Discovery and modification of cytochrome P450 for plant natural products biosynthesis

Xiaonan Liua,b,1, Xiaoxi Zhua,b,1, Hui Wangc, Tian Liud,a, Jian Chenga,∗∗, Huifeng Jianga,∗

a Key Laboratory of Systems Microbial Biotechnology, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin, 300308, China
b University of Chinese Academy of Sciences, Beijing, 100049, China
c College of Biotechnology, Tianjin University of Science & Technology, Tianjin, 300457, China
d State Key Laboratory for Conservation and Utilization of Subtropical Agro-bioresources, College of Life Science and Technology, Guangxi University, Nanning, Guangxi, 530004, China

ARTICLE INFO
Keywords: P450 enzymes Plant natural products Synthetic biology

ABSTRACT
Cytochrome P450s are widespread in nature and play key roles in the diversification and functional modification of plant natural products. Over the last few years, there has been remarkable progress in plant P450s identification with the rapid development of sequencing technology, “omics” analysis and synthetic biology. However, challenges still persist in respect of crystal structure, heterologous expression and enzyme engineering. Here, we reviewed several research hotspots of P450 enzymes involved in the biosynthesis of plant natural products, including P450 databases, gene mining, heterologous expression and protein engineering.

Backgrounds
Plant cytochrome (CYP) P450 gene family is the largest gene family in plants [1]. P450 enzymes (P450s) are widely involved in biosynthetic pathway of plant natural products (PNPs) due to the wide range of activities including hydroxylation, reduction, decarboxylation, sulfation, N- and O-demethylation, epoxidation, deamination, and dehalogenation [2]. Since the first P450 gene was debuted to the world in 1958, until now, more than 300,000 P450 genes have been collected in databases [3]. Unfortunately, only about 0.1% of them has been functionally characterized in the biosynthetic pathway of natural products. With the increasing market demand of PNPs, microbial cell factories have been becoming more and more popular to produce PNPs [4]. However, the expression level and catalytic efficiency of P450s are limited in the heterogeneous microbial hosts [5,6]. Therefore, identifying new catalytic functions and enhancing the heterologous activities of plant P450s have been the major challenges in establishing efficient and high-yielding microbial production platforms for complex PNPs [5]. In this review, we summarized the recent progress of P450 enzyme databases, P450s mining, heterologous expression and enzyme engineering (Fig. 1).

Functional and structural databases of P450 enzymes
With the development of modern biotechnologies, more and more P450 genes have been well-characterized including their catalytic functions, crystal structures and the roles in biosynthetic pathway of secondary metabolites. In the near future, the number of P450 genes will reach to more than one million. It is impossible to obtain the desired P450s for scientific researches and industrial applications from so many genes without well-organized information. Therefore it is necessary to construct databases to collect the sequences, structures and functions of P450s.

Functional distribution of the known P450s
The plant P450s were named and classified into CYP71-CYP99 and CYP701-CYP999 gene families according to the standard CYP450 nomenclature sequences in Nelson’s lab [3]. In general, P450s were grouped into a family when sharing more than 40 sequence identities [7]. To date, about 200 plant P450 genes, which are from 55 gene families, have been identified and characterized (Fig. 2). 135 secondary metabolites, including 86 terpenoids, 11 flavonoids, 18 alkaloids were modified by these P450 enzymes and the important reactions catalyzed...
by plant P450s were summarized in Table 1. Among of them, CYP71 family is mainly involved in many sesquiterpenes biosynthesis, including capsidiol, solavetivone, gossypol, artemisinin, parthenolide, zerumbone, caryophyllene and so on. CYP76, CYP701, CYP725, CYP720 and CYP714 families are mainly involved in diterpenes biosynthesis, including tanshinone, oryzalexin, phytocassane, oryzalexin, gibberellin, taxol, abietate [8–11]. Except the CYP725 and CYP720 families, all families in CYP85 clan (CYP716, CYP90, CYP724, CYP87, CYP705, CYP708, CYP85 and CYP88) are involved in triterpenes biosynthesis, including glycyrrhizin, maslinic acid, ginsenoside, betulinate, zanhic acid, brassinolide, diosgenin, cucurbitacin, mogroside, thalianol and so on [12–14]. Besides, CYP51, CYP92, CYP710, CYP72 families also play important roles in many triterpenoid saponins biosynthesis [12,15]. However P450s for monoterpenes biosynthesis are sporadically distributed in different gene families, such as CYP71D13 for menthol biosynthesis, CYP750B1 for thujone biosynthesis and CYP72A1 for strictosidine biosynthesis [16–19].

In addition, many P450 genes associated with the biosynthesis of flavonoids and alkaloids also have been identified. P450 genes in CYP73A subfamily (like cinnamate 4-hydroxylase) and CYP93 family (flavone synthases) help to form the skeleton of flavonoids skeleton. Post-modifications of flavonoids at C6, C8, C3’ and C5’ positions usually are hydroxylated by P450 enzymes [20]. For example, CYP75A and CYP75B (C3’ and C5’ flavonoid hydroxylases), play the key roles in cyanidin biosynthesis. P450 genes in CYP706X and CYP82D subfamilies (C6 and C8 flavonoid hydroxylases), are necessary for the biosynthesis of scutellarin, baicalin, wogonin and so on [4,21]. P450 enzymes from CYP80, CYP82 and CYP719 families usually are required for the biosynthesis of many alkaloids, such as magnoflorine, morphine, sanguinarine and noscapine. For example, CYP80B1 (the (S)–N-methylcocaine 3’-hydroxylase) plays an important role for biosynthesis of the key precursor of (S)-reticuline [22]; CYP719A1 helps to generate the skeleton of isoquinoline alkaloids [23]; P450 genes from CYP82N, CYP82P, CYP82X and CYP82Y subfamilies also contribute a lot on biosynthesis of the final product of alkaloids [23–27].

Construction of P450 gene databases

In 1996, the first P450 database was constructed, which systematically named the existing 481 P450 genes [28]. In 2007, a tool for searching and predicting P450 protein family, CYPED (Cytochrome P450 Engineering Database) was developed [29]. CYPED included 8613 sequences and 47 protein structures, and integrated several tools such as multiple sequence alignment, phylogenetic tree and HMM (Hidden Markov Model) profile, which could be used to analyze sequences, structures, and their relationship to biochemical properties. Subsequently, Nelson’s lab constructed the Cytochrome P450 Homepage [30], which manually organized and collected P450 genes with systematic nomenclature and sequence information. To date, the database have collected 300,000 P450 sequences and 41,000 named P450s [3].
Besides the collection of P450 sequences, a CYP-drug interactive analysis database (SuperCYP) was developed [31], which focused on the metabolic relationship between P450s and drugs. This database contained about 1170 drugs and 2785 CYP-drug interactions.

P450 gene databases had been developed for decades and have contributed a lot on gene mining, annotation and functional analysis. Taking the advantage of genome sequencing technology, more and more P450 genes have been annotated from the genome databases of some model species, such as http://www.p450.kvl.dk/ [32], http://www.arabidopsis.org/ [12] and http://www.phytozome.net/ [33], and also some plant genome sequencing projects such as the One thousand plants (1 KP) database [34], 10 KP project (https://db.cngb.org/10kp/), the JGI Open Green Genomes initiative, the PhytoMetaSyn project (www.phytometasyn.ca) and the Medicinal Plant Genome project [35]. It is necessary to construct more robust database of P450 genes according to the fast development of genome sequencing.

Crystal structure of P450 enzymes

Despite the advances of P450 gene discovery, only a limited number of P450s have been successfully crystallized [36]. The first crystal structure of cytochrome P450 was published in 1985 (CYP101A1) [37]. Until now, 963 P450 crystal structures (X-Ray resolution ≤3 Å) have been deposited in the Protein Data Bank (PDB). However, only five crystal structures of P450 are from plants (https://www.rcsb.org/). The first two crystal structures of P450 in plants are two CYP74s [38,39]. Due to their atypical mechanism, which does not require molecule oxygen and NADPH reductase, it is not suitable for modeling other typical plant P450s based on these two structures. Fortunately, the X-ray crystal structure of CYP76AH1 containing inhibitor 4-phenylimidazole (4-PI), which is a typical membrane-associated cytochrome P450 enzyme and plays a critical role in the biosynthetic pathway of tanshinones, was determined at the resolution of 2.6 Å [36].

In recent years, the Arabidopsis campesterol 22-hydroxylase CYP90B1 involved in brassinosteroids biosynthesis and carotenoid-ring hydroxylase CYP97A3 (PDB Number: 6J94) involved in lutein biosynthesis had been respectively crystallized by Shingo Nagano’s lab and Zhao’s lab [40]. The area of crystal structure analysis of plant P450 still has a long way to go.

Identification of P450 genes from plants

Identifying P450s is the key step to establish the original biosynthetic mechanisms of PNPps. In general, P450s are screened and identified from the plant by genomic or transcriptomic sequencing. For some plants without sequencing data, it is also possible to dig an alternative isozyme from database, which would catalyze the same biological reaction to realize PNPps biosynthesis. In this section, both strategies of P450s discovery will be discussed.
Table 1
Summary of the important reactions catalyzed by plant P450s.

| Reactions          | Substrates          | Products                       | Enzymes                                      | Sources                          |
|--------------------|---------------------|--------------------------------|----------------------------------------------|----------------------------------|
| Hydroxylation      | Apigenin            | Scutellalin                    | Erigeron breviscapus [4]                     |                                  |
| l-Tyrosine         | l-DOPA              |                                 | Beta vulgaris and Mirabilis jalapa [71]      |                                  |
| Taxa-4(5),11(12)-diene | Taxa-4(20),11(12)-dien-5a-ol | Taxa-4(5),11(12)-dien-5a-ol | Taxus species [83]                         |                                  |
| Coronaridine       | 10-Hydroxycoronaridine | Coronaridine                   | Tabernantheboga [84]                        |                                  |
| Dihydrochelerythrine | 10-Hydroxydihydrochelerythrine | Dihydrochelerythrine | California poppy [23]                       |                                  |
| Chrysin            | Baicalein           | Baicalein                      | Scutellaria baicalensis [21]                 |                                  |
| C-2α hydroxylated triterpenoids |                       |                                 | Creasagus pinnatifida and Centella asiatica [48] |                                  |
| Miltiradiene       | Ferruginol          | Ferruginol                     | Salvia miltiorrhiza [85]                     |                                  |
| (−)-5′-Desmethoxysteatin | (−)-5′-Desmethylysteatin |                                 | Podophyllum hexandrum (mayapple) [67]       |                                  |

(continued on next page)
Table 1 (continued)

| Reactions     | Substrates                                                                 | Products                                                                 | Enzymes                                                                 | Sources                        |
|---------------|----------------------------------------------------------------------------|--------------------------------------------------------------------------|-------------------------------------------------------------------------|--------------------------------|
| Decarboxylation | L-Tyrosine (EIZ)                                                          | (EZ) p-hydroxy-phenylacetaldoxime                                         | Taxus baccata [87]                                                       |
|               | L-Valine or L-isoleucine (Z)-2-methylpropanal oxime (Z)-2-methylbutanal oxime |                                                                          |                                                                        | Cassava (Manihot esculenta Crantz) [88] |
| Epoxidation   | β-amyrin                                                                  | 12,13β-epoxy-β-amyrin;                                                   |                                                                        | Oat (Avena strigose) [89,90]    |
|               | Lycosantaleine                                                            | Epoxy-Lycosantaleine                                                     |                                                                        | Solanum lycopersicum [91]       |
| GA12          | 16α, 17-epoxy GA12                                                        |                                                                          |                                                                        | Oryza sativa [92]               |

(continued on next page)
| Reactions                  | Substrates                  | Products                                         | Enzymes                     | Sources                          |
|---------------------------|-----------------------------|--------------------------------------------------|------------------------------|----------------------------------|
| Cyclization               | Tetrahydrocolumbamine       | Canadine                                         | Papaver somniferum [44]      |                                  |
|                           | (-) - mataresinol           | (−) - pluvialolide                                |                               |                                  |
|                           | (S)-Scoulerine              | (S)-Cheilanthiloline                              |                               | California poppy [23]            |
|                           | Obtusifoliol                | 4 alpha-methyl-5 alpha-ergosta-8, 14,24(28)-trien-3 beta-ol | Sorghum bicolor (L) [93]     |                                  |
| N-or O-demethylation      |                            |                                                  | Nicotiana tabacum [94]       |                                  |
|                           | Nicotine                    | Normicotine                                       |                               |                                  |
|                           | (-) - Deoxy-podophyllotoxin | (-) - 4’-Desmethyl-deoxy-podophyllotoxin          | Podophyllum hexandrum (mayapple) [67] |                                  |
|                           | Carbon-carbon double bond   | Narigenin                                         | Chrysanthemum indicum L [95] |                                  |
|                           | 24-epi-Campesterol          | Brassicasterol                                    | Arabidopsis thaliana [96]    |                                  |
Table 1 (continued)

| Reactions          | Substrates                  | Products                        | Enzymes                                      | Sources                           |
|--------------------|-----------------------------|---------------------------------|----------------------------------------------|-----------------------------------|
| Multistep reactions| 7-Deoxyloganetic acid       | Loganic acid and Secologanic acid| Camptotheca acuminata [97]                   |                                   |
|                    | Cucurbitadienol             | 11-hydroxy-24,25-epoxy-cucurbitadienol | Siraitia grosvenorii [98]                     |                                   |
|                    | Miltiradiene                | 8,12-Abietadienoic acid/Dehydroabietic acid | Tripterygium wfordii [43]                     |                                   |
|                    | Cucurbitadienol             | 11C-20H-Cuol                     | Cucumber (Cucumis sativus) [99]               |                                   |

(continued on next page)
| Reactions       | Substrates                              | Products          | Enzymes                                      | Sources                                      |
|-----------------|-----------------------------------------|-------------------|----------------------------------------------|----------------------------------------------|
| Cucurbitadienol | A variety of hydroxylation products     |                   | Bitter melon (Momordica charantia) [100]     |                                              |
| Oleanolic acid  | Gypsogenic acid                         |                   | Centella asiatica [46]                       |                                              |
| Lupeol          | Betulinic acid (BA)                     |                   | Rosmarinus officinalis (rosemary) [101]      |                                              |
**P450 enzyme discovery by next-generation sequencing**

Genome sequencing not only provides the most comprehensive genetic information for P450 gene mining, which guarantees to contain the desired P450s, but also provides more opportunities to study the mechanisms of PNPs biosynthesis by comparative genomic analysis. For example, comparative genomic analysis of cucumber (*Cucumis sativus*) uncovered several P450s which involved in cucurbitacins biosynthesis, such as CYP88L2 for cucurbitadienol C19 hydroxylation, CYP81Q58 for C25 hydroxylation [41], CYP87D20 for C20 hydroxylation [42]. A whole-genome sequencing of *Erigeron breviscapus* led to the discovery of P450 enzyme in breviscapine biosynthesis with a strategy of combining evolutionary genomics and synthetic biology [4]. Two P450 genes CYP82P2 and CYP82P3 were confirmed to function as 10-hydroxylase in dihydrobenzophenanthridine alkaloid biosynthesis by using a draft genome of *California poppy* [23]. The reference-grade genome of *Triporygium wilfordii* and the integrated genomic, transcriptomic, and metabolomics analysis resulted in the identification of CYP728B70, which catalyzes three steps of oxidation of a methyl to the acid moity of dehydroabietic acid in trilobide biosynthesis [43].

Comparative transcriptomic analysis could also be used to identify candidate P450 enzymes by studying expression profiles between the producing and non-producing plant species or high-yield and low-yield plant varieties or even different tissues in the same plant. For example, a 10-genie cluster involved in noscapine biosynthesis was discovered through comparative transcriptomic profiling of noscapine-producing and non-producing poppy varieties [44]. *De novo* assembly of tissue-specific transcriptomes of *Paris polyphylla* and *Trigonella foenum-graecum* demonstrated that both species employ independently P450 genes to catalyze oxidative 5,6-spiroketalization of cholesterol to yield diosgenin [45]. Based on identification of unique sequences in the transcriptome of *Centella asiatica* and heterologous expression in yeast cells, Kim et al. discovered a novel multifunctional C-23 oxidase, CYP714E19, which involves in asiaticoside biosynthesis [46].

**Identification of P450 isozymes from heterologous sources**

Although the rapid development of next-generation sequencing, only several hundreds of plant genomes could be acquired in NCBI (National Center for Biotechnology Information) database until now. More than ten thousands of genomes and millions of genes from other species have been deposited in NCBI. Hence, an alternative strategy of identifying the isozymes or alternative enzymes from heterologous sources would be possible to realize the biosynthesis of some important PNPs without genomic sequencing data. One of the most successful cases was the complete biosynthesis of opioids in the yeast by combining with the successful expression of 21 enzymes derived from multiple species [47]. In addition, P450 isoforms could also be identified from other plants, where the same PNP biosynthetic pathway independently evolved. For example, *Erigeron breviscapus* and *Scutellaria baicalensis* Georgi independently evolved a flavonoid 6-hydroxylase to catalyze apigenin into scutellarein from P450 subfamily CYP706X and CYP82D, respectively [4,21]. Both of a monocot medicinal plant Himalayan parsley and a eudicot culinary herb plant fenugreek independently recruited a pairs of P450s to convert the oxidative 5,6-spiroketalization of cholesterol into diosgenin [45]. It was reported that both *Craetaegus pinnatifida* and *Centella asiatica* had C-2α hydroxylation enzyme. Interestingly, the C-2α hydroxylation enzyme from *C. asiatica* could produce 233% higher product than that from *C. pinnatifida* in yeast cell factories [48].

**Optimization of P450 genes expression**

Nowadays, with the rapid development of synthetic biology, researchers are making efforts to use various biological systems to functionally express the P450 genes and produce plant-derived natural molecules with complex structure and high-valued. In this part, P450-related metabolic engineering works in yeast, *E. coli* or tobacco cells will be discussed in detail.

**Heterologous expression of P450 in yeast**

Due to the excellent eukaryotic properties, varieties of P450s have been heterologously expressed in the yeast for the biosynthesis of many PNPs, such as artemisic acid [49], triptolide [43], ganoderic acid [50], maslinic [48], various plant triterpenoids [51], vindoline [52], stricostidine [53], tropane alkaloids [54] have been successfully produced in yeast cell factories. Several means for optimization of P450s expression in yeast was established including optimization of P450 enzyme expression, enhancing the coupling between P450s and the cytochrome P450 reductase, increase of reduction power supply, expansion of the endoplasmic reticulum (ER) and so on. For example, to overcome the poor coupling between *Panax ginseng* P450 enzyme PPDS (protopanaxadiol synthetase) and the *Arabidopsis thaliana* cytochrome P450 reductase ATR1, the transmembrane domain truncation and self-sufficient PPDS-ATR1 fusion were constructed, which resulted in approximately 4.5-fold improvement of P450 catalytic activity [55]. In another study, the naringenin to eriodictyol and taxifolin conversion ratio reached 90.5% and 56.8% by screening three CPRs (cytochrome reductases) and optimizing the ratio of F3H—FPP to 1:2 [56]. Through increasing precursor supply, engineering rate-limiting pathway enzymes, optimizing enzyme expression levels and introducing modifications to the endogenous yeast metabolism to enhance NADPH supply, the biosynthesis of benzyl isouquinoline alkaloids in the yeast had been improved over 18,000-fold [47,57,58]. The disruption of the phosphatidic acid phosphatase (PAH1) resulted in a dramatic expansion of the ER, which led to the 8-fold increase of medicagenic acid by co-expressing three P450s [59]. Furthermore, over-expressing a key ER size regulatory factor (INO2) also improved the production of squalene and P450-mediated protopanaxadiol by 71-fold and 8-fold [60].

**Heterologous expression of P450 in E. coli**

Although the intrinsic P450 expression system is default, there were still many successful cases for metabolic engineering of P450 enzyme-mediated biosynthesis of PNPs in *E. coli*. The significant advances of P450 expression in bacterial expression systems, including N-terminal modifications, chaperonin coexpression, reduction of secondary mRNA structure, bacterial codon usage, selection of vector and host strain, as well as varying external growth conditions, have been comprehensively summarized by Zelasko et al. [61]. The N-terminal modifications of P450 enzymes for biosynthesis of PNPs were listed in Table 2. By optimizing P450 expression, reductase partner interactions, and N-terminal modifications, the highest titer of oxygenated taxanes achieved to $\sim 570 \pm 45$ mg/L in *E. coli* [62]. It has been well documented that cellular redox balance and co-factor availability could significantly influence the yield of metabolites in microorganisms. For example, the disruption of glutamate dehydrogenase (gdhA) could increase 10% of NADPH/NADP$^+$ ratio and also improved the product ent-kaurenic acid about 22.5% [63]. The expression of a fusion protein, containing P450 gene CYP714A2 and its reductase AtCPR2 from *A. thaliana*, improved the yield of steveriol to 38.4 mg/L in batch fermentation [64]. Co-culture both *E. coli* and *S. cerevisiae* would be a good strategy to adequately take advantages of both species. For example, benzyloquinoline alkaloids (BIs) were successfully synthesized in an *E. coli* and *S. cerevisiae* co-culture system. *E. coli* was engineered for biosynthesis of the branchpoint intermediate (S)-reticuline, and *S. cerevisiae* co-culture system. *E. coli* was engineered for biosynthesis of the branchpoint intermediate (S)-reticuline, and *S. cerevisiae* co-culture system. *E. coli* was engineered for biosynthesis of the branchpoint intermediate (S)-reticuline, and *S. cerevisiae* co-culture system.
coli and the efficient expression of P450 gene to hydroxylate taxadiene in S. cerevisiae, the synthetic consortium produced the highest yield of oxygenated taxanes with 33 mg/L [66].

**Heterologous expression of P450 genes in tobacco**

Nicotiana benthamiana, a wild relative of tobacco, has been increasingly used as a system for reconstituting biosynthetic pathway of PNPs [67,68]. N. benthamiana is able to be operated under Agrobacterium mediated transient expression which enables expression of multi-gene pathways or combinatorial biosynthetic libraries, without building multi-gene constructs. Moreover, many metabolic precursors for PNPs biosynthesis often could be present with sufficient quantities in N. benthamiana [67–69]. So far, several pathways for different types of natural products have been successfully reconstituted in N. benthamiana, including terpenoid [70], alkaloids [68], etoposine aglycone [71] or betalains [71]. For example, using a chloroplast compartmentalized metabolic engineering strategy, Yong Wang's team introduced taxadiene synthase, taxadiene-5a-hydroxylase, and a cytochrome P450 reductase into tabacco and demonstrated that tabacco has the potential as an alternative platform for taxol production [69]. Jing-Ke Weng's team revealed that instantaneous expression of P450s in tabacco is sufficient to elicit diosigenin production in N. benthamiana [45].

**P450 protein engineering**

Although P450 enzymes possess powerful catalytic function with irreplaceable regioselectivity and stereoselectivity, the natural P450 enzymes are still very limited because of low activity, poor stability and narrow substrate scope, which could not satisfy the requirements of rapid development of industrial biotechnology. The natural P450s have been engineered by many approaches such as directed evolution, random mutation, rational design and so on [72,73]. Here, we will discussed the two most useful methods for protein engineering of P450s.

A lot of efforts has been put into protein engineering, which not only improved the activity and stability of P450s, but also performed novel reactions [73], which was pioneered by the Nobel Laureate Frances H. Arnold [74–76]. Directed evolution has been widely applied to engineer the well-characterized P450s, such as P450BM3 and P450cam [77]. These extensive works of directed evolution and site-directed mutagenesis have been summarized up by Whitehouse et al. [78] and Li et al. [73]. However, there is very few research being done on plant derived P450s. Brühmann et al. applied directed evolution technology to engineer a 13-hydroperoxide lyase (CYP74B) from guava. A mutant with 15-fold increased product yield was generated by an iterative cycle of random mutagenesis. Establishment of a high-throughput screening method is very important for P450s. Brühmann et al. applied directed evolution technology to engineer a 13-hydroperoxide lyase (CYP74B) from guava. A mutant with 15-fold increased product yield was generated by an iterative cycle of in vitro DNA recombination using L-ShufflingTM and rational random mutagenesis using EvoSightTM [79]. By engineering the SR5s (Substrate Recognition Sites) of CYP76AH15, which is responsible for the first putative oxygenation step of the forskolin pathway, about 6 fold improvement of catalytic activity was observed with the variant A99I [80]. Although many successful examples of P450s engineering have been done, it is still restricted by high-throughput screening method. Establishment of a high-throughput screening method is very important for P450 enzyme engineering. A fluorescence-based screen was established for CYP76AD1, which exhibits tyrosine hydroxylase activity based on the further conversion of the product levodopa (L-DOPA) to...
betaxanthin by a plant DOPA dioxygenase. However, the CYP76AD1 mutants with higher activity of tyrosine hydroxylase also conducted DOPA oxidative activity, which could branch the metabolic flux toward undesired synthesis of L-dopaquinone [81]. Depending on a C-terminal GFP (Green fluorescent protein) fusions, a high-throughput screening platform was constructed. Several highly expressing an robustly performing chimeric designs of cytochrome P450s were identified. This results provided a framework for the broad application of N-terminal tags in the design of microbial cell factories [81].

Conclusions and future prospects

Cytochrome P450 enzymes are one of the dominating driving forces of plant natural products diversity. Currently, P450s in most biosynthetic pathways of PNsps remind unclear. Fortunately, the advances in genomics, bioinformatics and synthetic biology speed up the elucidation of PNsps biosynthetic pathways. We have seen notable progress in gene mining, heterologous expression and protein engineering. The collection of P450 genes in databases also provided a huge improvement for P450 studies [82]. However, post-modification by P450 is still a speed limited step for the biosynthesis of PNsps. At present, the directed evolution and semi-rational design have been applied to improve the performance of P450 enzymes. In the future, desired modifications of the target PNsps would be realized by de novo design of P450 genes.

CRediT authorship contribution statement

Xiaoxian Liu: Writing - original draft. Xiaoli Zhu: Writing - original draft. Hui Wang: Data curation. Tian Liu: Data curation. Jian Cheng: Writing - review & editing. Huifeng Jiang: Writing - review & editing.

Declaration of competing interest

The authors declare no conflict of interest.

Acknowledgments

This work was supported by grants from the National Key R&D Program of China (No. 2019YFA0905700) to J.C., the National Natural Science Foundation of China (NSFC, Grant No. 31901026) and the China Postdoctoral Science Foundation (Grant No. 2019M661032) to X.L. and Tianjin Science and technology plan project (Grant No. 19PT0114018) and Tianjin Synthetic Biotechnology Innovation Capability Improvement Project (TSBICIP-KJGG-007 & TSBICIP-KJGG-002X) H.J.

References

[1] Nelson D, Weerk-Reichart D. A P450-centric view of plant evolution. Plant J 2011;66:194–211.
[2] Jeffreys LN, Girvan HM, McLean KJ, Munro AW. Characterization of cytochrome P450 enzymes and their applications in synthetic biology. Methods Enzymol 2018;608:189–261.
[3] Nelson DR. Cytochrome P450 diversity in the tree of life. Biochimica et biophysica acta. Proteins and proteomics 2018;1866:141–54.
[4] Liu X, et al. Engineering yeast for the production of breviscapine by genomic engineering. Nature communications 2018;9:448.
[5] Li S, Li Y, Smolke CD. Strategies for microbial synthesis of high-value phytochemicals. Plant J 2019;80:793–901.
[6] Lupien S, Karp F, Wildung M, Croteau R. Regiospecific cytochrome P450 limonene hydroxylases from mint (mentha) species: cDNA isolation, characterization, and functional expression of (7)-4S-Limonene-3-hydroxylase and (7)-4S-Limonene-6-hydroxylase. Arch Biochem Biophys 1999;368: 0–192.
[7] Gesell A, et al. The gymnosperm cytochrome P450 CYP750B1 catalyzes stereospecific monoterpene hydroxylation of (S)-sabinene in thujone biosynthesis in western redcedar. Plant Physiol 2015;168:94–106.
[8] Bathe U, Tissier A. Cytochrome P450-dependent enzymes: a driving force of plant diterpene diversity. Phytochemistry 2019;161:149–62.
[9] Turk EM, et al. CYP72B1 inactivates brassinosteroid hormones: an intersection between photomorphogenesis and plant steroid signal transduction. Plant Physiol 2003;133:1643–53.
[10] Liao W, et al. Transcriptome assembly and systematic identification of novel cytochrome P450s in taxus chinensis. Front Plant Sci 2017;8:1468.
[11] Magome H, et al. CYP71F1 and CYP71F2 encode gibberellin 13-oxidases that reduce gibberellin activity in rice. Proc Natl Acad Sci U S A 2013;110:1947–52.
[12] Bak S, Beisson F, Bishop G, Hamberger B, Werck-Reichart D. Cytochromes P450. Arabidopsis Book 2013;1:134.
[13] Li C, Zhu Y, Guo X, Sun C, Chen S. Transcriptome analysis reveals ginsenoside biosynthetic genes, microRNAs and simple sequence repeats in Panax ginseng C. A. Meyer. BMC Genomics 2013;14:245.
[14] Bathé U, Tissier A. Cytochrome P450-dependent enzymes: a driving force of plant diterpene diversity. Phytochemistry 2019;161:149–62.
[15] Liu X, et al. Engineering yeast for the production of breviscapine by genomic engineering. Nature communications 2018;9:448.
[16] Pauli, et al. Molecular cloning and functional heterologous expression of two allelles encoding (S)-N-methylscoulerine 3'-hydroxylase (CYP80B1), a new methyl jasmonate-inducible cytochrome-P450-dependent mono-oxgenase of benzylisoquinoline alkaloid biosynthesis. J Biol Chem 2015;290:793–901.
[17] Horii K, et al. Mining of the uncharacterized cytochrome P450 genes involved in alkaloid biosynthesis in California poppy using a draft genome sequence. Plant Cell Physiol 2018;59:222–33.
[18] Liao W, et al. Transcriptome assembly and systematic identification of novel cytochrome P450s in taxus chinensis. Front Plant Sci 2017;8:1468.
[19] Lee SL, Hirta T, Scott AI. Indole alkaloid biosynthesis in Catharanthus roseus — involvement of geissoschizine and 19-epiajmalicine. Tetrahedron Lett 1979;20:691–4.
[20] Panyi RP, Parajuli P, Kofias MA, Sohng JK. Microbial production of natural and novel flavonoids. Biotechnol Adv 2016;34:63–64.
[21] Zhao Q, et al. Two CYP82D enzymes function as flavone hydroxylases in the biosynthesis of root-specific 4′-deoxyflavonols in Scutellaria baicalensis. Mol Plant 2017;11:335–48.
[22] Pauli, et al. Molecular cloning and functional heterologous expression of two allelles encoding (S)-N-methylscoulerine 3'-hydroxylase (CYP80B1), a new methyl jasmonate-inducible cytochrome-P450-dependent mono-oxgenase of benzylisoquinoline alkaloid biosynthesis. Plant Physiol 2015;168:94–106.
[23] Pauli, et al. The gymnosperm cytochrome P450 CYP750B1 catalyzes stereospecific monoterpene hydroxylation of (S)-sabinene in thujone biosynthesis in western redcedar. Plant Physiol 2015;168:94–106.
[24] Karp F, Wildung M, Croteau R. Regiospecific cytochrome P450 limonene hydroxylases from mint (mentha) species: cDNA isolation, characterization, and functional expression of (7)-4S-Limonene-3-hydroxylase and (7)-4S-Limonene-6-hydroxylase. Arch Biochem Biophys 1999;368: 0–192.
[25] Gesell, et al. The gymnosperm cytochrome P450 CYP750B1 catalyzes stereospecific monoterpene hydroxylation of (S)-sabinene in thujone biosynthesis in western redcedar. Plant Physiol 2015;168:94–106.
[26] Vetter H-P, et al. Molecular analysis and heterologous expression of an inducible cytochrome P-450 protein from periwinkle (catharanthus roseus L.). Plant Physiol 1992;100:998–1007.
[27] Liu X, et al. Engineering yeast for the production of breviscapine by genomic engineering. Nature communications 2018;9:448.
[28] Pauli, et al. Molecular cloning and functional heterologous expression of two allelles encoding (S)-N-methylscoulerine 3′-hydroxylase (CYP80B1), a new methyl jasmonate-inducible cytochrome-P450-dependent mono-oxgenase of benzylisoquinoline alkaloid biosynthesis. J Biol Chem 2015;290:793–901.
[29] Horii K, et al. Mining of the uncharacterized cytochrome P450 genes involved in alkaloid biosynthesis in California poppy using a draft genome sequence. Plant Cell Physiol 2018;59:222–33.
[30] Liao W, et al. Transcriptome assembly and systematic identification of novel cytochrome P450s in taxus chinensis. Front Plant Sci 2017;8:1468.
[31] Lee SL, Hirta T, Scott AI. Indole alkaloid biosynthesis in Catharanthus roseus — involvement of geissoschizine and 19-epiajmalicine. Tetrahedron Lett 1979;20:691–4.
[32] Panyi RP, Parajuli P, Kofias MA, Sohng JK. Microbial production of natural and novel flavonoids. Biotechnol Adv 2016;34:63–64.
[33] Zhao Q, et al. Two CYP82D enzymes function as flavone hydroxylases in the biosynthesis of root-specific 4′-deoxyflavonols in Scutellaria baicalensis. Mol Plant 2017;11:335–48.
[34] Pauli, et al. Molecular cloning and functional heterologous expression of two allelles encoding (S)-N-methylscoulerine 3'-hydroxylase (CYP80B1), a new methyl jasmonate-inducible cytochrome-P450-dependent mono-oxgenase of benzylisoquinoline alkaloid biosynthesis. Plant Physiol 2015;168:94–106.
[35] Pauli, et al. The gymnosperm cytochrome P450 CYP750B1 catalyzes stereospecific monoterpene hydroxylation of (S)-sabinene in thujone biosynthesis in western redcedar. Plant Physiol 2015;168:94–106.
[112] Kitaoka N, Wu Y, Xu M, Peters RJ. Optimization of recombinant expression enables discovery of novel cytochrome P450 activity in rice diterpenoid biosynthesis. Appl Microbiol Biotechnol 2015;99:7549–58.

[113] Li J, et al. Production of plant-specific flavones baicalein and scutellarein in an engineered E. coli from available phenylalanine and tyrosine. Metab Eng 2019;52:124–33.

[114] Schoch GA, Attias R, Belghazi M, Dansette PM, Werck-Reichhart D. Engineering of a water-soluble plant cytochrome P450, CYP73A1, and NMR-based orientation of natural and alternate substrates in the active site. Plant Physiol 2003;133:1198–208.