

Pathway enzyme engineering for flavonoid production in recombinant microbes



Jian Zha^{a,**}, Xia Wu^a, Guoli Gong^a, Mattheos A.G. Koffas^{b,c,*}

^a School of Food and Biological Engineering, Shaanxi University of Science and Technology, Xi'an, 710021, Shaanxi, China

^b Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, 12180, NY, USA

^c Department of Chemical and Biological Engineering, Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, 12180, NY, USA

ARTICLE INFO

Keywords:

Flavonoids
Phenylpropanoic acids
Isoflavonoids
Anthocyanins
Flavonols
Flavones

ABSTRACT

Metabolic engineering of microbial strains for the production of flavonoids of industrial interest has attracted great attention due to its promising advantages over traditional extraction approaches, such as independence of plantation, facile downstream separation, and ease of process and quality control. However, most of the constructed microbial production systems suffer from low production titers, low yields and low productivities, restricting their commercial applications. One important reason of the inefficient production is that the expression conditions and the detailed functions of the flavonoid pathway enzymes are not well understood. In this review, we have collected the biochemical properties, structural details, and genetic information of the enzymes in the flavonoid biosynthetic pathway as a guide for the expression and analysis of these enzymes in microbial systems. Additionally, we have summarized the engineering approaches used in improving the performances of these enzymes in recombinant microorganisms. Major challenges and future directions on the flavonoid pathway are also discussed.

1. Introduction

Flavonoids are a class of polyphenolic compounds naturally present in plants, which contain a C6–C3–C6 skeleton (Fig. 1). These compounds are produced against biotic and abiotic stresses when plants are exposed to external stimuli such as pathogen invasion, environmental stresses, physical injury, etc. (Chouhan et al., 2017; Treutter, 2006). In plants, the biosynthesis of flavonoids starts with the deamination of phenylalanine or tyrosine to cinnamic acid or coumaric acid by phenylalanine ammonia lyase (PAL) or tyrosine ammonia lyase (TAL) (Fig. 2). Next, cinnamic acid or coumaric acid is converted to corresponding CoA ester under the action of 4-coumarate: CoA ligase (4CL). The CoA ester is subsequently condensed with malonyl-CoA at a 1:3 M ratio by chalcone synthase (CHS) to form chalcone, which is then isomerized to form (2S)-flavanones catalyzed by chalcone isomerase (CHI). Flavanones can be modified by various enzymes, such as hydroxylases, glycosyltransferases, and methyltransferases, to produce diverse flavonoids. For example, hydroxylation of (2S)-flavanones at C3-position by flavanone 3 β -hydroxylase gives rise to dihydroflavonols, whose reduction is then catalyzed by dihydroflavonol 4-reductase (DFR) at C4-position to produce

leucoanthocyanidins. Leucoanthocyanidins are reduced by leucoanthocyanidin reductase (LAR) to generate flavan-3-ols. Leucoanthocyanidins or flavan-3-ols under the action of anthocyanidin synthase (ANS) can be used to synthesize anthocyanidins, which are then glycosylated at C3 by flavonoid 3-O-glucosyltransferase (3GT), yielding anthocyanins (Zha and Koffas, 2017). Diverse modifications such as hydroxylation, methylation, prenylation and glycosylation can occur to the baseline flavonoid molecules, giving rise to a large number of flavonoids with diverse structures and different properties.

Flavonoids are very important for physiological activities of plants and they also have wide applications for humans. A wide range of flavonoids have been used as pharmaceuticals, nutraceuticals, and food additives, etc. (Chouhan et al., 2017; Trantas et al., 2009). Traditionally, the major way of supplying flavonoids is extraction from plants. However, this method has a few limitations, such as heavy dependence on land and climate, seasonal and regional fluctuations in supply, and variations in flavonoid content from different sources (Tsai et al., 2008; Zhu et al., 2017). Another way of providing flavonoids is chemical synthesis, which, however, often requires the use of chemically hazardous catalysts, and is only limited to flavonoids with simple chemical structures (such as

* Corresponding author. Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, 12180, NY, USA.

** Corresponding author.

E-mail addresses: zhajian1985@sust.edu.cn (J. Zha), koffam@rpi.edu (M.A.G. Koffas).

<https://doi.org/10.1016/j.mec.2019.e00104>

Received 5 September 2019; Received in revised form 10 October 2019; Accepted 13 October 2019

2214-0301/© 2019 Published by Elsevier B.V. on behalf of International Metabolic Engineering Society. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

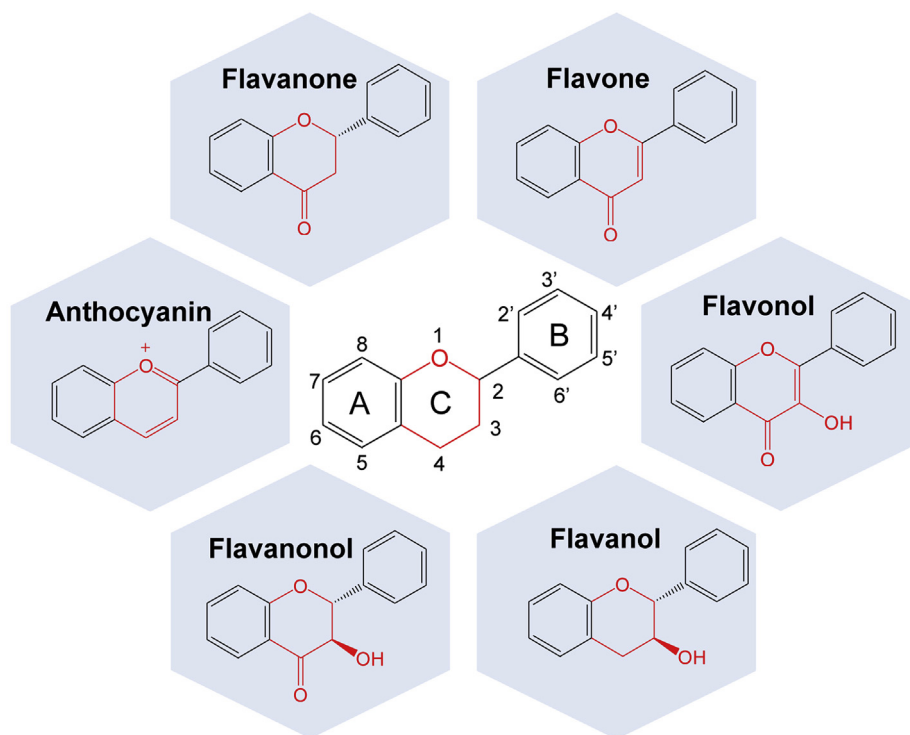


Fig. 1. The core skeleton of flavonoids. The common flavonoids contain six subgroups, including flavanone, flavone, flavonol, flavanol, flavanonol, and anthocyanin.

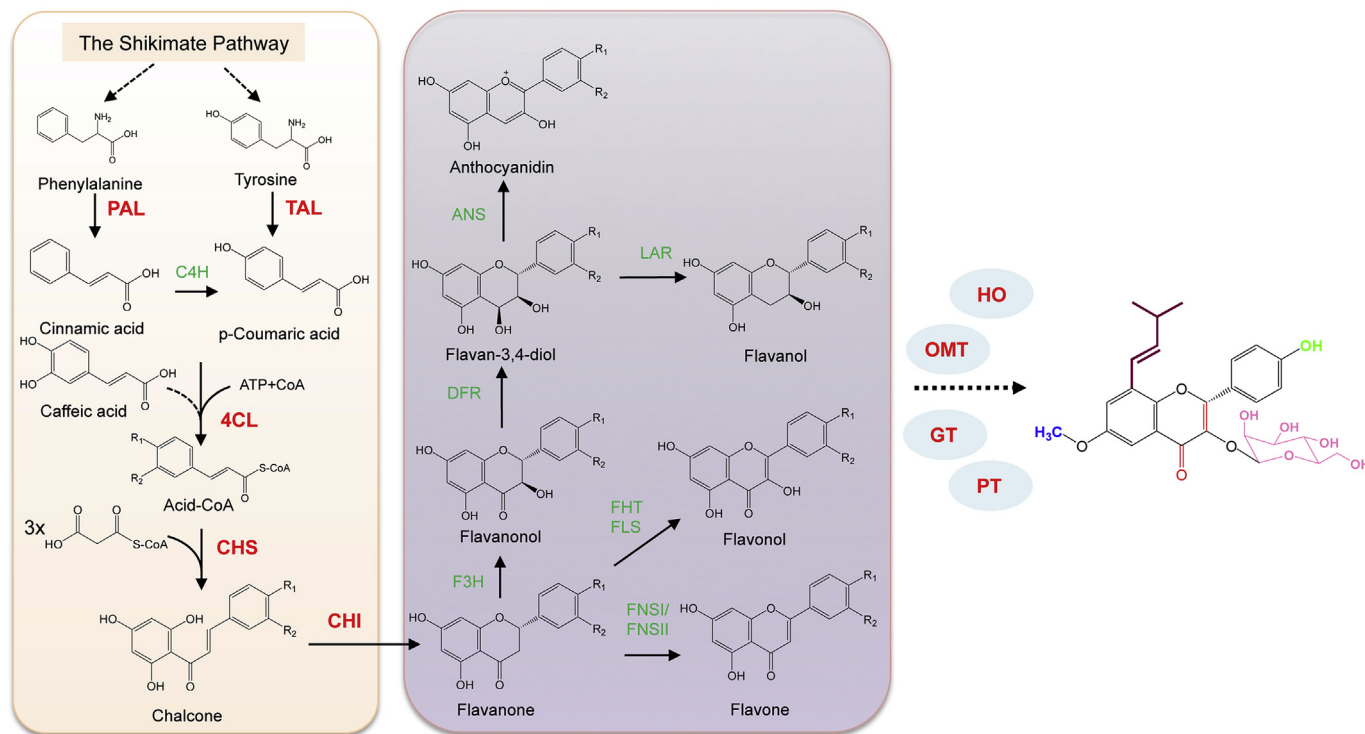


Fig. 2. The flavonoid biosynthesis pathway in plants. The enzymes marked in bold red are the ones described in this review. PAL, phenyl ammonia lyase; TAL, tyrosine ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate: CoA ligase; CHS, chalcone synthase; CHI, chalcone synthase; FNSI, soluble flavone synthase; FNSII, membrane-bound flavone synthase; FHT, flavanone 3 β -hydroxylase; FLS, flavanol synthase; F3H, flavanone 3-hydroxylase; DFR, dihydroflavonol 4-reductase; LAR, leucocyanidin reductase; ANS, anthocyanidin synthase; HO, hydroxylase; OMT, O-methyltransferase; GT, glycotransferase; PT, prenyltransferase. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

naringenin) due to the presence of multiple active functional groups such as hydroxyls and the requirement for accurate modifications of such groups at particular positions (Wang et al., 2015). As an alternative,

microbial cells can be employed to synthesize flavonoids in a sustainable and economical manner by metabolic engineering approaches (Chemler and Koffas, 2008; Wang et al., 2016b).

Microbial cell factories hold some distinct advantages over plant extraction and chemical synthesis in terms of flavonoid production. For example, this strategy is ecofriendly, has facile downstream processing, and provides stable production. Microbial strains can be easily grown, and sophisticated metabolic engineering tools are available for elaborate metabolic engineering modifications. Moreover, this approach also has the potential to produce new flavonoid derivatives which may have novel pharmaceutical or nutraceutical applications (Fowler et al., 2011; Leonard et al., 2008). Metabolic engineering of microbes for the production of flavonoids has been carried out for two decades, and the microbes engineered include *E. coli*, *Saccharomyces cerevisiae*, *Streptomyces venezuelae*, and *Corynebacterium glutamicum* (Kallscheuer et al., 2017; Pandey et al., 2016). A variety of flavonoid compounds has been successfully synthesized in these microbial strains. However, the production ability, including titer, yield and productivity, has not fulfilled the industrial requirements.

To optimize the microbial cell factories of flavonoid production for industrial applications, many efforts have been attempted to optimize the introduced pathways, including selection of gene orthologues, promoters, and ribosome binding sites, enhancement of cosubstrate supply, and regulation of product relocation. The relevant work has been well summarized in recent reviews (Chouhan et al., 2017; Pandey et al., 2016). However, the performances of the resultant strains still do not meet the requirement for commercial applications. Given that the pathway enzymes are key to efficient bioproduction, and that the detailed information of many of these enzymes is still under exploration, it is critical to find novel insights from these enzymes and propose new approaches to optimize the pathways and improve the production titers.

In this review, we summarize the major enzymes involved in flavonoid biosynthesis, and review their biochemical properties, structural details, and genetic information. In addition, we cover the strategies applied for functionally expressing these enzymes in recombinant microbes to improve flavonoid production.

2. Plant enzymes involved in the biosynthetic pathway of flavonoids

Construction of an efficient microbial platform to produce diverse flavonoids in microorganisms requires an extensive understanding of the metabolic pathway enzymes in their natural hosts, i.e., plants, and their expression conditions in engineered microbial hosts. Such information is dependent on the exploration of kinetics of the enzymatic reaction, the substrate preference, the structure, the enzyme evolution, the biochemical properties, etc., which will be the focus of this section.

2.1. PAL/TAL

PAL or TAL catalyzes the first step of the general phenylpropanoid pathway, the nonoxidative deamination of phenylalanine or tyrosine. This step is a joint reaction common to the production of many metabolites including flavonoids, coumarins, phytoalexins, and lignin (Fig. 2). The deamination reaction is carried out using a 3,5-dihydro-5-methylidene-4H-imidazole-4-one (MIO) prosthetic group to activate the substrate, which is formed through spontaneous cyclization and dehydration of an alanine-serine-glycine segment of the polypeptide backbone in the active site of PAL (Schroeder et al., 2008). This enzyme is present in the form of a tetramer, and is inhibited by its product trans-cinnamic acid and the phenylalanine analogue aminoxyphenylpropionic acid (Mavandad et al., 1990).

Many PALs, no matter of plant origin or microbial origin, present TAL activity, and are also termed TALs. Interestingly, almost all the PALs and TALs show activity towards both phenylalanine and tyrosine, whereas most of these enzymes have a higher preference for phenylalanine over tyrosine (Vannelli et al., 2007). The TAL from the red yeast *Rhodotorula glutinis* (RgTAL) has the strongest substrate preference for tyrosine over phenylalanine and the highest specific activity although the TAL activity

is around half of its PAL activity (Vannelli et al., 2007). Therefore, it is the most commonly used TAL for microbial production of flavonoids with tyrosine being the substrate.

Although PALs and TALs share many similarities and sometimes can be used interchangeably, in certain cases, a high TAL activity with low PAL activity is necessary for the highly efficient production of certain compounds that use tyrosine as the precursor, such as naringenin. Given that the crystal structures are available for PALs from *Rhodobacter sphaeroides*, *Rhodospiridium toruloides*, and *Petroselinum crispum*, as well as histidine ammonia-lyase from the bacterium *Pseudomonas putida* (Louie et al., 2006), homology modeling and rational design based on structural information can help with the generation of novel TALs with higher tyrosine specificity and activity.

2.2. 4CL

In plants, 4CLs are present in multiple isoforms and encoded by a gene family. These enzymes belong to the adenylate-forming enzyme family and catalyze the formation of phenylpropanoic acid-CoA complexes, which are further utilized for biosynthesis of phenylpropanoids. The catalytic process of 4CLs contains two steps, in which the substrate is converted to adenylate intermediate and then to thioester (Li and Nair, 2015).

4CLs are present in the form of multiple isozymes, which exhibit distinct substrate affinity that appears to coincide with specific metabolic functions (Lavhale et al., 2018). Physiological studies have shown that specific isoforms of 4CLs are differentially regulated upon exposure to diverse stimuli, leading the metabolic flux toward particular metabolites as desired, such as lignin for structural support, flavones and flavonols for UV protection, anthocyanins as pigments for the attraction of pollinators and seed distributors, and isoflavonoids and furanocoumarins as phytoalexins for pathogen defense (Douglas, 1996). Thus, 4CLs are one of the key branch point enzymes in the phenylpropanoid pathway, regulating the production of different phenylpropanoids according to the needs of plants.

4CL genes can be classified into two broad groups. Class I 4CL genes play important roles in the biosynthesis of lignin and other phenylpropanoid derivatives, and are found in most dicotyledonous plants, such as *Arabidopsis At4CL1* and *At4CL2*, *Populus Pt4CL1*, and soybean *Gm4CL2 Pt4CL1* (Sun et al., 2013). Class II 4CL genes are associated with the production of flavonoids and antioxidants, and are mainly found in monocotyledonous plants and gymnosperms, as well as certain dicotyledonous plants, such as *Arabidopsis At4CL3*, *Pt4CL2*, and rice 4CL. Evolutionary studies indicate that 4CLs separately evolve after monocot and dicot segregation (Cukovica et al., 2001). Given the diverse functions and substrate preferences of 4CLs, the selection of appropriate gene/enzyme for flavonoid production is of great importance. For example, *in vitro* assays show that 4CL from *Petroselinum crispum* has higher enzymatic activity than that from rice or *Plagiochasma appendiculatum* (Zang et al., 2019).

In the heterologous expression of 4CLs, Santos et al found 4CL-mediated suppression of TAL activity and recovery of TAL activity through adequate pathway balancing, suggesting that accumulation of coumaroyl-CoA could inhibit the activity of TALs, which was not observed in previous *in vivo* or *in vitro* studies (Santos et al., 2011).

2.3. CHS

CHS catalyzes the condensation of one molecule of phenylpropanoic acid-CoA and three molecules of malonyl-CoA to form one molecule of chalcone. CHSs belong to the family of type III polyketide synthases (PKSs), which exist as homodimers based on a ketosynthase domain (Dao et al., 2011). Type III PKSs belong to the "initiation enzyme" group characterized by a Cys-His-Asn catalytic triad. In contrast, type I and type II PKSs, which contain a Cys-His-His catalytic triad, are "elongation enzymes" (Dao et al., 2011; Yu et al., 2012). CHSs and other type III PKSs,

such as stilbene synthase, curcuminoid synthase, and pentaketide synthase, can take different numbers of malonyl-CoA molecules to form diversified skeleton structures, followed by various modifications to form a wide range of secondary metabolites. Type III PKSs show very high similarity in the amino acid sequences, indicating that they may evolve from the same origin and diverge into different groups during long-term evolution due to the requirement of plants to cope with environmental challenges.

The crystal structures of CHS from Alfalfa and apples are available. The active pocket of Alfalfa CHS and STS from *Pinus sylvestris* was determined and reconstituted (T132A or S131A mutation) to produce both chalcones and stilbenes (Austin et al., 2004). Other mutations in the active sites not only change the reaction types, but also invert the substrate specificity. For example, the active site of CHS from *Hypericum sampsonii* was mutated to L263M-F265Y-S388G and the triple mutant showed higher substrate affinity of benzoyl-CoA than 4-coumaroyl-CoA (Liu et al., 2003).

In metabolic engineering of microorganisms to produce flavonoids and other phenylpropanoids, the CHS-catalyzed reaction is a critical step. The condensation is a dual-substrate reaction, and requires the supply of malonyl-CoA, which is involved in many cellular reactions (Wu et al., 2015). However, it is very difficult for CHSs to compete for malonyl-CoA binding with other endogenous enzymes which have very efficient reactivity with malonyl-CoA after thousands of years of evolution in nature. So far, metabolic engineering efforts have mainly focused on selection of CHS gene orthologues from different plants, whereas direct protein engineering to improve the specific activity has not been reported.

2.4. CHI

CHIs catalyze the intramolecular cyclization of chalcones to form tricyclic (2S)-flavanones. Chalcones can spontaneously cyclize in solution to produce an enantiomeric mixture of flavanones, so expression of CHIs in plants guarantees the generation of biologically active (2S)-flavanones for plants (Jez and Noel, 2002). Furthermore, the rate of enzyme-catalyzed cyclization is 10^7 -fold higher than that of the spontaneous cyclization (Jez and Noel, 2002).

CHIs are generally grouped into two classes according to their substrates. Class I CHIs can only convert 6'-hydroxychalcone into 5-hydroxyflavanone and are ubiquitous among plants (Shimada et al., 2003). Class II CHIs, mostly found in leguminous plants, can convert both 6'-hydroxychalcone and 6'-deoxychalcone to 5-hydroxyflavanone and 5-deoxyflavanone, respectively.

The intramolecular cyclization reaction catalyzed by CHIs requires a hydroxyl group at the 2'-position in chalcone. The reaction begins with the deprotonation of the 2'-hydroxyl of chalcone, causing the 2'-oxy anion to attack the β -carbon atom of the α , β -unsaturated double bond to produce the 2'-acidic hydroxyl group (Yin et al., 2019). The catalytic activity of CHI is pH-dependent. For example, the activity at pH 6 is about half of that at pH 7 (Jez and Noel, 2002).

Researchers have identified more than 3,000 nucleotide sequences of CHIs and have tested their expression in *E. coli* and *S. cerevisiae* (Yin et al., 2019). A 926 bp-CHI from *Ginkgo biloba* L. was successfully expressed in *E. coli* and its activity was confirmed by the formation of flavonoids (Cheng et al., 2011). The expression of CHI from *Milletia pinnata* was conducted in *S. cerevisiae*, which enhanced yeast's tolerance to salt treatment (Wang et al., 2013). Additionally, overexpression of CHIs has been used in hairy roots to confirm their functions and their enhancing effects on flavonoid levels (Li et al., 2006). These studies demonstrate that CHI-catalyzed reaction is a rate-limiting step in flavonoid biosynthesis, and efficient expression of CHIs in microorganisms is essential for construction of microbial cell factories for flavonoid production.

2.5. Hydroxylases

Hydroxylation of flavonoids is an important post modification for the biosynthesis of a diverse range of flavonoids. To generate diversity in structural chemistry of flavonoids and to produce novel therapeutically or nutraceutically functional compounds, flavonoids are hydroxylated by various hydroxylases.

Hydroxylation typically occurs at C3, C6 carbons and the B ring. Flavanone 3-hydroxylase can add a hydroxyl at C3 position with the help of iron and ascorbate (Zhang et al., 2018). Recently, a flavone 6-hydroxylase from plant *Scutellaria baicalensis* was identified and the removal of the membrane recognition signal led to functional cytoplasmic expression of this enzyme and the production of 6-hydroxyflavones (Li et al., 2019). However, the substrate specificity of this enzyme has not been studied.

Hydroxylation of the B ring has been extensively studied. The enzymes corresponding to hydroxylation at specific positions have been discovered and tested in microbial strains, such as flavanone 3'-hydroxylase for hydroxylation at C3' position. Most of these hydroxylases are P450 enzymes which are anchored on membranes. The hydroxylation process requires cytochrome P450 reductases (CPRs), which are also located on membranes such as endoplasmic reticulum membrane and help to transfer electrons to receptors. In plants, multiple P450s can use a single CPR and an approximate P450/CPR ratio is 15:1 (Jensen and Møller, 2010). So in metabolic engineering for flavonoid production, selection of CPRs and modulation of P450 hydroxylase/CPR ratio for optimal hydroxylation is generally required.

In nature, apart from P450 hydroxylases, there are non-P450 hydroxylases. Jones et al identified a non-P450 hydroxylase, HpaBC, native to *E. coli* and *P. aeruginosa* with high activity towards phenylpropanoic acids as well as flavanones. Expression of HpaBC could implement the *ortho*-hydroxylation of *p*-coumaric acid and naringenin but not the flavan-3-ol afzelechin (Jones et al., 2016).

2.6. Methyltransferases

Methylation helps to maintain the physiological functions of flavonoids in plants, as it improves the stability and membrane permeability of flavonoid compounds (Koirala et al., 2016). Methylation in plants is catalyzed by methyltransferases, which transfer S-adenosyl-L-methionine (SAM) to flavonoids at hydroxyl groups, with a preference for C7-hydroxyl in A ring and hydroxyls in B ring, although the reason for such a position preference is unknown (Pandey et al., 2016).

Methyltransferases from plants are SAM-dependent caffeoyl-CoA O-methyltransferases (OMTs). However, a few of the microbially originated SAM-dependent OMTs are also found to methylate plant flavonoids. For example, the OMT from *Streptomyces peucetius* ATCC 27952 (SpOMT2884) was functionally overexpressed in *E. coli* and used for the methylation of flavonoid compounds (Kim et al., 2005). So far, the crystal structure of the flavonoid methyltransferase *Bacillus cereus* BcOMT2 has been resolved, and the SAM-binding domain and the metal-dependent catalytic site have been identified, which provides theoretical basis for structural analysis and engineering of other methyltransferases (Cho et al., 2007).

In microbial production of methylated flavonoids, the supply of SAM is a critical issue besides the substrate specificity and activity of OMTs. The generation of SAM undergoes feedback repression by methionine biosynthesis regulator *MetJ* based on the intracellular SAM concentrations, thus limiting its accumulation and the rate of methylation (Cress et al., 2017). To increase the availability of SAM, CRISPRi-mediated silencing of *MetJ* was performed for the production of peonidin 3-O-glucoside (a methylated flavonoid compound) from catechin in *E. coli*, and a twofold increase in the production titer was achieved (Cress et al., 2017).

2.7. Glycosyltransferases

Glycosyltransferases conduct glycosylation of flavonoids in plants. Glycosylation, one of the most essential and common modifications, is the process by which a sugar unit is transferred from an activated nucleotide diphosphate sugar, such as UDP-glucose, to a flavonoid molecule. Glycosylation increases the chemical complexity and diversity of natural products, and introduces changes in physical and biological properties of flavonoids, such as stabilizing the formed flavonoids or facilitating the regulation of cellular metabolism (Weymouth-Wilson, 1997).

The flavonoid glycotransferases belong to the GT1 family (Lairson et al., 2008). The most common glycosylation sites in flavonoids are 3-hydroxyl and 7-hydroxyl, although some glycosylation occurs to the 5-hydroxyl in A ring, and the hydroxyl moieties in B ring. The sugar units for monosaccharide generation include glucose, rhamnose, galactose, arabinose, and xylose. The sugar units for disaccharides are rutinose, sophorose, and sambubiose. Some studies have shown the

biotransformation of flavonoid glycosides in recombinant *E. coli* or other microbial cells (Koirala et al., 2014; Pandey et al., 2016; Ruprecht et al., 2019; Shrestha et al., 2018; Wang et al., 2016a). In addition, several nucleotide sugar biosynthetic pathways are engineered in *E. coli* for regioselective biotransformation and high titer production of flavonoid glycosides (Leonard et al., 2008; Yan et al., 2008).

2.8. Prenyltransferases

Plant flavonoid prenyltransferases contain signal peptides at the N-terminus and several transmembrane α -helices to locate them to plastids (Stec and Li, 2012). They share common conserved aspartate-rich motifs NQxxDxxxD and KDxxDxxGD, with the former being involved in binding prenyl donor and the latter in binding flavonoids (Stec and Li, 2012). The prenyl moiety increases the lipophilicity of flavonoids, thus changing their membrane-permeation capabilities (Yang et al., 2015). Most of the prenylated flavonoids are produced by biotransformation using plant and fungal whole cells. The enzymes that can catalyze the formation of

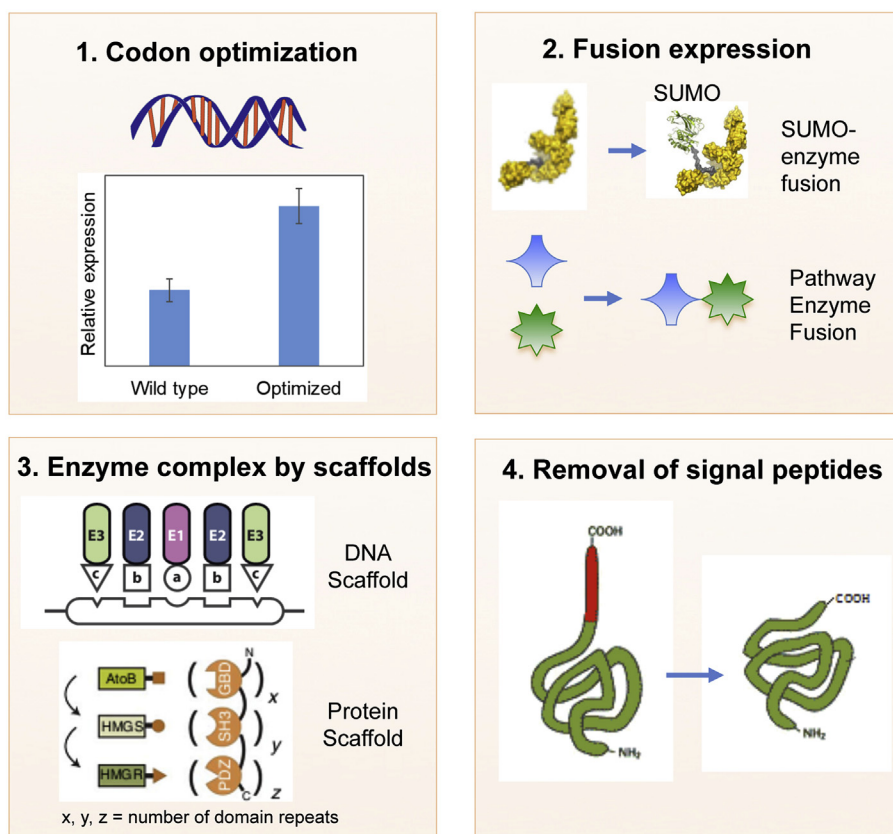
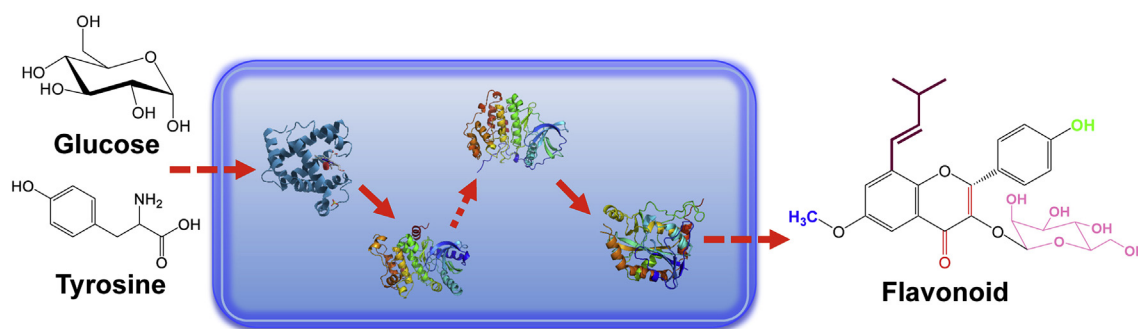


Fig. 3. The approaches frequently used in engineering of flavonoid pathway enzymes for improved functions in recombinant microorganisms. The strategies include codon optimization, fusion expression, enzyme scaffolding (Conrado et al., 2011; Dueber et al., 2009), and removal of signal peptides for expression of P450 pathway enzymes.

prenylated flavonoids are very limited. APT(*Orf2*) from *Streptomyces* was identified to transfer the geranyl group to various flavonoid compounds such as daidzein, formononetin, genistein, fisetin, naringenin, and resveratrol (Kuzuyama et al., 2005). The prenyltransferase NovQ from *S. niveus* was discovered to transfer dimethylallyl group to phenylpropanoids and B ring of flavonoids (Ozaki et al., 2009). A naringenin 8-dimethylallyltransferase (SfN8DT-1) from *Sophora flavescens* was expressed in yeast strain W303-1A and the recombinant yeast could produce 566 µg/L of 8-dimethylallyl naringenin from exogenously added 0.2 mM naringenin (Sasaki et al., 2009). In future studies, improving the supply of dimethylallyl pyrophosphate by engineering native pathway and engineering prenyltransferases can allow the efficient production of diverse flavonoids using recombinant microbes.

3. Engineering strategies applied in improving enzyme functionality

3.1. Codon optimization for better expression

Codon optimization is a commonly used strategy to improve the expression of pathway enzymes (Lanza et al., 2014). Because of the codon bias between the native plant hosts and engineered microbial hosts, expression of plant flavonoid pathway genes in microbes is often difficult. As a result, codon optimization is generally employed according to the preference of codon usage in the microbial species to be used (Fig. 3). In *E. coli*, the expression of flavonoid pathway genes without codon optimization is usually successful based on our experience in producing diverse flavonoids for two decades. However, their expression in other microorganisms could be problematic. In *C. glutamicum*-based naringenin production, expression of the non-codon-optimized pathway genes was unsuccessful (Kallscheuer et al., 2016); however, codon-optimization of these genes led to acceptable levels of naringenin biosynthesis. Due to the importance of codon-optimization, it is usually performed at the very beginning of the engineering procedure in microbial production of flavonoids.

3.2. Fusion engineering for improved enzyme expression and activity

Codon-optimization only aims at making the gene sequences recognizable by the microbial hosts; however, it does not guarantee that the genes/enzymes can be expressed correctly. In some cases, the expression of flavonoid pathway enzymes is confronted with enzyme misfolding and the formation of inclusion bodies, so the soluble expression of these enzymes is of particular significance. A frequently used approach to increase the soluble expression of enzymes is fusion of a protein or peptide tag, which on its own is highly soluble in the host strain even at a very high expression level, such as maltose-binding protein (MBP) and small ubiquitin-like modifier (SUMO) (Fig. 3). For example, to engineer *C. glutamicum* for the biosynthesis of cyanidin 3-*O*-glucoside from catechin, SUMO tag was fused to the N-terminus of flavonoid 3-*O*-glucosyltransferase to improve its expression. The fusion enzyme allowed a 1.1-fold increase in the production of cyanidin 3-*O*-glucoside (Zha et al., 2018).

Beyond optimization of a single pathway enzyme for improved expression and hence increased enzyme activity, translational fusion of multiple enzymes in successive steps is another effective method of improving enzyme performance (Fig. 3). Such fusions can maximize the local concentrations of substrates for each enzyme in the fusion system, allowing multiple reactions to proceed efficiently while minimizing the degradation of unstable intermediates or reducing the negative effects of toxic intermediates (Springob et al., 2003). Using this strategy, it has been shown in *E. coli* that the translational fusion of flavonoid 3-*O*-glucosyltransferase from *Arabidopsis* to the N-terminus of anthocyanidin synthase from *Petunia* could allow better conversion of catechin to cyanidin 3-*O*-glucoside compared with the tandem expression of the two enzymes, and the fused enzyme complex could catalyze the successive

biochemical reactions 16.9% more efficiently than the uncoupled enzymes (Yan et al., 2008).

3.3. Assembly of pathway enzymes by scaffolds

Channeling metabolic intermediates efficiently to improve the performance of the introduced pathway is a fundamental strategy in pathway optimization (Wu et al., 2016). Fusion of pathway enzymes is a common and useful strategy, but very often such direct fusion causes misfolding and loss of function. Alternatively, DNA or protein scaffolding provides a way of assembling pathway enzymes with enhanced catalytic activities. To form a DNA scaffold, first, a specific DNA-binding protein (such as zinc finger domains) is fused to each enzyme at the genetic level (different enzymes need to be fused with different DNA-binding proteins). Then, the specific DNA sequences that interact with the DNA-binding proteins are placed on a plasmid and their copies can be varied to simply regulate the composition of DNA-protein complex (the DNA scaffold). Using such an approach, a threefold increase of resveratrol production was achieved in *E. coli* expressing 4CL and stilbene synthase with the supplementation of 4-coumarate (Conrado et al., 2011).

Beyond DNA scaffolds, protein scaffolds are also very useful in the regulation of enzyme performance. By adding a specific peptide ligand to the C-terminus of each enzyme in the mevalonate biosynthetic pathway, including AtoB, HMGS, and HMGR, the pathway enzymes formed an enzyme complex through the binding between ligands and the three cognate domains, which were tethered with glycine-serine linkers to provide a protein scaffold (Fig. 3). The copies of the separate domains could be varied to regulate the composition of the enzyme complex to achieve highest mevalonate production (Dueber et al., 2009). The modulation achieved a 77-fold increase in mevalonate production compared to the unscaffolded pathway. Zhao et al applied this method to the production of catechin from eriodictyol and observed only slight improvement of the product titer (Zhao et al., 2015). The potential issue with such protein scaffolds lies in the rigidity of the protein domains, which could negatively affect the catalytic process.

3.4. Removal of the signal peptide to improve expression

Heterologous expression of some flavonoid pathway enzymes, such as hydroxylases, in prokaryotes is sometimes challenging, and typically, the genes need to be modified prior to their functional expression. Most hydroxylases in natural flavonoid biosynthetic pathways are P450 enzymes, and their catalytic process requires the involvement of P450 reductases. Both P450 enzymes and P450 reductases contain signal peptides to locate the enzymes to membranes of eukaryotic organelles. However, prokaryotes lack such organelles and the cells cannot recognize the signal peptides with plant origin, causing the misfolding and malfunctioning of the expressed enzymes. Thus, the elimination of signal peptides in P450 hydroxylases and P450 reductases is generally required (Fig. 3). For example, the first four codons of P450 F3'5'H from *Catharanthus roseus* were removed to achieve its functional expression, and the fifth codon was replaced with ATG as the new start codon, while the sixth codon was changed from leucine to alanine (Leonard et al., 2006). The new F3'5'H was fused with a truncated P450 reductase from *C. roseus* to form a chimeric protein, which resulted in the successful formation of quercetin by feeding coumaric acid. In another case, the truncated version of flavone 6-hydroxylase from *Scutellaria baicalensis* showed a 10-fold higher activity than the full-length version in the production of baicalein in *E. coli*, and fusion of the truncated enzyme with an expression tag elevated the activity by another 2-fold (Li et al., 2019).

Although removal of signal peptides can improve the expression and activity of P450 enzymes, its effect is still very limited in the sense that the engineered enzymes heterologously expressed in microorganisms are usually not as active as the natural enzymes produced in plants (Ajikumar et al., 2010). A possible reason is that electron transfer performed by the truncated P450 hydroxylases is not as efficient as that by the native

enzymes anchored in membranes of eukaryotic organelles (Ortiz de Montellano, 2010). Optimization of the molar ratio and the relative distance between the P450 hydroxylase and the P450 reductase can greatly solve this problem (Li et al., 2019); nonetheless, this process is largely trial-and-error and time consuming. A complete understanding of the electron transfer process in the truncated enzyme complex and reconstruction of an efficient channel for electron transfer will give rise to more efficient and soluble hydroxylases for application in prokaryotic strains.

4. Conclusions and future perspectives

Metabolic engineering of pathway enzymes for microbial production of flavonoids proves to be a promising way of supplying flavonoid compounds for industrial applications. It also allows the biosynthesis of rare flavonoid compounds, such as pyranoanthocyanins (Akdemir et al., 2019; Solopova et al., 2019). With the understanding of pathway enzymes, researchers have successfully reconstituted the biosynthetic pathways of various flavonoid compounds in *E. coli*, *S. cerevisiae*, *S. venezuelae*, and *C. glutamicum*. Production of some flavonoids with the supplementation of precursors has reached production titers of over 1 g/L (Wang et al., 2016a). In this review, we have provided the genetic, biochemical, and expression information on the pathway enzymes and have summarized the engineering approaches for functional enzyme expression and improved enzyme performances.

Given the current achievements in flavonoid biosynthesis in engineered microbes, the production capacity in most of the constructed producing systems still cannot meet the requirements for industrial use. One critical issue is that the detailed expression condition (the expression level and the extent of soluble expression), the stability (enzyme degradation), the location (enzyme diffusion and interaction with other cellular components or compounds), the activity, and the kinetic property of each expressed pathway enzyme in the host cell are not clearly investigated, thus limiting our understanding of the enzyme behavior and further hindering pathway engineering on the enzyme level for better functionality and a higher production yield and titer.

The expression levels of pathway enzymes can be investigated by proteomic tools, such as iTRAQ. Coupled with the metabolomics analysis, bottleneck pathway enzymes can be identified and engineered to improve the flavonoid pathway function. As for the location of these enzymes, labeling technologies such as fusion of fluorescent proteins combined with fluorescence microscopy can provide a powerful tool to track the expression and diffusion over the whole catalytic process. It is thus anticipated that the issues with the expressed enzymes will be addressed with the available systems biology tools, biochemical analysis techniques, and enzyme engineering techniques. The big stride in understanding of the enzyme reactions inside cells will drive the reconstruction of commercially available microbial cell factories of diverse flavonoids for applications in industry in the near future.

Acknowledgements

J.Z. and X.W. would like to thank the startup funds from Shaanxi University of Science and Technology.

References

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

Akdemir, H., Silva, A., Zha, J., Zagorevski, D.V., Koffas, M.A.G., 2019. Production of pyranoanthocyanins using *Escherichia coli* co-cultures. *Metab. Eng.* 55, 290–298.

Austin, M.B., Bowman, M.E., Ferrer, J.-L., Schröder, J., Noel, J.P., 2004. An aldol switch discovered in stilbene synthases mediates cyclization specificity of type III polyketide synthases. *Chem. Biol.* 11, 1179–1194.

Chemler, J.A., Koffas, M.A., 2008. Metabolic engineering for plant natural product biosynthesis in microbes. *Curr. Opin. Biotechnol.* 19, 597–605.

Cheng, H., Li, L., Cheng, S., Cao, F., Wang, Y., Yuan, H., 2011. Molecular cloning and function assay of a chalcone isomerase gene (GbCHI) from *Ginkgo biloba*. *Plant Cell Rep.* 30, 49–62.

Cho, J.-H., Lim, Y., Ahn, J.-H., Rhee, S., 2007. Crystallization and preliminary X-ray diffraction analysis of BcOMT2 from *Bacillus cereus*: a family of O-methyltransferase. *J. Microbiol. Biotechnol.* 17, 369–372.

Chouhan, S., Sharma, K., Zha, J., Guleria, S., Koffas, M.A.G., 2017. Recent advances in the recombinant biosynthesis of polyphenols. *Front. Microbiol.* 8, 2259.

Conrado, R.J., Wu, G.C., Boock, J.T., Xu, H., Chen, S.Y., Lebar, T., Turnšek, J., Tomšič, N., Avbelj, M., Gaber, R., Koprivnjak, T., Mori, J., Glavnik, V., Vovk, I., Bencina, M., Hodnik, V., Anderlüh, G., Dueber, J.E., Jerala, R., DeLisa, M.P., 2011. DNA-guided assembly of biosynthetic pathways promotes improved catalytic efficiency. *Nucleic Acids Res.* 40, 1879–1889.

Cress, B.F., Leitz, Q.D., Kim, D.C., Amore, T.D., Suzuki, J.Y., Linhardt, R.J., Koffas, M.A.G., 2017. CRISPRi-mediated metabolic engineering of *E. coli* for O-methylated anthocyanin production. *Microb. Cell Factories* 16, 10.

Cukovic, D., Ehrling, J., Ziffle Jessica, A.V., Douglas Carl, J., 2001. Structure and evolution of 4-coumarate:Coenzyme A ligase (4CL) gene families. *Biol. Chem.* 382, 645.

Dao, T.T.H., Linthorst, H.J.M., Verpoorte, R., 2011. Chalcone synthase and its functions in plant resistance. *Phytochem. Rev.* 10, 397.

Douglas, C.J., 1996. Phenylpropanoid metabolism and lignin biosynthesis: from weeds to trees. *Trends Plant Sci.* 1, 171–178.

Dueber, J.E., Wu, G.C., Malmirchegini, G.R., Moon, T.S., Petzold, C.J., Ullal, A.V., Prather, K.L., Keasling, J.D., 2009. Synthetic protein scaffolds provide modular control over metabolic flux. *Nat. Biotechnol.* 27, 753–759.

Fowler, Z.L., Shah, K., Panepinto, J.C., Koffas, M.A.G., 2011. Development of non-natural flavanones as antimicrobial agents. *PLoS One* 6, e25681.

Jensen, K., Moller, B.L., 2010. Plant NADPH-cytochrome P450 oxidoreductases. *Phytochemistry* 71, 132–141.

Jez, J.M., Noel, J.P., 2002. Reaction mechanism of chalcone isomerase: pH dependence, diffusion control, and product binding differences. *J. Biol. Chem.* 277, 1361–1369.

Jones, J.A., Collins, S.M., Vernacchio, V.R., Lachance, D.M., Koffas, M.A.G., 2016. Optimization of naringenin and p-coumaric acid hydroxylation using the native *E. coli* hydroxylase complex. *HpaBC. Biotechnol. Prog.* 32, 21–25.

Kallscheuer, N., Vogt, M., Bott, M., Marienhagen, J., 2017. Functional expression of plant-derived O-methyltransferase, flavanone 3-hydroxylase, and flavonol synthase in *Corynebacterium glutamicum* for production of pterostilbene, kaempferol, and quercetin. *J. Biotechnol.* 258, 190–196.

Kallscheuer, N., Vogt, M., Stenzel, A., Gatgens, J., Bott, M., Marienhagen, J., 2016. Construction of a *Corynebacterium glutamicum* platform strain for the production of stilbenes and (2S)-flavanones. *Metab. Eng.* 38, 47–55.

Kim, D.H., Kim, B.-G., Lee, Y., Ryu, J.Y., Lim, Y., Hur, H.-G., Ahn, J.-H., 2005. Regiospecific methylation of naringenin to poncirin by soybean O-methyltransferase expressed in *Escherichia coli*. *J. Biotechnol.* 119, 155–162.

Koirala, N., Pandey, R.P., Van Thang, D., Jung, H.J., Sohng, J.K., 2014. Glycosylation and subsequent malonylation of isoflavonoids in *E. coli*: strain development, production and insights into future metabolic perspectives. *J. Ind. Microbiol. Biotechnol.* 41, 1647–1658.

Koirala, N., Thuan, N.H., Ghimire, G.P., Thang, D.V., Sohng, J.K., 2016. Methylation of flavonoids: chemical structures, bioactivities, progress and perspectives for biotechnological production. *Enzym. Microb. Technol.* 86, 103–116.

Kuzuyama, T., Noel, J.P., Richard, S.B., 2005. Structural basis for the promiscuous biosynthetic prenylation of aromatic natural products. *Nature* 435, 983–987.

Lairson, L.L., Henrissat, B., Davies, G.J., Withers, S.G., 2008. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* 77, 521–555.

Lanza, A.M., Curran, K.A., Rey, L.G., Alper, H.S., 2014. A condition-specific codon optimization approach for improved heterologous gene expression in *Saccharomyces cerevisiae*. *BMC Syst. Biol.* 8, 33.

Lavhale, S.G., Kalunke, R.M., Giri, A.P., 2018. Structural, functional and evolutionary diversity of 4-coumarate-CoA ligase in plants. *Planta* 248, 1063–1078.

Leonard, E., Yan, Y., Fowler, Z.L., Li, Z., Lim, C.G., Lim, K.H., Koffas, M.A.G., 2008. Strain improvement of recombinant *Escherichia coli* for efficient production of plant flavonoids. *Mol. Pharm.* 5, 257–265.

Leonard, E., Yan, Y., Koffas, M.A., 2006. Functional expression of a P450 flavonoid hydroxylase for the biosynthesis of plant-specific hydroxylated flavonols in *Escherichia coli*. *Metab. Eng.* 8, 172–181.

Li, F.-X., Jin, Z.-P., Zhao, D.-X., Cheng, L.-Q., Fu, C.-X., Ma, F., 2006. Overexpression of the *Saussurea medusa* chalcone isomerase gene in *S. involucrata* hairy root cultures enhances their biosynthesis of apigenin. *Phytochemistry* 67, 553–560.

Li, J., Tian, C., Xia, Y., Mutanda, I., Wang, K., Wang, Y., 2019. Production of plant-specific flavones baicalein and scutellarein in an engineered *E. coli* from available phenylalanine and tyrosine. *Metab. Eng.* 52, 124–133.

Li, Z., Nair, Satish K., 2015. Structural basis for specificity and flexibility in a plant 4-coumarate:CoA ligase. *Structure* 23, 2032–2042.

Liu, B., Falkenstein-Paul, H., Schmidt, W., Beerhues, L., 2003. Benzophenone synthase and chalcone synthase from *Hypericum androsaemum* cell cultures: cDNA cloning, functional expression, and site-directed mutagenesis of two polyketide synthases. *Plant J.* 34, 847–855.

Louie, G.V., Bowman, M.E., Moffitt, Michelle C., Baiga, T.J., Moore, Bradley S., Noel, J.P., 2006. Structural determinants and modulation of substrate specificity in phenylalanine-tyrosine ammonia-lyases. *Chem. Biol.* 13, 1327–1338.

Mavandad, M., Edwards, R., Liang, X., Lamb, C.J., Dixon, R.A., 1990. Effects of trans-cinnamic acid on expression of the bean phenylalanine ammonia-lyase gene family. *Plant Physiol.* 94, 671.

- Ortiz de Montellano, P.R., 2010. Hydrocarbon hydroxylation by cytochrome P450 enzymes. *Chem. Rev.* 110, 932–948.
- Ozaki, T., Mishima, S., Nishiyama, M., Kuzuyama, T., 2009. NovQ is a prenyltransferase capable of catalyzing the addition of a dimethylallyl group to both phenylpropanoids and flavonoids. *J. Antibiot.* 62, 385.
- Pandey, R.P., Parajuli, P., Koffas, M.A., Sohng, J.K., 2016. Microbial production of natural and non-natural flavonoids: pathway engineering, directed evolution and systems/synthetic biology. *Biotechnol. Adv.* 34, 634–662.
- Ruprecht, C., Bönisch, F., Ilmberger, N., Heyer, T.V., Haupt, E.T.K., Streit, W.R., Rabausch, U., 2019. High level production of flavonoid rhamnosides by metagenome-derived glycosyltransferase C in *Escherichia coli* utilizing dextrans of starch as a single carbon source. *Metab. Eng.* 55, 212–219.
- Santos, C.N.S., Koffas, M.A.G., Stephanopoulos, G., 2011. Optimization of a heterologous pathway for the production of flavonoids from glucose. *Metab. Eng.* 13, 392–400.
- Sasaki, K., Tsurumaru, Y., Yazaki, K., 2009. Prenylation of flavonoids by biotransformation of yeast expressing plant membrane-bound prenyltransferase SfN8DT-1. *Biosci. Biotechnol. Biochem.* 73, 759–761.
- Schroeder, A.C., Kumaran, S., Hicks, L.M., Cahoon, R.E., Halls, C., Yu, O., Jez, J.M., 2008. Contributions of conserved serine and tyrosine residues to catalysis, ligand binding, and cofactor processing in the active site of tyrosine ammonia lyase. *Phytochemistry* 69, 1496–1506.
- Shimada, N., Aoki, T., Sato, S., Nakamura, Y., Tabata, S., Ayabe, S.-I., 2003. A cluster of genes encodes the two types of chalcone isomerase involved in the biosynthesis of general flavonoids and legume-specific 5-deoxy(iso)flavonoids in *Lotus japonicus*. *Plant Physiol.* 131, 941–951.
- Shrestha, A., Pandey, R.P., Dhakal, D., Parajuli, P., Sohng, J.K., 2018. Biosynthesis of flavone C-glucosides in engineered *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 102, 1251–1267.
- Solopova, A., van Tilburg, A.Y., Foito, A., Allwood, W.J., Stewart, D., Kulakauskas, S., Kuipers, O.P., 2019. Engineering *Lactococcus lactis* for the production of unusual anthocyanins using tea as substrate. *Metab. Eng.* 54, 160–169.
- Springob, K., Nakajima, J., Yamazaki, M., Saito, K., 2003. Recent advances in the biosynthesis and accumulation of anthocyanins. *Nat. Prod. Rep.* 20, 288–303.
- Stec, E., Li, S.-M., 2012. Mutagenesis and biochemical studies on AuaA confirmed the importance of the two conserved aspartate-rich motifs and suggested difference in the amino acids for substrate binding in membrane-bound prenyltransferases. *Arch. Microbiol.* 194, 589–595.
- Sun, H., Li, Y., Feng, S., Zou, W., Guo, K., Fan, C., Si, S., Peng, L., 2013. Analysis of five rice 4-coumarate:coenzyme A ligase enzyme activity and stress response for potential roles in lignin and flavonoid biosynthesis in rice. *Biochem. Biophys. Res. Commun.* 430, 1151–1156.
- Trantas, E., Panopoulos, N., Ververidis, F., 2009. Metabolic engineering of the complete pathway leading to heterologous biosynthesis of various flavonoids and stilbenoids in *Saccharomyces cerevisiae*. *Metab. Eng.* 11, 355–366.
- Treutter, D., 2006. Significance of flavonoids in plant resistance: a review. *Environ. Chem. Lett.* 4, 147.
- Tsai, P.-J., Wu, S.-C., Cheng, Y.-K., 2008. Role of polyphenols in antioxidant capacity of napiergrass from different growing seasons. *Food Chem.* 106, 27–32.
- Vannelli, T., Qi, W.W., Sweigard, J., Gatenby, A.A., Sariaslani, F.S., 2007. Production of *p*-hydroxycinnamic acid from glucose in *Saccharomyces cerevisiae* and *Escherichia coli* by expression of heterologous genes from plants and fungi. *Metab. Eng.* 9, 142–151.
- Wang, H., Hu, T., Huang, J., Lu, X., Huang, B., Zheng, Y., 2013. The expression of *Milletia pinnata* chalcone isomerase in *Saccharomyces cerevisiae* salt-sensitive mutants enhances salt-tolerance. *Int. J. Mol. Sci.* 14, 8775–8786.
- Wang, H., Yang, Y., Lin, L., Zhou, W., Liu, M., Cheng, K., Wang, W., 2016a. Engineering *Saccharomyces cerevisiae* with the deletion of endogenous glucosidases for the production of flavonoid glucosides. *Microb. Cell Factories* 15, 134.
- Wang, J., Guleria, S., Koffas, M.A.G., Yan, Y., 2016b. Microbial production of value-added nutraceuticals. *Curr. Opin. Biotechnol.* 37, 97–104.
- Wang, Q., Yang, J., Zhang, X.-m., Zhou, L., Liao, X.-L., Yang, B., 2015. Practical synthesis of naringenin. *J. Chem. Res.* 455–457.
- Weymouth-Wilson, A.C., 1997. The role of carbohydrates in biologically active natural products. *Nat. Prod. Rep.* 14, 99–110.
- Wu, G., Yan, Q., Jones, J.A., Tang, Y.J., Fong, S.S., Koffas, M.A., 2016. Metabolic burden: cornerstones in synthetic biology and metabolic engineering applications. *Trends Biotechnol.* 34, 652–664.
- Wu, J., Du, G., Chen, J., Zhou, J., 2015. Enhancing flavonoid production by systematically tuning the central metabolic pathways based on a CRISPR interference system in *Escherichia coli*. *Sci. Rep.* 5, 13477.
- Yan, Y., Li, Z., Koffas, M.A.G., 2008. High-yield anthocyanin biosynthesis in engineered *Escherichia coli*. *Biotechnol. Bioeng.* 100, 126–140.
- Yang, X., Jiang, Y., Yang, J., He, J., Sun, J., Chen, F., Zhang, M., Yang, B., 2015. Prenylated flavonoids, promising nutraceuticals with impressive biological activities. *Trends Food Sci. Technol.* 44, 93–104.
- Yin, Y.-C., Zhang, X.-D., Gao, Z.-Q., Hu, T., Liu, Y., 2019. The research progress of chalcone isomerase (CHI) in plants. *Mol. Biotechnol.* 61, 32–52.
- Yu, D., Xu, F., Zeng, J., Zhan, J., 2012. Type III polyketide synthases in natural product biosynthesis. *IUBMB Life* 64, 285–295.
- Zang, Y., Zha, J., Wu, X., Zheng, Z., Ouyang, J., Koffas, M.A.G., 2019. *In vitro* naringenin biosynthesis from *p*-coumaric acid using recombinant enzymes. *J. Agric. Food Chem.* <https://doi.org/10.1021/acs.jafc.9b00413>.
- Zha, J., Koffas, M.A.G., 2017. Production of anthocyanins in metabolically engineered microorganisms: current status and perspectives. *Synth. Syst. Biotechnol.* 2, 259–266.
- Zha, J., Zang, Y., Mattozzi, M., Plassmeier, J., Gupta, M., Wu, X., Clarkson, S., Koffas, M.A.G., 2018. Metabolic engineering of *Corynebacterium glutamicum* for anthocyanin production. *Microb. Cell Factories* 17, 143.
- Zhang, Z., He, Y., Huang, Y., Ding, L., Chen, L., Liu, Y., Nie, Y., Zhang, X., 2018. Development and optimization of an in vitro multienzyme synthetic system for production of kaempferol from naringenin. *J. Agric. Food Chem.* 66, 8272–8279.
- Zhao, S., Jones, J.A., Lachance, D.M., Bhan, N., Khalidi, O., Venkataraman, S., Wang, Z., Koffas, M.A., 2015. Improvement of catechin production in *Escherichia coli* through combinatorial metabolic engineering. *Metab. Eng.* 28, 43–53.
- Zhu, L., Huang, Y., Zhang, Y., Xu, C., Lu, J., Wang, Y., 2017. The growing season impacts the accumulation and composition of flavonoids in grape skins in two-crop-a-year viticulture. *J. Food Sci. Tech. MYS.* 54, 2861–2870.