBPF1* transgene mRNA levels were normalized versus *CrBPF1* endogenous gene mRNA levels in the un-induced control line at 0 h rather than against *CrBPF1* transgene mRNA levels in the un-induced control line at 0 h as the control line lacks the *CrBPF1* transgene. **(B)** Relative *CrBPF1* endogenous gene mRNA levels are indicated as  $\Delta\Delta\text{Ct}$ . A positive  $\Delta\Delta\text{Ct}$  value indicates that *CrBPF1* endogenous gene mRNA levels are higher in the indicated hairy root line grown for the indicated time under the indicated conditions than in the un-induced control line at 0 h. Negative  $\Delta\Delta\text{Ct}$  values indicate the reverse situation. Results are the average  $\Delta\Delta\text{Ct}$  of three biological replicates, with two technical replicates per biological replicate. Error bars indicate standard deviations. *CrBPF1* mRNA levels in the un-induced versus induced *CrBPF1*-OE cultures differed at the same time point with: \* =  $p \leq 0.05$ , \*\* =  $p \leq 0.01$  according to a Student's t-test. The results of Student's t-tests for the un-induced versus induced cultures of the control line are not depicted.

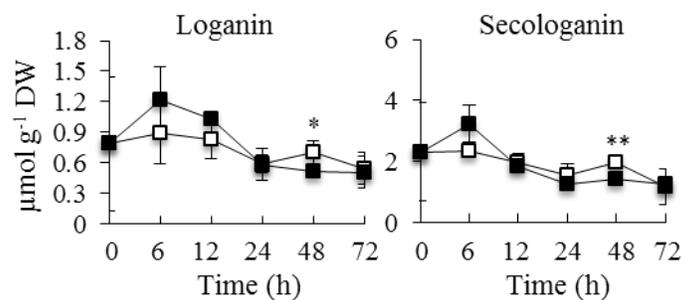


Figure 3.6 Time course analysis of metabolite levels from the terpenoid pathway. Results depicted are the following: un-induced CrBPF1-OE line (□, solid line), β-estradiol induced CrBPF1-OE line (■, solid line). Metabolite levels in the un-induced versus induced CrBPF1-OE cultures differed at the same time point with: \* =  $p \leq 0.1$ , \*\* =  $p \leq 0.05$  according to a Student's t-test.

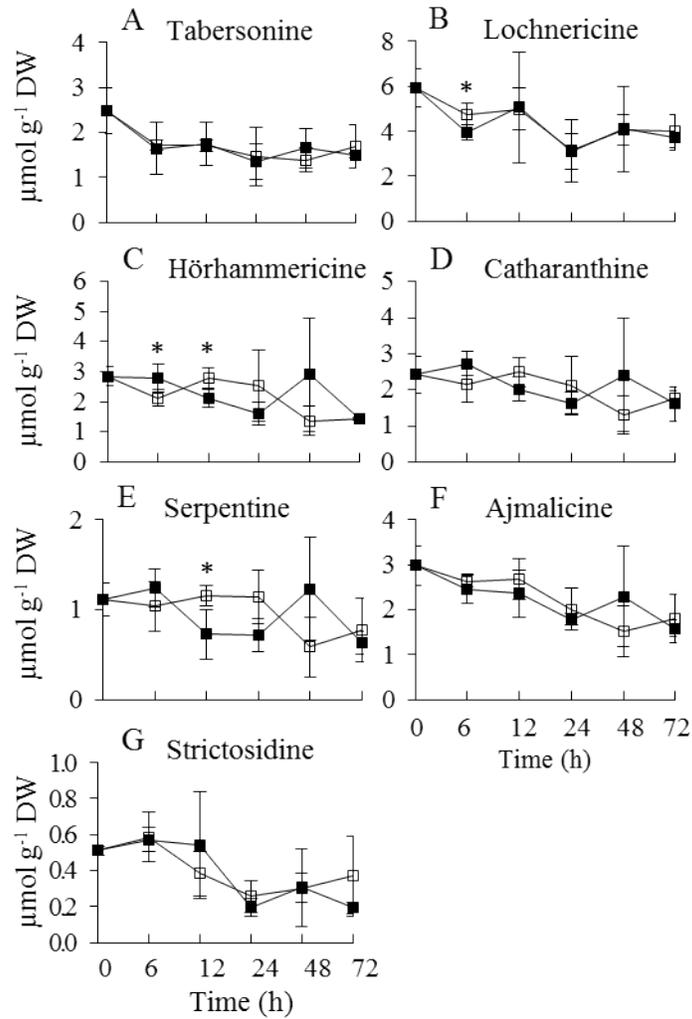


Figure 3.7 Time course analysis of TIA metabolite levels. Results depicted are the following: un-induced CrBPF1-OE line ( $\square$ , solid line) and  $\beta$ -estradiol induced CrBPF1-OE line ( $\blacksquare$ , solid line). Metabolite levels shown are the following: (A) tabersonine levels, (B) lochnericine levels (C) hörhammericine levels, (D) catharanthine levels, (E) serpentine levels, (F) ajmalicine levels and (G) strictosidine levels. Results are the average values of three biological replicates. Error bars indicate standard deviations. Metabolite levels in the un-induced versus induced CrBPF1-OE cultures differed at the same time point with:  $* = p \leq 0.1$  according to a Student's t-test.

## References

1. Gidding, C.E.M., et al., *Vincristine revisited*. *Critical Reviews in Oncology Hematology*, 1999. **29**(3): p. 267-287.
2. Ishikawa, H., et al., *Total synthesis of vinblastine, vincristine, related natural products, and key structural analogues*. *J Am Chem Soc*, 2009. **131**(13): p. 4904-16.
3. Sears, J.E. and D.L. Boger, *Total Synthesis of Vinblastine, Related Natural Products, and Key Analogues and Development of Inspired Methodology Suitable for the Systematic Study of Their Structure-Function Properties*. *Accounts of Chemical Research*, 2015. **48**(3): p. 653-662.
4. Menke, F.L.H., et al., *A novel jasmonate- and elicitor-responsive element in the periwinkle secondary metabolite biosynthetic gene Str interacts with a jasmonate- and elicitor-inducible AP2-domain transcription factor, ORCA2*. *Embo Journal*, 1999. **18**(16): p. 4455-4463.
5. van der Fits, L., et al., *A Catharanthus roseus BPF-1 homologue interacts with an elicitor-responsive region of the secondary metabolite biosynthetic gene Str and is induced by elicitor via a JA-independent signal transduction pathway*. *Plant Molecular Biology*, 2000. **44**(5): p. 675-685.
6. Peebles, C.A.M., et al., *Transient effects of overexpressing anthranilate synthase alpha and beta subunits in Catharanthus roseus hairy roots*. *Biotechnology Progress*, 2005. **21**(5): p. 1572-1576.

CHAPTER IV. PRODUCTION OF WAX ESTERS IN YEAST *YARROWIA LIPOLYTICA*

This chapter is a manuscript in preparation for submission to ACS synthetic biology. I contributed in the majority of the work.

Le Zhao<sup>1</sup>, Fuyuan Jing<sup>2</sup>, James Yu<sup>3</sup>, Suzanne Sandmeyer<sup>3</sup>, Jacqueline Shanks<sup>1</sup>, Zengyi Shao<sup>1</sup>

<sup>1</sup> Dept. of Chemical and Biological Engineering, Iowa State University

<sup>2</sup> Dept. of Biochemistry, Biophysics & Molecular Biology, Iowa State University

<sup>3</sup> Dept. of Biological Chemistry, University of California, Irvine

### Introduction

Wax esters are an important commodity and have wide applications in skincare products, cosmetics, inking industry, high performance lubricants, coating etc. There are some commercial natural sources for wax production, such as honeycomb, wool, Brazilian palm, jojoba, and sperm whale. Among all the natural waxes, sperm whale oil and jojoba oil are outstanding in lubrication, since they contain a high percentage of monounsaturated wax esters, keeping them in liquid form at room temperature with high oxidation stability [1, 2]. However, whale hunting was banned in 1970s, making the sperm whale oil supply terminate. And the tight supply of jojoba oil also makes its application limited to high price/low volume additives in personal care products. For the synthetic wax esters, several patents reported to produce sperm whale oil-type or jojoba oil-type of wax esters by chemical synthesis or enzyme catalysis, using expensive plant oils as substrates [3-5].

Scientists found that some bacteria, such as *Acinetobacter sp.*, *Marinobacter sp.*, *Psychrobacter sp.*, and *Rhodococcus sp.*, naturally accumulate wax esters in certain conditions [6]. The studies of those bacteria are still in their infancy due to either the uncharacterized pathways or the lack of genetic tools. Besides working on the native wax-producing bacteria, the heterologous production of wax ester in both *E. coli* and *S. cerevisiae* were also studied [7, 8].

Currently, the reported best wax ester titer in microorganisms was 80 mg/L, and it was produced in *Acinetobacter sp.* ADP1 [9].

In this study, we chose *Yarrowia lipolytica* as the host to produce wax esters heterologously. *Y. lipolytica* is oleaginous yeast, naturally accumulating abundant of triacylglycerol (TAG). And plenty of genetic tools have been developed. By engineering *Y. lipolytica* strain, the titer of TAG reached 85 g/L [10]. The aim of this study was to explore a new potential host, *Y. lipolytica*, for wax ester production by taking advantage of its natural capability of lipid hyper-accumulation and the ease of genetic manipulation.

## **Materials and methods**

### **Strains, media, and cultivation methods**

*Y. lipolytica* Po1f strain was obtained from ATCC. The rest of the studied *Y. lipolytica* strains were genetically modified from Po1f. *Escherichia coli* strain NEB5 $\alpha$  was used for recombinant DNA manipulation. We used YPD or other synthetic auxotrophic yeast culture media to select or culture yeast cells, and used Luria broth with 100  $\mu$ g/mL ampicillin to select and culture recombinant *E. coli* strains. *Y. lipolytica* strains were grown in glass tubes at 30 °C and 250 rpm for aerobic growth. *E. coli* strains were grown at 37 °C and 250 rpm. All chemicals were purchased from ThermoFisher Scientific or Sigma Aldrich.

### **Gene deletion in *Yarrowia lipolytica***

To knock out the gene encoding phosphatidic acid phosphohydrolase (PAP) in TAG biosynthetic pathway, we used Cre-Lox recombination system. LoxP-Ura3-LoxP DNA fragment was amplified from plasmid p3992 obtained from Dr. Suzaane Sandmeyer's group in University of California, Irvine. Around 1 kb of promoter region and terminator region of the PAP gene

were amplified from *Y. lipolytica* Po1f genome as upstream and downstream homologous arms in the knockout fragment. Gibson assembly was applied to assemble the promoter DNA fragment, LoxP-Ura3-LoxP fragment, and the corresponding terminator fragment to construct the knockout fragment, namely Ppap-LoxP-Ura3-LoxP-Tpap (Figure 4.1). After the Ppap-LoxP-Ura3-LoxP-Tpap cassette was constructed and sequence was confirmed, it was transferred into *Y. lipolytica* Po1f as a linear DNA fragment via chemical transformation. Colony PCR was conducted to select colonies in which the LoxP cassette was integrated into the right position of the genome. Once we got the colony with the original PAP gene substituted by the LoxP cassette, we transferred p3902 (gift from Suzaane Sandmeyer) containing CreA and Leu2 selection marker into the colony to eliminate the Ura3 marker gene between two LoxP elements. After the chemical transformation, we spread the cells on Sc-Leu plate to maintain the CreA plasmid. Then we replicated colonies from the Sc-Leu plate to Sc-Leu+FOA (1 g/L) plate to select the colonies with the deletion of Ura3 marker. Colony PCR was conducted to check whether the selected colonies have lost their Ura3 markers between LoxP elements. Then the confirmed strain was grown in YPD medium for many generations to lose p3902 CreA-Leu2 plasmid. By spreading proper amount of cultured cells on YPD plate and making replicated Sc-Leu plate, we screened out the colonies which had successfully lost p3902 plasmid. The same procedures were conducted to generate MFE1 knockout.

Le3 (PEX10 knockout strain) was a gift from Yu Jiang's lab in China. They developed the CRISPR protocol in *Y. lipolytica* and used that technique to knock out PEX10 gene [11].

### **Construction and transformation of plasmid and linear DNA**

Plasmid pLZ1 was constructed with four basic elements: an Ura3 selection marker from *Y. lipolytica* genome, ARS18 with replicative function from *Y. lipolytica* genome, ampicillin

resistance gene (*bla*), as well as a replicative function sequence working in *E. coli*. pLZ2 plasmid is very similar to pLZ1 plasmid, the only difference is that pLZ2 contains a *Leu2* selection marker instead of a *Ura3* selection marker.

Genes encoding ScFAR from jojoba (*Simmondsia chinensis*), MmFAR from house mouse (*Mus musculus*), MhFAR from *Marinobacter hydrocarbonoclasticus*, and WS/DGAT from *Acinetobacter sp.* strain ADP1 were codon optimized with our in-house program and synthesized by Integrated DNA Technologies (IA, USA). These FARs and WS/DGAT expression cassettes which include a *TEF1* (including *TEF1* original intron) promoter and a *CYC1* terminator were inserted into pLZ1 or pLZ2 plasmids separately to form different expression plasmids listed in Table 4.1.

The constructed plasmids were transferred into *E. coli* NEB5 $\alpha$  for propagation and the sequence of each plasmid was confirmed by DNA sequencing. Then the correct plasmids were extracted from *E. coli* separately and transferred into *Y. lipolytica* strains via chemical transformation. The constructed strains were culture in synthetic complete selective media at 30 °C.

### **Random integration of genes into *Y. lipolytica***

Hygromycin phosphotransferase (HPT) gene and *ura3* gene were used as selection markers for gene random integration into *Y. lipolytica* genome. HPT gene from *E. coli* was codon optimized and synthesized together with a *Leu2* minimal promoter (-94bp ~ -1bp) and a short synthetic terminator *Tsynth8* [12]. The whole DNA sequence was confirmed after assembling to plasmid and being sequenced. Overlap extension PCR was used to assemble HPT and MhFAR, and the PCR product was chemically transferred into *Y. lipolytica* to allow the

HPT-MhFAR fragment to be randomly integrated into the genome. And the same procedure was applied to integrate WS-Ura3 cassette into chromosome.

### **Nile Red staining and lipid separation via thin layer chromatography (TLC)**

To roughly estimate lipid production, we used Nile Red to stain the lipid composition in cells. We took 20  $\mu$ l of cell culture and transferred it into Nunc F96 MicroWell black plate. Then we added 180  $\mu$ l staining buffer, 40  $\mu$ M Nile Red in 90% phosphate-buffered saline (PBS) and 10% dimethyl sulfoxide (DMSO), into the 20  $\mu$ l of sample, and put it in a Synergy HT microplate reader. The program for the microplate reader was as follows: excitation wavelength was 480nm, emission wavelength was 580nm, continuously shook the plate for 30 min, and recorded the kinetic data with 10 min interval.

For TLC separation, we took 400  $\mu$ l of cell culture and span down the cells, kept the cell pellet, and added 200  $\mu$ l of methanol, 100  $\mu$ l of hexane, and 200  $\mu$ l of water one after another with vortex in the interval. After well mixed, we centrifuged the sample and loaded 10  $\mu$ l of hexane layer (upper layer) to the TLC plate. C18:0FA/C18:0-ol was used as wax ester standard, and the TLC plate was developed with the solvent hexane/diethyl ether/acetic acid, 90/10/1, v/v/v. After running on TLC, we sprayed primuline solution (5 mg in 100 ml of acetone/water, 80/20, v/v) evenly on the TLC plate with a Sigma glass sprayer, and visualized the plate under UV illumination.

### **Total lipid extraction and analysis**

Two ml of liquid cell culture were harvested after growing for 72 hours at 30 °C, 250 rpm. Cells were pelleted with centrifugation and the supernatant was dumped. The cell pellet was resuspended with 700  $\mu$ l water, and then lyophilized. Cell dry weight was recorded. Proper

amount of internal standard, pentadecanoic acid (C15FA), was spiked into each sample. The following lipid extraction method was adapted from Folch method [13]. Sample was well-dispersed in 2 ml of chloroform and 1 ml of methanol with sonication and vortex. The homogenate was centrifuged and 1 ml of the liquid phase was transferred into a new glass tube and completely dried with nitrogen evaporator. Then we used 500  $\mu$ l of hexane and 500  $\mu$ l of water to partition the dried extract. After mixing and centrifugation, 200  $\mu$ l of the upper hexane phase containing lipids was taken out and dried.

Before running GC-MS, samples were silylated with MSTFA reagent (Thermo Scientific) followed by the recommended protocol. GC-MS analysis was performed in Agilent Technologies gas chromatograph coupled to Agilent mass spectrometer. The GC column we used was a DB-Wax capillary column (15 m  $\times$  250  $\mu$ m  $\times$  0.25  $\mu$ m). The injection volume was 1  $\mu$ L at a carrier gas flow of 3 mL/min helium with splitless mode. The GC program was as follows: At first, the temperature increased from 80  $^{\circ}$ C to 230  $^{\circ}$ C linearly with a rate of 15  $^{\circ}$ C per minute, then elevated to 340  $^{\circ}$ C with a rate of 8.5  $^{\circ}$ C per minute, and held at 340  $^{\circ}$ C for another 7 minutes. The mass spectrometer source temperature was tuned to 280  $^{\circ}$ C. MS data was processed using the Chemstation software.

From the GC-MS data, different types of free fatty alcohols and wax esters were quantified with corresponding standards purchased from Nu-Check Prep, Inc. (MN, USA).

### **Lipid saponification and analysis**

To quantify the total fatty acid, we needed to extract all the fatty acid group from all types of lipids, such as TAG, wax ester, sterol ester etc. That is called saponification. We took about 10 mg of freeze dried sample and added 1 ml 1 N KOH in 90% aqueous methanol, and reacted at 80  $^{\circ}$ C for 1 hour. After cooling to room temperature, we added 1 ml 2 N HCl to acidify

the solution, and then used 2 ml hexane to extract the fatty acids. We took the hexane layer sample, did silylation and ran GC-MS to determine the total fatty acid. The silylation protocol and GC-MS protocol were the same as described above.

### **Copy number analysis**

To measure the copy number of WS gene on genome for DI strain or on plasmid for SI strain, total DNA (including genomic DNA and plasmid DNA) was extracted from *Y. lipolytica* cells following our previous protocol [14] and was quantified by real-time PCR. ACT1 gene encoding actins was chosen as a reference gene. The purified PCR products amplifying partial ACT1 gene and WS gene were used as template to construct standard curve. A 16-fold serial dilution of those two PCR products, ranging from 1 ng/μL to  $9.5 \times 10^{-7}$  ng/μL (1/1049576 ng/μL), were prepared. Real-time PCR was performed on a StepOnePlus System (Life Technologies, Carlsbad, CA) using SYBR Green as the DNA dye. And the protocol was described previously [14].

## **Results**

### **Knocking out the substrate competitive pathways to increase acyl-CoA pool**

Acyl-CoA is a key metabolite in fatty acid metabolism and catabolism pathways (Figure 4.2). Naturally, the inputs of acyl-CoA come from the fatty acid elongation pathway, the degradation of sterol ester, and the degradation of triacylglycerol. On the other hand, the outputs of acyl-CoA include that it reacts with sterol to form sterol ester; binds to glycerol molecules in a series way to form triacylglycerol; converts to free fatty acids; or is further degraded through beta-oxidation pathway. The wax ester synthetic pathway is also derived from acyl-CoA. FAR converts acyl-CoA to fatty alcohol, and WS combines one molecule of acyl-CoA and one

molecule of fatty alcohol to form a wax ester (Figure 4.3). In other words, acyl-CoA is the sole substrate for wax ester production.

One of the direct strategies to improve wax ester production is to increase the substrate acyl-CoA pool by knocking out or knocking down the substrate competitive pathways. Therefore, we constructed several knockout strains based on *Y. lipolytica* Po1f (Table 4.1). To disrupt the TAG pathway, the gene encoding phosphatidic acid phosphohydrolase (PAP) was knocked out to form a new strain Le1. The reason we chose PAP to manipulate is that its substrate, phosphatidic acid (PA), is a major component of cell membranes. It has been reported that dysfunction of the enzymes involved in PA synthesis can result in severe defects of various cellular processes [15]. Based on Le1 strain, we further knocked out the gene encoding multifunctional beta-oxidation protein (MFE1) to construct a new strain Le2. And by knocking out the gene encoding peroxisome assembly protein 10 (PEX10) in Po1f, we got Le3 strain. Several literatures have reported that lipid accumulation increased by disrupting the  $\beta$ -oxidation pathway via knocking out either MFE1 gene or PEX10 gene [16, 17]. So we expected that the metabolic flux could flow to wax ester synthetic pathway by disrupting the TAG pathway and  $\beta$ -oxidation pathway.

Then we tested the fatty acid types in Po1f, Le1, Le2, and Le3 strains. Figure 4.4 showed that *Y. lipolytica* mainly produced fatty acid with 16 and 18 carbon chain lengths, both saturated and unsaturated ones. Though the fatty acid type distributions were slightly different among those four strains, they all contained a high percentage of unsaturated fatty acids. The unsaturated fatty acids (C16:1  $\Delta$ 9, C18:1  $\Delta$ 9, C18:2  $\Delta$ 9 $\Delta$ 12) took about 70% of the total fatty acids.

## Heterologous expression of the wax ester synthetic pathway in the four *Y. lipolytica* strains

There are two enzymes involved in the wax ester pathway, fatty acyl-CoA reductase (FAR) and wax synthase (WS) (Figure 4.3). The FAR and WS genes were co-expressed in the four *Y. lipolytica* strain, Po1f, Le1, Le2, and Le3, separately. Besides comparing the wax ester productions and types in different host strains, we also compared the effect of three different FARs on wax ester production since the substrate specificity of FARs from different sources varies. The three FARs we studied were ScFAR from jojoba plant (*Simmondsia chinensis*), MmFAR from house mouse (*Mus musculus*), and MhFAR from a marinobacter (*Marinobacter hydrocarbonoclasticus*). And the WS gene we used encodes the bifunctional protein wax ester synthase/acyl coenzyme A (acyl-CoA):diacylglycerol acyltransferase (WS/DGAT) from *Acinetobacter sp.* strain ADP1. This enzyme has been used in several researches to produce biodiesel-like fuels as fatty acid ethyl esters [18, 19]. All the combinations of the engineered strains are listed in Table 4.1.

Figure 4.5 summarized the titers of the total free fatty alcohols and total wax esters with three FARs in four host strains. Comparing the wax ester titers among the three FARs, ScFAR produced the least, and MhFAR accumulated the most. This was led by the compatibility of the original acyl-CoA types in *Y. lipolytica* and the substrate specificity of those FARs and WS/DGAT. Specifically, the main types of fatty acid produced by *Y. lipolytica* are C16 and C18 saturated and unsaturated fatty acids. In previous study, WS/DGAT has a very broad substrate specificity towards fatty alcohol (C2 – C30) and acyl-CoA (C2 – C20), and prefers to utilize C14, C16, C18, and C18:1 fatty alcohols and C14, C16, C16:1, and C18 acyl-CoA [20]. This indicates that WS/DGAT we chose has a good compatibility with the main acyl-CoA types in *Y. lipolytica*. While for FARs, ScFAR prefers to utilize C18, C20:1, and C22:1 acyl-CoAs [21],

MmFAR prefers C16, C18, C18:1, C18:2 acyl-CoAs [22], and MhFAR prefers C16, C18, C20, C18:1 acyl-CoAs [23]. Because there is rare very long chain acyl-CoA (>C18 acyl-CoA) in *Y. lipolytica*, we observed the tiny amount of wax ester accumulation for the strains with ScFAR expression. Although both MmFAR and MhFAR have good compatibility with the major acyl-CoA types in *Y. lipolytica*, the big difference in wax ester titers between those two types of strains may partially be contributed by different enzyme activities when being expressed in the host strains.

In addition, comparing the performance of same FAR in different host strains, we noticed that Po1f, Le1, and Le3 produced relatively equivalent amount of wax esters. In strains with ScFAR, the order of wax ester amounts were Le1 > Po1f > Le3; while in strains expressing either MmFAR or MhFAR, Po1f performed the best, followed by Le3, then Le1. But for Le2 strain, wax ester was never detected. For knocking out the substrate competitive pathways, we expected a larger acyl-CoA pool leading to an increase of wax ester production, but the results were opposite to what we expected. A TLC analysis for the lipid distributions in Mh-Po1f, Mh-Le1, Mh-Le2, and Mh-Le3 was then performed (Figure 4.6), and we found that those knockout strains had more free fatty acids than Po1f strain, indicating that the fatty acids were not in the acyl-CoA activated form in those knockout strains.

### **Distributions of free fatty alcohol types and wax ester types produced in engineered strains**

Besides quantifying the total titers of free fatty alcohol and wax ester in the engineered strains, we also analyzed the distributions of different types of fatty alcohols and wax esters as shown in Figure 4.7 and Figure 4.8. Figure 4.7a indicated that ScFAR specifically catalyzed C18:0 acyl-CoA to C18:0-ol. A recent *in vitro* study reported that ScFAR exhibited the highest activity to C18:0 acyl-CoA followed by C20:1 acyl-CoA and C22:1 acyl-CoA [21]. Since *Y.*

*lipolytica* doesn't contain too much very long chain acyl-CoA (>C18), that's why we only observed C18:0-ol in the ScFAR strains. Figure 4.7b and 4.7c showed that MmFAR and MhFAR expressions led to the production of C16:0-ol, C18:0-ol, and C18:1-ol, and both are in favor of C18:0-ol. The distributions are partially consistent with the ones reported in literatures. MmFAR preferred C16:0, C18:0, C18:1, and C18:2 acyl-CoAs in insect cells [24] and MhFAR also preferred medium chain acyl-CoAs as C16:0, C18:0, C20:0 and C18:1 in *in vitro* experiments [23].

For wax ester types in Figure 4.8, strains containing ScFAR mainly produced C18:0FA/C18:0-ol and didn't produce any unsaturated wax ester; strains with MmFAR preferred to produce C18:0FA/C16:0-ol; and strains with MhFAR produced C16:0FA/C18:0-ol the most. All of the data showed that WS/DGAT has broad substrate specificity to those long chain acyl-CoAs (C16 and C18) and long chain fatty alcohols (C16 and C18), which is consistent with the reported WS/DGAT substrate specificity [20]. Figure 4.8 also reflected that the titers of saturated wax esters was much higher than the titers of unsaturated wax esters, this is probably due to the WS/DGAT enzyme localization in *Y. lipolytica*. WS/DGAT from the prokaryote *Acinetobacter sp.* ADP1 most likely stayed in the cytosol of *Y. lipolytica* cells, whereas most of the unsaturated fatty acyl-CoAs stay in organelles such as endoplasmic reticulum. The access to those unsaturated fatty acyl-CoAs and fatty alcohols for WS/DGAT seems to be an important question to be answered in future studies.

### **Random genome integration to avoid plasmid instability issue**

In our experiments, we used ARS18 (1.3kb) in all of the constructed plasmids [25]. Although in theory the centromere in ARS18 keeps the plasmid in higher mitotic stability, we still found serious plasmid instability issue in all the experiments: (1) The wax ester production

among different colonies on the same selection petri dishes were very different with large standard deviation; (2) The recovered strains from the frozen stocks produced much less wax esters compared with previous results.

To solve the plasmid instability issue, at first, we randomly integrated MhFAR together with a selection marker gene encoding hygromycin phosphotransferase (HPT) into *Y. lipolytica* Po1f's genome and kept WS/DGAT in the plasmid form with Leu2 as the selection marker (pLZ2-WS), and selected the transformants in Sc-Leu+200  $\mu\text{g/mL}$  hygromycin B petri dishes. We named the strain as Single Integration (SI) strain (Table 4.1). We randomly picked 48 colonies and cultured in Sc-Leu+200  $\mu\text{g/mL}$  hygromycin B liquid medium for three days, then extracted total lipids and analyzed with thin layer chromatography (TLC). We found that, among all the 48 colonies, only 11 of them (marked with yellow color) had relatively bright and big wax ester spots (compared with C18:0FA/C18:0-ol wax ester standards) (Figure 4.9). We quantified the wax esters and free fatty alcohols in the 11 colonies via GC-MS, and their titers were shown in Figure 4.10a. The titers of wax esters varied among the 11 samples probably due to the different locus MhFAR randomly inserted in the chromosome. Colony #35 produced the highest amount of wax ester ( $\sim 1.16$  g/L), and was about 2 fold of the best plasmid based engineered strain Mh-Po1f (0.57 g/L). However, the over 1 g/L titer could not be repeated any more when we conducted the nitrogen limited experiment for #35 SI strain (Figure 4.10b). Free fatty alcohol accumulated at a constant level ( $\sim 80$  mg/L), but the wax ester production dropped to 276 mg/L. It indicated that the plasmid pLZ2-WS was still not stable in the SI strain, so we decided to integrate the WS gene into genome, following the MhFAR random integration.

Firstly, we lost the pLZ2-WS plasmid in the #35 colony, only leaving the MhFAR in the chromosome, and then randomly integrated WS together with URA3 selection marker into the

genome, selected the transformants in Sc-Ura petri dishes. We named this strain as Double Integration (DI) strain (Table 4.1). Twenty colonies were randomly picked and their total lipids were analyzed via TLC. Different from the SI strain, almost all of the 20 colonies produced wax esters (Figure 4.11), then we picked 6 from the 20 colonies, cultured them in Sc-Ura liquid medium and analyzed the lipid production through GC-MS. The highest wax ester titer from the 6 colonies of DI strain was 84 mg/L (Figure 4.12a). Compared with the #35 SI strain, the production of wax ester was less in those DI strains (Figure 4.12b). Meanwhile, the copy number of WS gene was measured in both SI strain and DI strain at three time points (24, 48, 72 hours) (Figure 4.13). We found that the copy number of WS gene randomly integrated into the genome was one in the DI strain, while the copy number dropped from 2.1 (24 hours) to 1.5 (72 hours) in the SI strain. These results indicated that the copy number of WS gene in SI strain was greater than the one in DI strain, which could explain that SI strain produced more wax esters than DI strain did. And the decrease of the WS copy number in SI strain implied the plasmid instability issue.

### **Nitrogen limited fermentation to optimize culture medium for wax ester production**

Research on the microbial lipid production has shown that a high carbon to nitrogen ratio leads to the increased production of lipids [26]. Metabolic flux analysis and lipidomics were also studied in *Y. lipolytica* under carbon limited and nitrogen limited conditions to prove lipid accumulation in nitrogen limited condition [27]. In this project, we studied the effect of nitrogen limited fermentation on wax ester production. The common synthetic complete medium contains 20 g/L glucose and 5 g/L ammonia sulfate. We chose three concentrations of glucose, equal or higher than normal condition (20 g/L, 40 g/L, and 80 g/L) and three concentrations of ammonia sulfate (AS), equal or lower than normal condition (0.2 g/L, 1 g/L, and 5 g/L), keeping the rest of

medium composition same. We found that 40 g/L glucose led to the highest wax ester content when AS concentration varied, and 1 g/L AS also worked the best no matter how glucose concentration changed. In combination, the best medium composition for wax ester production contained 40 g/L glucose and 1 g/L AS (Figure 4.14).

### Discussions

In this study, we investigated three FARs from different species, ScFAR from jojoba, MmFAR from house mouse, and MhFAR from *Marinobacter hydrocarbonoclasticus*. It has been found that MmFAR is a transmembrane protein localized in the peroxisome [24]. And partial evidence has shown that ScFAR is also a membrane protein associated with the endoplasmic reticulum (ER) [28]. Since the sources of MhFAR and WS are both prokaryote, MhFAR and WS stay in the cytosol when being introduced in *Y. lipolytica*. We think that the different compartments FAR and WS enzymes located in would affect the metabolite transport across those organelle membranes, the distribution of free fatty alcohol among the organelles and cytosol would have effect on wax ester production.

On the TLC plate for analyzing lipid distributions in the 48 SI colonies, the patterns of lipid distributions present several situations: (1) Colonies like #5 and #9, rarely produced any free fatty alcohols nor wax esters. That is probably either because the integrated MhFAR was inactive or because only the selective marker gene encoding HPT was completely integrated into the genome. (2) Colonies like #1 and #4, accumulate significantly amount of free fatty alcohols, but much less wax esters compared with wax standards. This is most likely caused by the instability of the pLZ2-WS plasmid. (3) Some colonies accumulated less free fatty alcohols and less wax esters compared with other samples. For example, colony #14 produced less than colony #20, we think this may be led by the different genome locus of the integrated MhFAR.

When we constructed those substrate competitive pathway knockout strains, Le1, Le2, and Le3, we expected that the knockout strains produce more wax esters than the wild type strain Po1f, however the results contradicted to our expectation. TLC data indicated that there were more free fatty acid accumulated in Le1, Le2, and Le3 (Figure 4.6). Acyl-CoA serves as the precursors for free fatty acids, TAG, sterol esters, and many other types of lipids (Figure 4.2). The TLC result indicated that the pathways derived from acyl-CoA are regulated and interact with each other, providing a buffer capability for acyl-CoA pool. In our case, knocking out either the TAG pathway or  $\beta$ -oxidation pathway led to the accumulation of free fatty acids, and the acyl-CoA pool was not increased. To further test this hypothesis, the transcription levels of the genes in all the acyl-CoA derived pathways could be investigated. On the other hand, TLC is a rough analytical method for studying lipid distribution. A more comprehensive lipid quantification method using HPLC-CAD (Charged Aerosol Detection) reported by several literatures [29, 30] and companies (DIONEX, THERMOFISHER) may provide us more clues how the fluxes in the lipid pathway changed in those knockout strains.

### **Conclusion**

Currently, the highest wax ester titer in this study was 570 mg/L produced by Mh-Po1f. It was the highest wax ester production in microbes reported so far and was about 7 fold higher than the wax esters produced in *Acinetobacter* ADP1 (80 mg/L) [31]. And we found the different FARs had a great effect on the wax ester production, wax ester content and titer also boosted under certain nitrogen limited condition.

**Table**

Table 4.1 Strains used and constructed in this study.

Host Strain Name	Genotype
<b><i>Yarrowia lipolytica</i> host strains</b>	
Po1f	MatA, leu2-270, ura3-302, xpr2-322, axp-2
Le1	Derived from Po1f, ΔPAP
Le2	Derived from Po1f, ΔPAP, ΔMFE1
Le3	Derived from Po1f, ΔPEX10
<b><i>Yarrowia lipolytica</i> engineered strains</b>	
<b>Po1f background</b>	
Sc-Po1f	Po1f, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-ScFAR, pLZ2-WS
Mm-Po1f	Po1f, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-MmFAR, pLZ2-WS
Mh-Po1f	Po1f, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-MhFAR, pLZ2-WS
SI	Po1f, leucine <sup>+</sup> , MhFAR-HPT (integration), pLZ2-WS
DI	Po1f, uracil <sup>+</sup> , MhFAR-HPT (integration), WS-Ura3 (integration)
<b>Le1 background</b>	
Sc-Le1	Le1, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-ScFAR, pLZ2-WS
Mm-Le1	Le1, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-MmFAR, pLZ2-WS
Mh-Le1	Le1, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-MhFAR, pLZ2-WS
<b>Le2 background</b>	
Sc-Le2	Le2, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-ScFAR, pLZ2-WS
Mm-Le2	Le2, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-MmFAR, pLZ2-WS
Mh-Le2	Le2, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-MhFAR, pLZ2-WS
<b>Le3 background</b>	
Sc-Le3	Le3, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-ScFAR, pLZ2-WS
Mm-Le3	Le3, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-MmFAR, pLZ2-WS
Mh-Le3	Le3, leucine <sup>+</sup> , uracil <sup>+</sup> , pLZ1-MhFAR, pLZ2-WS

## Figures



Figure 4.1 Phosphatidic acid phosphohydrolase (PAP) loxP cassette. Ppap, promoter of PAP; Ura3, the selection marker; Tpap, terminator of PAP.

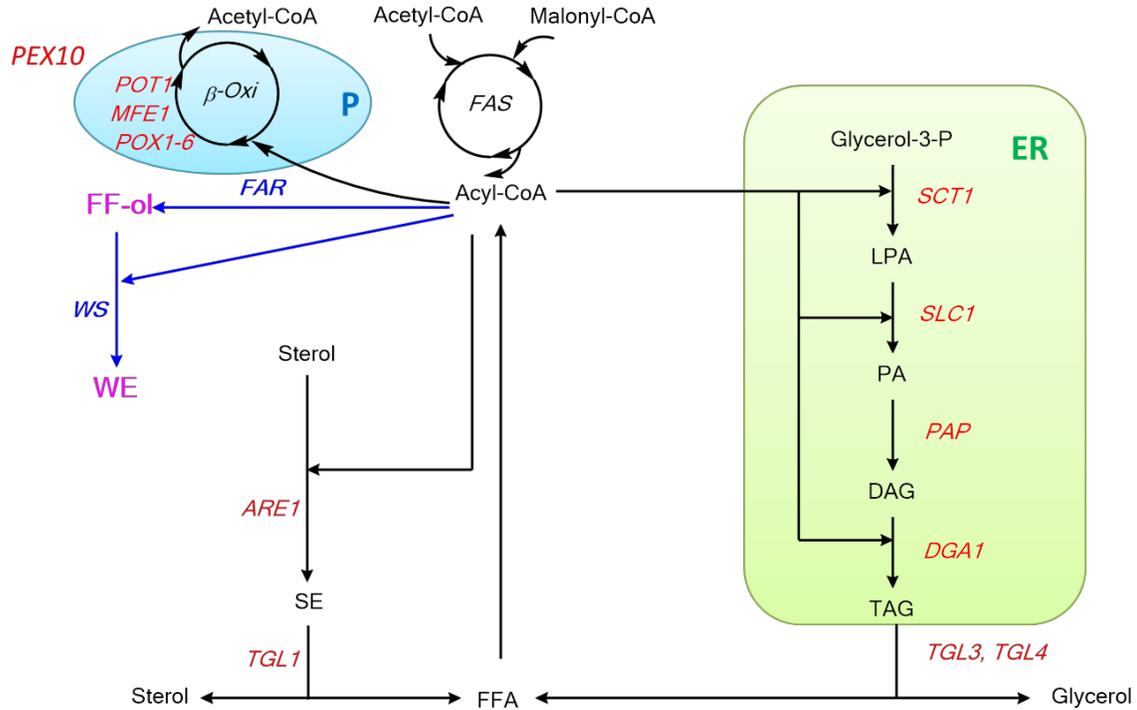


Figure 4.2 The overview of the lipid metabolism and catabolism pathways. FAS, fatty acid synthesis; Glycerol-3-P, glycerol-3-phosphate; LPA, lysophosphatidic acid; PA, phosphatidic acid; DAG, diacylglycerol; TAG, triacylglycerol; FFA, free fatty acid; SE, sterol ester; FF-ol, free fatty alcohol; WE, wax ester;  $\beta$ -Oxi, beta-oxidation pathway; Genes encoding each reactions in the pathways are in red color; P, peroxisome; ER, endoplasmic reticulum. Acyl-CoA is produced from fatty acid synthesis pathway, it is involved in the TAG synthesis and SE synthesis in ER, and can also be converted to FFA. In addition, TAG and SE can be degraded to FFA, and recycle to acyl-CoA. Acyl-CoA or FFA is transferred into peroxisome, and is further degraded via  $\beta$ -oxidation pathway. The introduced wax ester pathway also utilizes acyl-CoA as sole precursor.



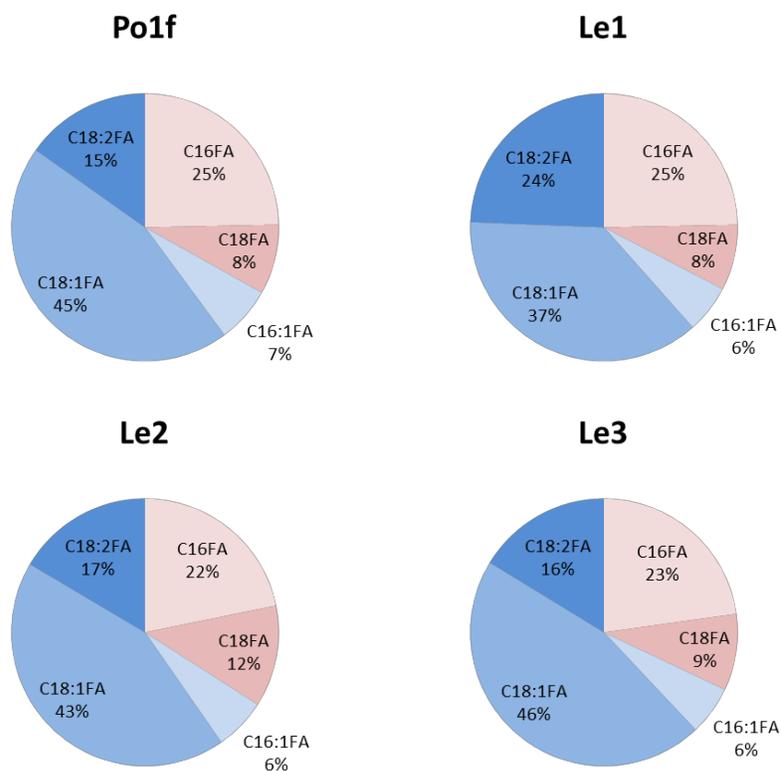


Figure 4.4 Fatty acid type distributions in Po1f, Le1, Le2, and Le3 strains. *Y. lipolytica* mainly produces C16 and C18 saturated and unsaturated fatty acids. The unsaturated fatty acids (C16:1, C18:1, and C18:2) take 70% (in weight) of the total fatty acids.

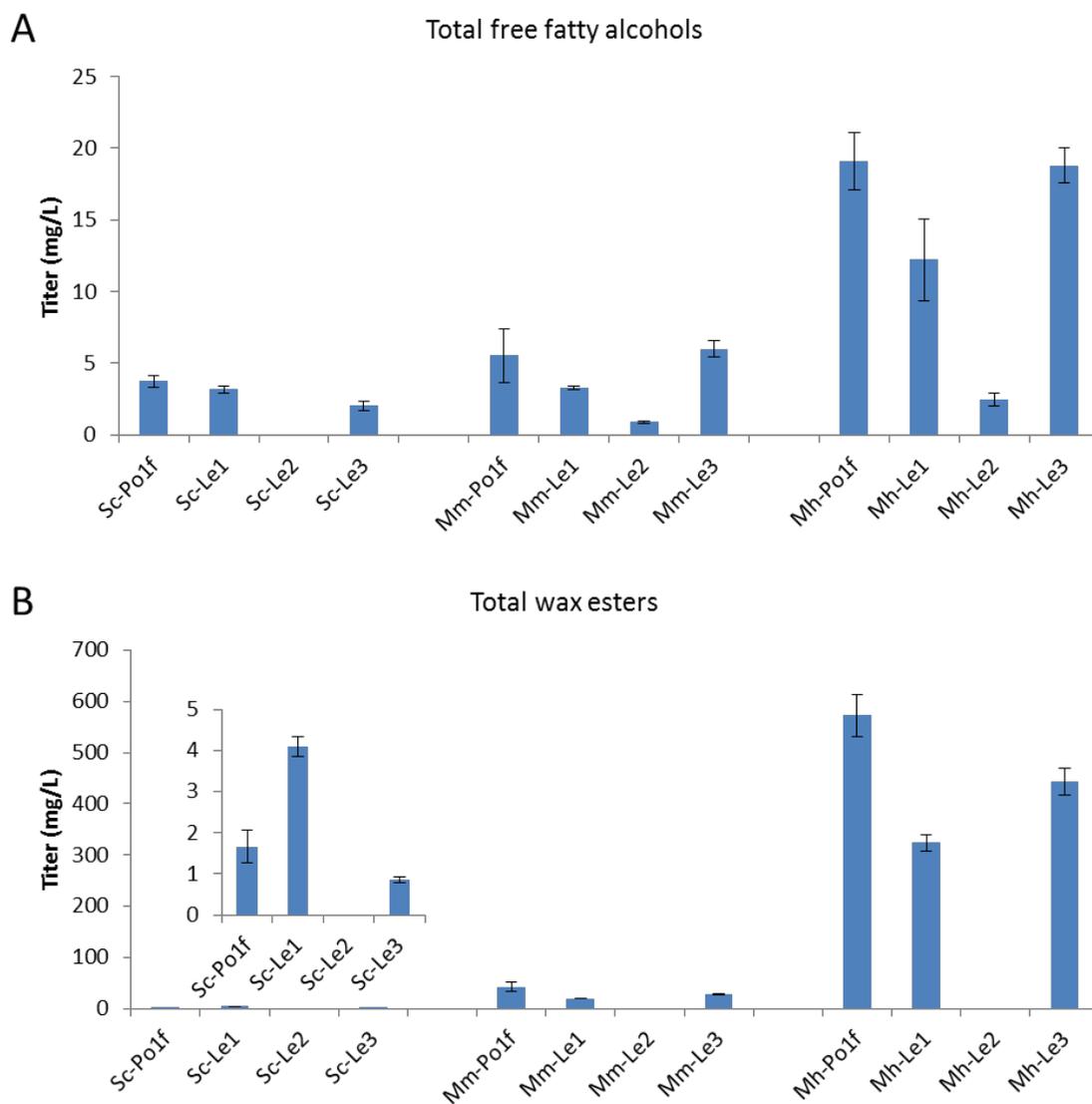


Figure 4.5 Free fatty alcohol (A) and wax ester titers (B) in all the engineered strains. “Sc” represents ScFAR; “Mm” represents MmFAR; “Mh” represents MhFAR. The results showed that MhFAR strains produced the most wax esters compared with ScFAR strains and MmFAR strains. In most cases, parent strain Po1f produced more wax esters and free fatty alcohols than the knockout strains. Le2 gave trace amount of free fatty alcohols and didn’t produce any wax esters.

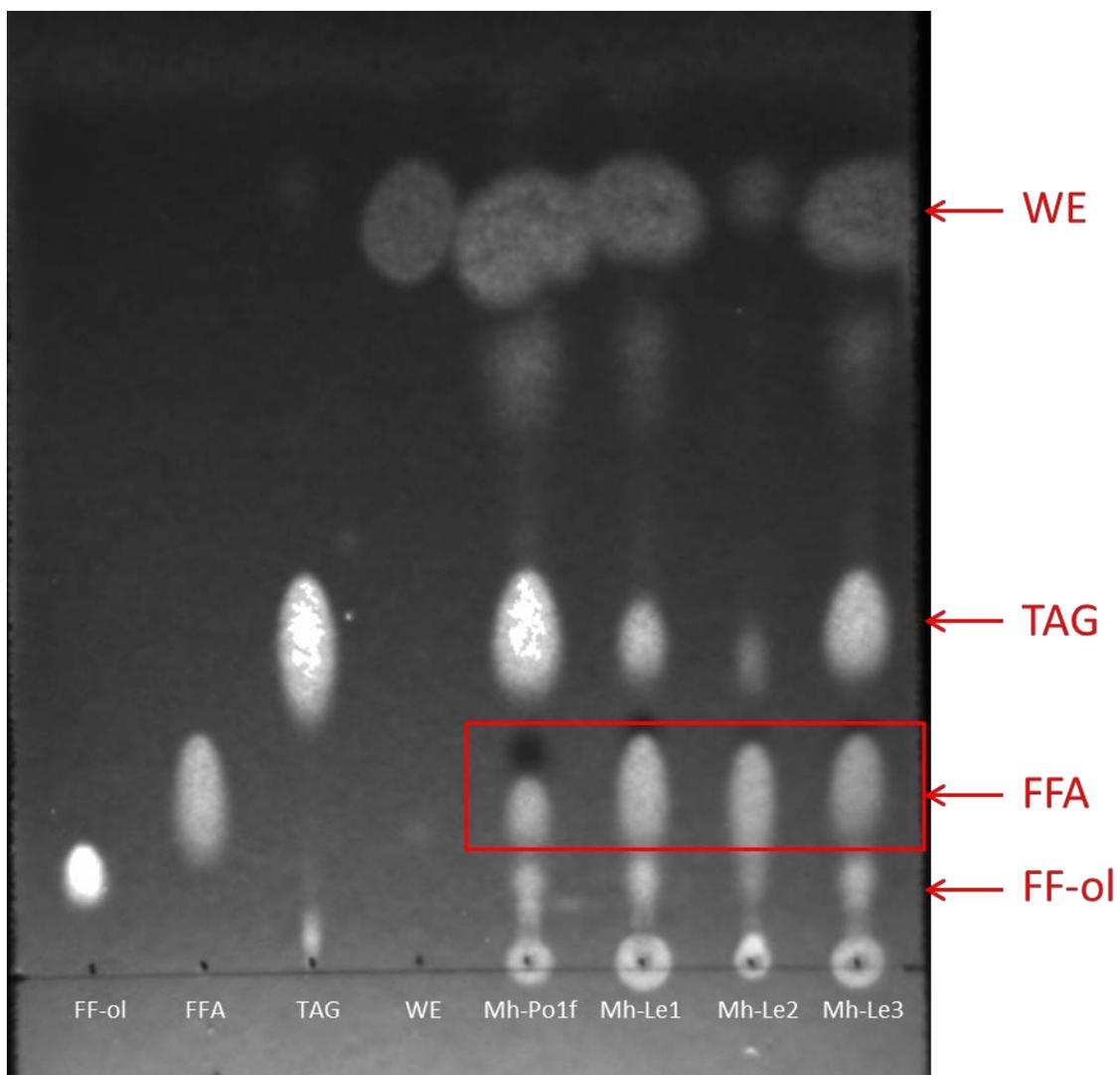


Figure 4.6 Thin layer chromatography analysis of lipid distribution in Mh-Po1f, Mh-Le1, Mh-Le2, and Mh-Le3 strains. FF-ol, free fatty alcohol standard; FFA, free fatty acid standard; TAG, triacylglycerol standard; WE, wax ester standard. Mh-Le1, Mh-Le2, and Mh-Le3 accumulated more FFA than Mh-Po1f did.

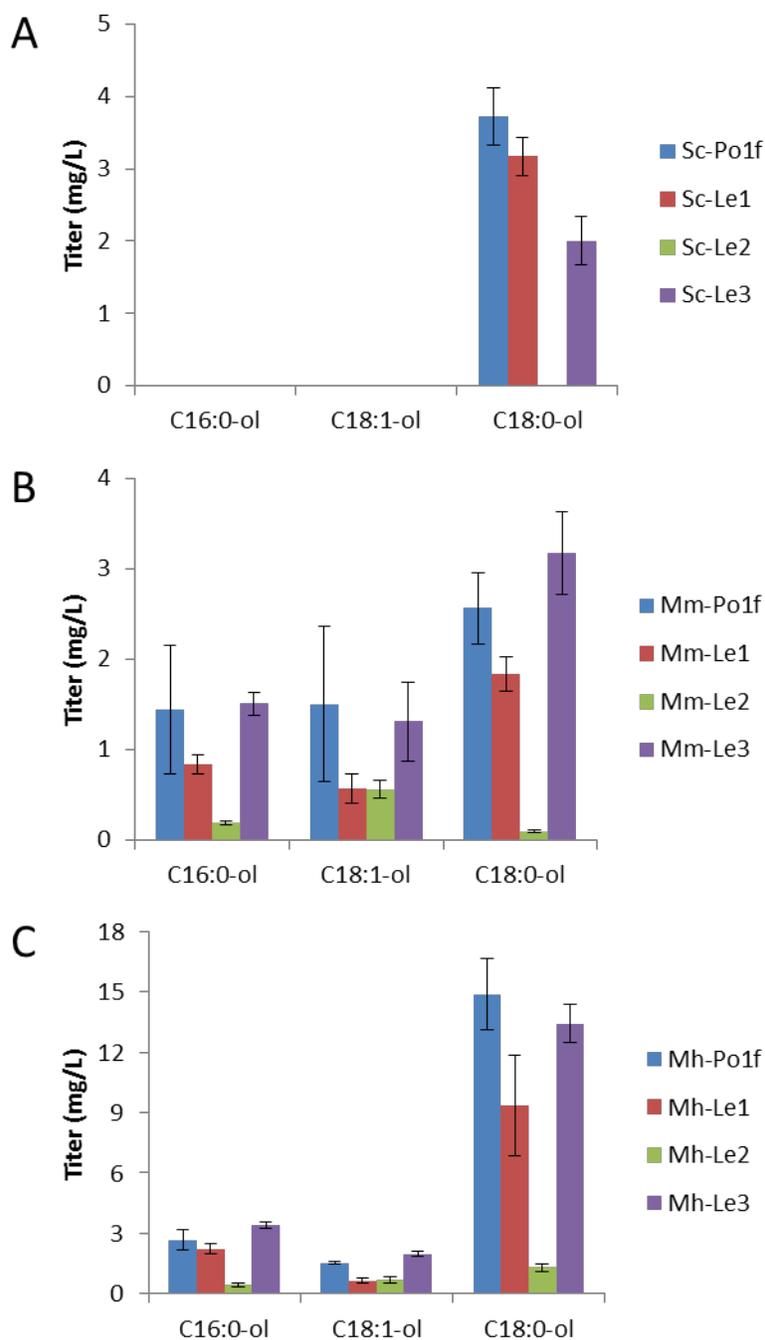


Figure 4.7 Titers of different free fatty alcohol types in all the engineered strains. (A) are the fatty alcohol titers of ScFAR strains. The ScFAR strains only produced C18:0 fatty alcohol. (B) are the fatty alcohol titers of MmFAR strains. (C) are the fatty alcohol titers of MhFAR strains. Strains in both (B) and (C) accumulated C16:0, C18:0 saturated fatty alcohols and C18:1 unsaturated fatty alcohol.

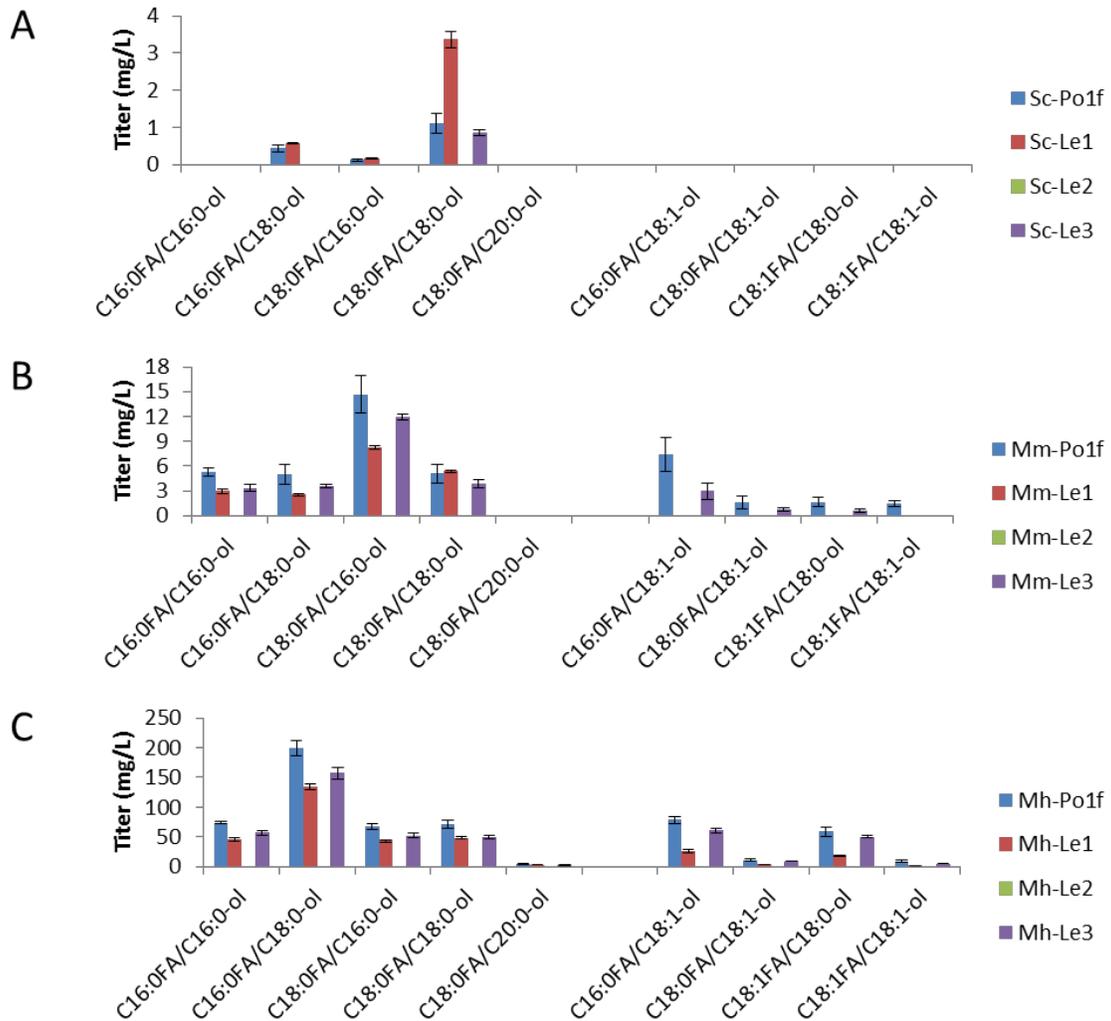


Figure 4.8 Titers of different wax ester types in all the engineered strains. (A) are the wax ester titers of ScFAR strains; (B) are the ones of MmFAR strains; (C) are the ones of MhFAR strains. The wax ester types are diverse among different FAR strains, and they all produced more saturated wax esters than unsaturated ones.

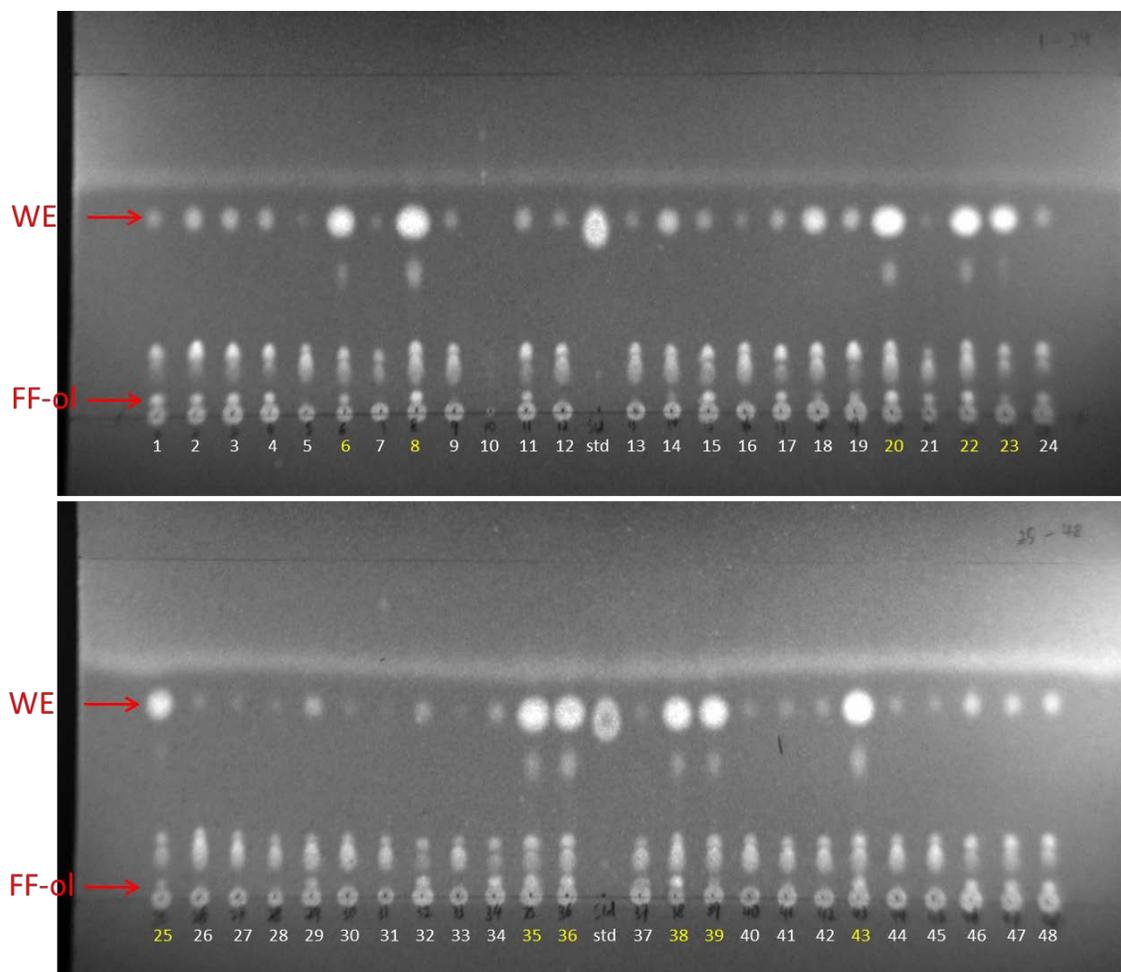


Figure 4.9 Thin layer chromatography of lipid distribution in 48 colonies of SI strain. std: C18:0FA/C18:0-ol wax ester. The lipids in colony 6, 8, 20, 22, 23, 25, 35, 36, 38, 39, 43 (marked in yellow color) were further analyzed via GC-MS.

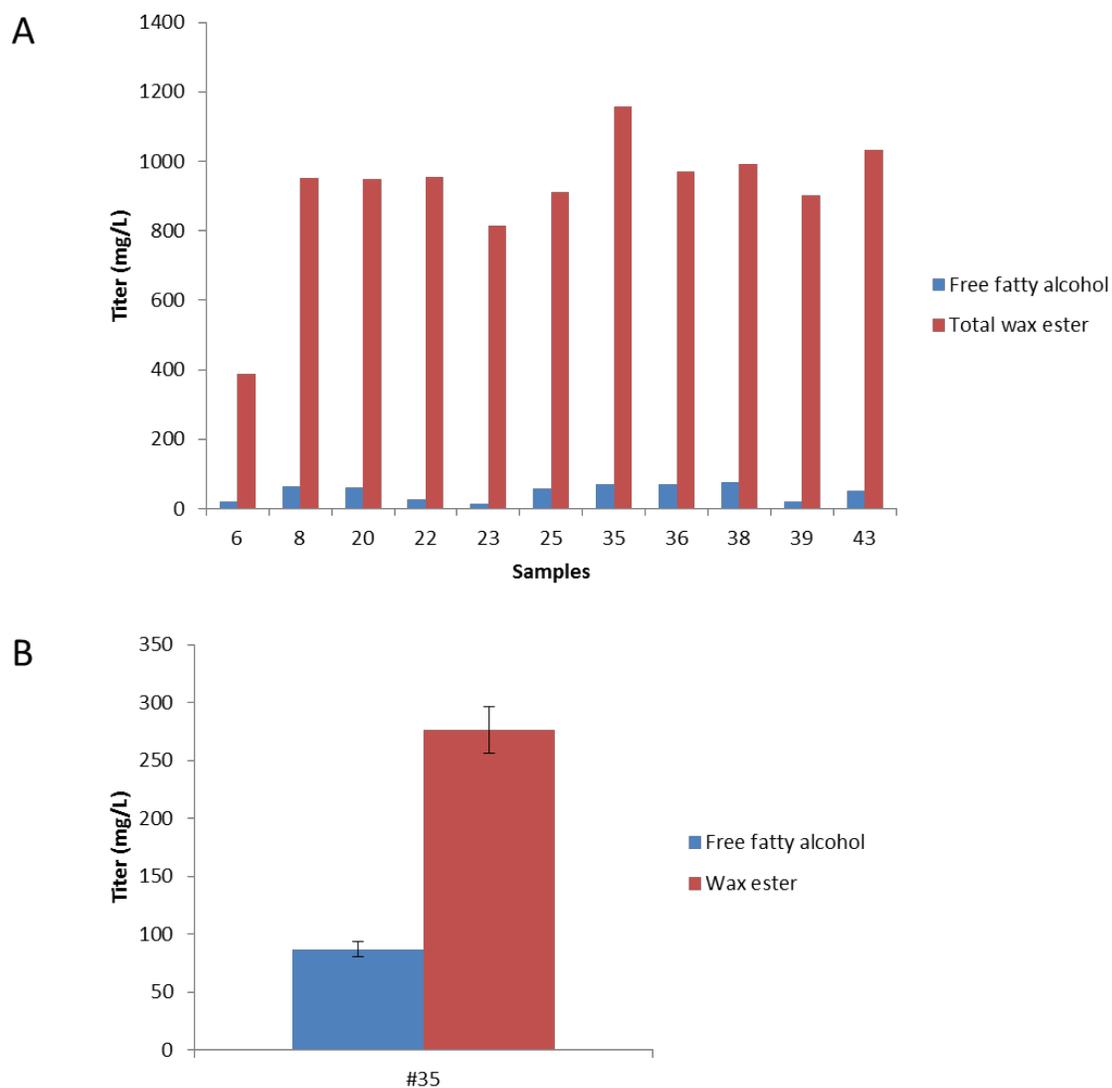


Figure 4.10 (A) Titrers of free fatty alcohols and wax esters produced in the 11 colonies of SI strain. Colony #35 produced the highest titer of wax esters (1100 mg/L), and its free fatty alcohol titer was 70 mg/L. (B) Free fatty alcohol titer (~80 mg/L) and wax ester titer (~270 mg/L) of colony #35 dropped in the later fermentation experiment.

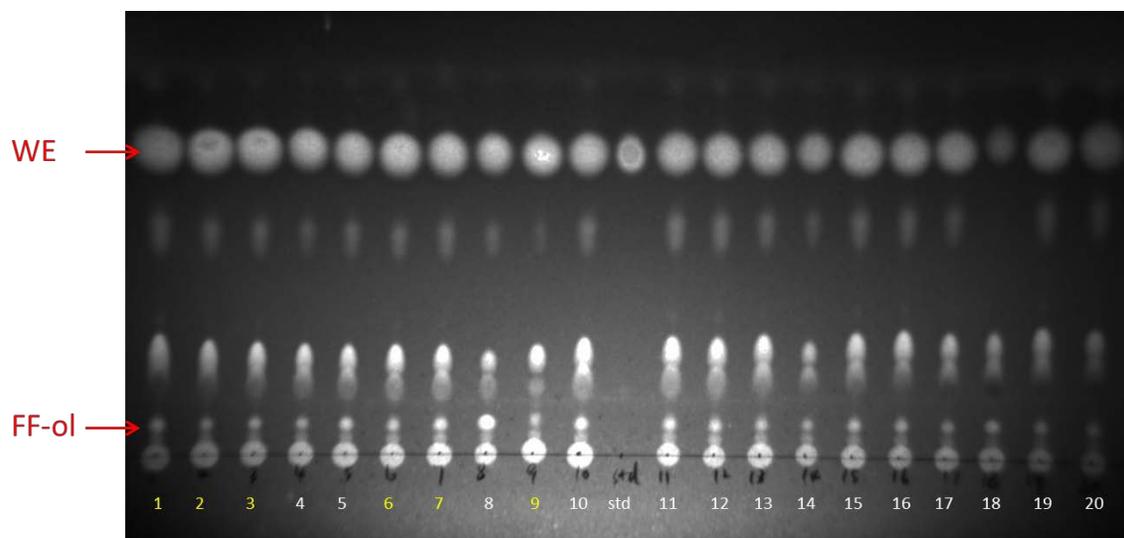


Figure 4.11 Thin layer chromatography of lipid distribution in 20 colonies of DI strain. std: C18:0FA/C18:0-ol wax ester. The lipids in colony 1, 2, 3, 6, 7, 9 (marked in yellow color) were further analyzed via GC-MS.

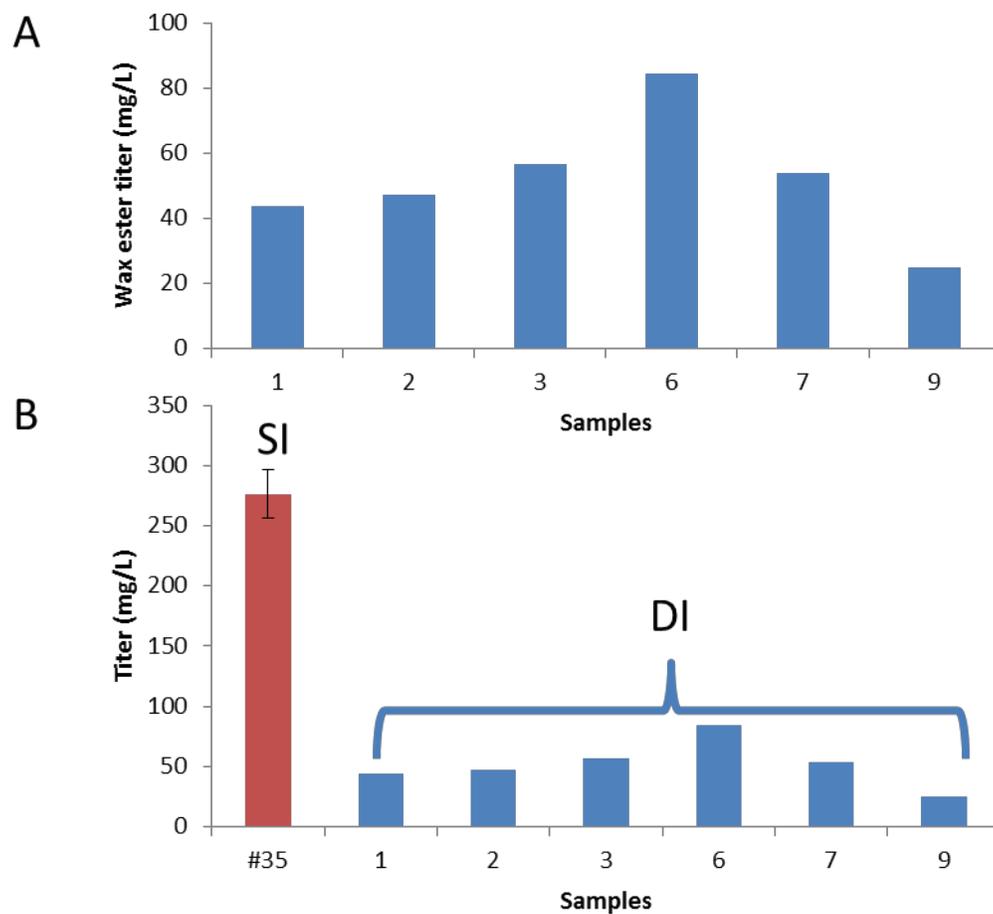


Figure 4.12 (A) Titers and contents of wax esters produced by 6 colonies of DI strain. Colony 6 produced highest amount of wax esters (84mg/L). (B) DI strains produced about 1/4 of the wax ester titer in SI #35 strain.

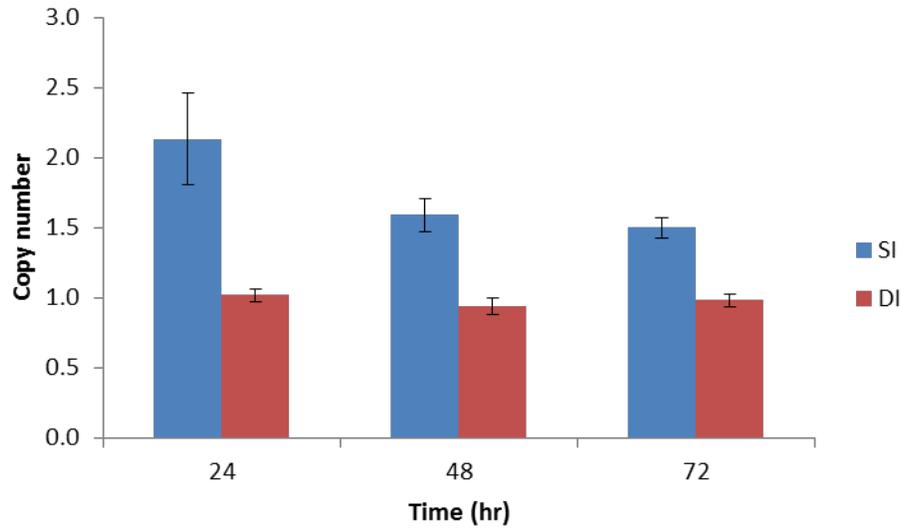
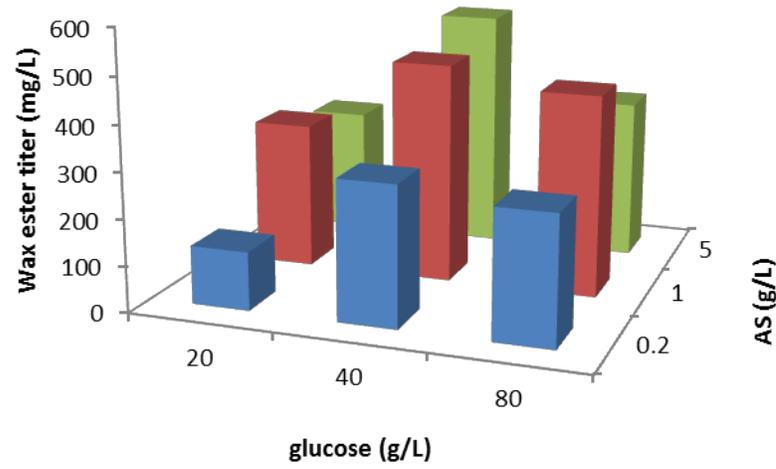


Figure 4.13 Copy numbers of WS gene in SI and DI strains. In DI strain, the genome contained one copy of WS gene; while in SI strain, the copy number of the WS gene or the plasmid decreased from 2.1 (24 hour) to 1.5 (72 hour).

A



B

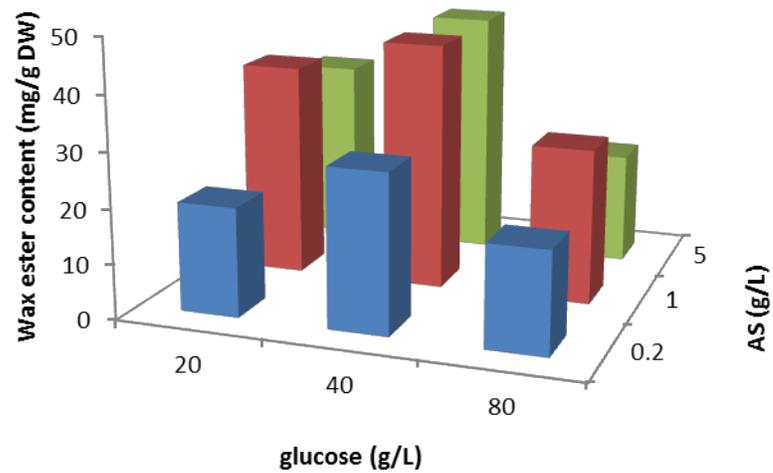


Figure 4.14 Wax ester titer (A) and content (B) in SI strain #35 under different glucose and ammonium sulfate concentration. AS, ammonium sulfate. The highest wax ester titer and content were both achieved under 40 g/L glucose and 1 g/L AS condition.

## References

1. Spencer, G.F. and W.H. Tallent, *Sperm Whale Oil Analysis by Gas-Chromatography and Mass-Spectrometry*. Journal of the American Oil Chemists Society, 1973. **50**(6): p. 202-206.
2. Heilweil, I.J., *Review of Lubricant Properties of Jojoba Oil and Its Derivatives*. Journal of the American Oil Chemists Society, 1988. **65**(1): p. 32-32.
3. Steinke, G., et al., *High-yield preparation of wax esters via lipase-catalyzed esterification using fatty acids and alcohols from crambe and camelina oils*. Journal of Agricultural and Food Chemistry, 2001. **49**(2): p. 647-651.
4. Ucciani, E., et al., *Enzymatic synthesis of some wax-esters*. Fett-Lipid, 1996. **98**(6): p. 206-210.
5. Hallberg, M.L., D.B. Wang, and M. Harrod, *Enzymatic synthesis of wax esters from rapeseed fatty acid methyl esters and a fatty alcohol*. Journal of the American Oil Chemists Society, 1999. **76**(2): p. 183-187.
6. Ishige, T., et al., *Wax ester production by bacteria*. Current Opinion in Microbiology, 2003. **6**(3): p. 244-250.
7. Kalscheuer, R., et al., *Neutral lipid biosynthesis in engineered Escherichia coli: Jojoba oil-like wax esters and fatty acid butyl esters*. Applied and Environmental Microbiology, 2006. **72**(2): p. 1373-1379.
8. Heilmann, M., et al., *Production of wax esters in plant seed oils by oleosomal cotargeting of biosynthetic enzymes*. Journal of Lipid Research, 2012. **53**(10): p. 2153-2161.
9. Santala, S., et al., *Real-Time monitoring of intracellular wax ester metabolism*. Microbial Cell Factories, 2011. **10**.
10. Friedlander, J., et al., *Engineering of a high lipid producing Yarrowia lipolytica strain*. Biotechnol Biofuels, 2016. **9**: p. 77.
11. Gao, S.L., et al., *Multiplex gene editing of the Yarrowia lipolytica genome using the CRISPR-Cas9 system*. Journal of Industrial Microbiology & Biotechnology, 2016. **43**(8): p. 1085-1093.
12. Curran, K.A., et al., *Short Synthetic Terminators for Improved Heterologous Gene Expression in Yeast*. Acs Synthetic Biology, 2015. **4**(7): p. 824-832.
13. Folch, J., M. Lees, and G.H.S. Stanley, *A Simple Method for the Isolation and Purification of Total Lipides from Animal Tissues*. Journal of Biological Chemistry, 1957. **226**(1): p. 497-509.
14. Cao, M., et al., *Centromeric DNA Facilitates Nonconventional Yeast Genetic Engineering*. Acs Synthetic Biology, 2017.

15. Athenstaedt, K. and G. Daum, *Phosphatidic acid, a key intermediate in lipid metabolism*. European Journal of Biochemistry, 1999. **266**(1): p. 1-16.
16. Gajdos, P., et al., *Single cell oil production on molasses by Yarrowia lipolytica strains overexpressing DGA2 in multicopy*. Applied Microbiology and Biotechnology, 2015. **99**(19): p. 8065-8074.
17. Blazeck, J., et al., *Harnessing Yarrowia lipolytica lipogenesis to create a platform for lipid and biofuel production*. Nat Commun, 2014. **5**: p. 3131.
18. Kalscheuer, R., T. Stolting, and A. Steinbuchel, *Microdiesel: Escherichia coli engineered for fuel production*. Microbiology-Sgm, 2006. **152**: p. 2529-2536.
19. Steen, E.J., et al., *Microbial production of fatty-acid-derived fuels and chemicals from plant biomass*. Nature, 2010. **463**(7280): p. 559-62.
20. Stoveken, T., et al., *The wax ester synthase/acyl coenzyme A:diacylglycerol acyltransferase from Acinetobacter sp. strain ADP1: characterization of a novel type of acyltransferase*. J Bacteriol, 2005. **187**(4): p. 1369-76.
21. Miklaszewska, M. and A. Banas, *Biochemical characterization and substrate specificity of jojoba fatty acyl-CoA reductase and jojoba wax synthase*. Plant Science, 2016. **249**: p. 84-92.
22. Cheng, J.B. and D.W. Russell, *Mammalian wax biosynthesis. II. Expression cloning of wax synthase cDNAs encoding a member of the acyltransferase enzyme family*. J Biol Chem, 2004. **279**(36): p. 37798-807.
23. Hofvander, P., T.T.P. Doan, and M. Hamberg, *A prokaryotic acyl-CoA reductase performing reduction of fatty acyl-CoA to fatty alcohol*. Febs Letters, 2011. **585**(22): p. 3538-3543.
24. Cheng, J.B. and D.W. Russell, *Mammalian wax biosynthesis - I. Identification of two fatty acyl-coenzyme A reductases with different substrate specificities and tissue distributions*. Journal of Biological Chemistry, 2004. **279**(36): p. 37789-37797.
25. Fournier, P., et al., *Colocalization of Centromeric and Replicative Functions on Autonomously Replicating Sequences Isolated from the Yeast Yarrowia-Lipolytica*. Proceedings of the National Academy of Sciences of the United States of America, 1993. **90**(11): p. 4912-4916.
26. Sattur, A.P. and N.G. Karanth, *Production of Microbial Lipids .2. Influence of C/N Ratio - Model Prediction*. Biotechnology and Bioengineering, 1989. **34**(6): p. 868-871.
27. Kerkhoven, E.J., et al., *Regulation of amino-acid metabolism controls flux to lipid accumulation in Yarrowia lipolytica*. npj Systems Biology and Applications, 2016. **2**(1).
28. Metz, J.G., et al., *Purification of a jojoba embryo fatty acyl-coenzyme A reductase and expression of its cDNA in high erucic acid rapeseed*. Plant Physiology, 2000. **122**(3): p. 635-644.

29. Moreau, R.A., *The analysis of lipids via HPLC with a charged aerosol detector*. *Lipids*, 2006. **41**(7): p. 727-734.
30. Khoomrung, S., et al., *Rapid Quantification of Yeast Lipid using Microwave-Assisted Total Lipid Extraction and HPLC-CAD*. *Analytical Chemistry*, 2013. **85**(10): p. 4912-4919.
31. Santala, S., et al., *Rewiring the wax ester production pathway of *Acinetobacter baylyi* ADPI*. *Acs Synthetic Biology*, 2014. **3**(3): p. 145-51.

## CHAPTER V. SUMMARY

There are two complex metabolic pathways involved in my research projects, one is TIA pathway in *C. roseus*, and the other one is lipid pathway in *Y. lipolytica*. The commonality in these two distinct pathways is that they both involve one key metabolite: for TIA pathway, tabersonine is the key metabolite; and for lipid pathway, acyl-CoA is the key metabolite. The characteristics of these two key metabolites are: (1) the pathways derived from the key metabolites are highly branched; (2) the key metabolites are tightly regulated. For example, in Figure 2.5b, the tabersonine level kept constant in both induced hairy root lines and uninduced hairy root lines. It is also reported by various literature that acyl-CoA is tightly regulated in many organisms [1, 2]. In the wax ester project, we conjectured that the reason for low wax ester production in knockout strains was the tightly regulated acyl-CoA.

One of the research goals for my projects is to increase the metabolic flux to selected high-value products derived from the key metabolites. The “push and pull” strategy of pathway engineering has been considered in the projects. “Push” means to increase the precursor pools, and in my projects the precursors are tabersonine and acyl-CoA. “Pull” means to improve the enzyme conversion capability in the target biosynthetic pathways, such as the enzymes in the vindoline pathway and the ones in the wax ester pathway.

For the “push” strategy, we knocked out a few competitive pathways in the wax ester project, but did not observe expected improvement of wax ester production, which indicated the acyl-CoA pool may not be increased. In the TIA projects, we overexpressed a few transcriptional activators and knocked down a few transcriptional repressors, expecting to increase the TIA metabolite levels, but the TIA levels either kept constant or went up in transient, especially tabersonine level was not changed. All these results indicated that the “push” strategy was not

working efficiently, presumably mainly because the key metabolites, acyl-CoA and tabersonine, are tightly regulated. The multiple outputs of the key metabolites make it challenging to lead the flux to the desired pathways.

For the “pull” strategy, we found that the expression of three FARs from different sources affected wax ester titer significantly. And in the TIA projects, overexpression of the first two enzymes in the vindoline pathway also drove the metabolic flux to the vindoline pathway efficiently. All these data showed that improving enzyme conversion in the target biosynthetic pathways is an efficient way to enhance the flux to products.

In conclusion, the “pull” strategy is more applicable than the “push” strategy in these two distinct cases, and it suggests a direction for future research. Specifically in the wax ester project, improving the FAR and WS enzyme activities would be an effective strategy to further improve the wax ester production.

### References

1. Ellis, J.M., C.E. Bowman, and M.J. Wolfgang, *Metabolic and tissue-specific regulation of acyl-CoA metabolism*. Plos One, 2015. **10**(3): p. e0116587.
2. Black, P.N., N.J. Faergeman, and C.C. DiRusso, *Long-chain acyl-CoA-dependent regulation of gene expression in bacteria, yeast and mammals*. J Nutr, 2000. **130**(2S Suppl): p. 305S-309S.