Available online at www.sciencedirect.com
ScienceDirect
Chinese Journal of Natural Medicines 2020, **18**(6): 417-424
doi: 10.1016/S1875-5364(20)30049-2
Chinese Journal of Natural Medicines

•Review•

Advances in biosynthesis of triterpenoid saponins in medicinal plants

YAO Lu[1, 2], LU Jun[1, 2], WANG Juan[1, 2*], GAO Wen-Yuan[1, 2*]

[1] *Tianjin Key Laboratory for Modern Drug Delivery and High Efficiency, School of Pharmaceutical Science and Technology, Tianjin University, Tianjin 300072, China;*
[2] *Key Laboratory of Systems Bioengineering, Ministry of Education, Tianjin University, Tianjin 300072, China*

Available online 20 Jun., 2020

[ABSTRACT] In recent years, biosynthesis of triterpenoid saponins in medicinal plants has been widely studied because of their active ingredients with diverse pharmacological activities. Various oxidosqualene cyclases, cytochrome P450 monooxygenases, uridine diphosphate glucuronosyltransferases, and transcription factors related to triterpenoid saponins biosynthesis have been explored and identified. In the biosynthesis of triterpenoid saponins, the progress of gene mining by omics-based sequencing, gene screening, gene function verification, catalyzing mechanism of key enzymes and gene regulation are summarized and discussed. By the progress of the biosynthesis pathway of triterpenoid saponins, the large-scale production of some triterpenoid saponins and aglycones has been achieved through plant tissue culture, transgenic plants and engineered yeast cells. However, the complex biosynthetic pathway and structural diversity limit the biosynthesis of triterpenoid saponins in different system. Special focus can further be placed on the systematic botany information of medicinal plants obtained from omics large dataset, and triterpenoid saponins produced by synthetic biology strategies, gene mutations and gene editing technology.

[KEY WORDS] Medicinal plant; Cytochrome P450 monooxygenases; Uridine diphosphate glucuronosyltransferases; Transcription factor; Triterpenoid saponins

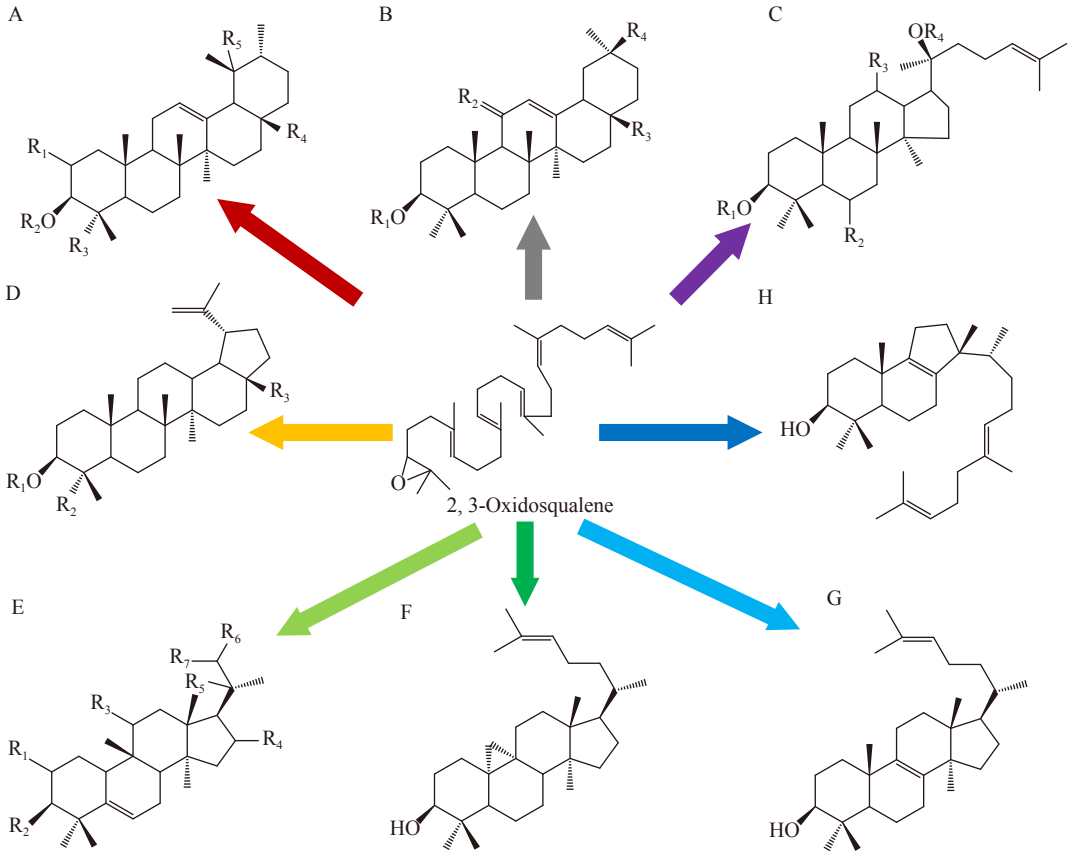
[CLC Number] Q67 **[Document code]** A **[Article ID]** 2095-6975(2020)06-0417-08

Introduction

Medicinal plants growth is influenced by various environmental factors, which induces the diverse change of active ingredients [1, 2]. Many rare medicinal plants are difficult to cultivate and grow slowly [3, 4]. Besides, active ingredients of some medicinal plants have low contents, complex structures, and difficulties in chemical synthesis [5, 6]. But by biotechnology, the production of active ingredients has many advantages, such as a short growth cycle, standardized production processes, and controllable quality [7, 8]. Synthetic biology provides an effective strategy for the sustainable development of traditional Chinese medicine. Oleanolic acid production was improved to 606.9 ± 9.1 mg·L^{-1} by reconstruction of cellular galactose regulatory network and further fermentation optimization in *Saccharomyces cerevisiae*, which was 7.6-fold higher than the reported maximum production [9]. In *S. cerevisiae*, by repurposing UGT51, an inherently promiscuous glycosyltransferase, about 300 mg·L^{-1} of ginsenoside Rh_2 was obtained which was the highest titer reported [10]. By increasing the copy number of *Uni25647* and pairing cytochrome P450 reductases from various plant sources, the production of glycyrrhetinic acid was increased to 18.9 ± 2.0 mg·L^{-1}, which was 946.5-fold higher than previously reported data [11].

Triterpenoid saponins are most studied in medicinal plant with diverse structures skeletons of α-amyrin, β-amyrin, dammarenediol, lupeol, etc. [12-14], further reactions of diverse structures skeletons by cytochrome P450 monooxygenases (CYP450s) and uridine diphosphate glucuronosyltransferases (UGTs) induced the production of numerous triterpenoid saponins (Fig. 1). Biologically, triterpenoid saponins are considered defensive compounds against external stress [15] (Fig. 2). Because of their various pharmacological effects, they are meaningful to humans [16, 17]. Table 1 shows the summary of CYP450s, UGTs and transcription factors (TFs) involved in triterpenoid saponins biosynthesis, including their plant sources, catalytic sites and moieties [8,18-43]. Recent studies focus on CYP450s, UGTs and TFs. Different plant

[Received on] 12-Oct.-2019
[Research funding] This work was supported by the Central Significant Increase or Decrease Program of China (No. 2060302-1604-05) and NSFC (No. 81703639).
[*Corresponding author] E-mail: drwangjuan@163.com (WANG Juan); pharmgao@tju.edu.cn (GAO Wen-Yuan)
These authors have no conflict of interest to declare.



A	R_1	R_2	R_3	R_4	R_5
Ursolic acid	−H	−H	$-CH_3$	−COOH	−H
Asiatic acid	−OH	−H	$-CH_2OH$	−COOH	−H
Ziyuglycoside I	−H	−Ara	$-CH_3$	−COOGlc	−H

B	R_1	R_2	R_3	R_4
Oleanic acid	−H		−COOH	$-CH_3$
Glycyrrhizic acid	−GlcA-GlcA	=O	$-CH_3$	−COOH
Ginsenodide R_0	−Glc-GlcUA		−COOGlc	$-CH_3$

C	R_1	R_2	R_3	R_4
Protopanaxadiol	−H	−H	−OH	−H
Protopanaxatriol	−H	−OH	−OH	−H
Gininsenodide Rh_1	−H	−OGlc	−OH	−H
Gininsenodide Rh_2	−Glc	−H	−OH	−H
Notoginsenodide R_1	−H	$-Oglc\text{-}xy_1$	−OH	−Glc

D	R_1	R_2	R_3
Betulinic acid	−H	$-CH_3$	−COOH
Pulchinenoside C	−Ara-Rha	$-CH_2OH$	−Glc-Glc-Rha

E	R_1	R_2	R_3	R_4	R_5	R_6	R_7
Mogrol	−H	−OH	−OH	−H	−H	−H	$-CH_2CHOH\text{-}C(CH_3)_2OH$
Mogroside	−H	−Glc-Glc	−OH	−H	−H	−H	$-CH_2CH(Glc)_3\text{-}C(CH_3)_2OH$
Calebassine	−OH	=O	=O	−OH	−OH	=O	$-CH{=}CH\text{-}C(CH_3)_2OC{=}OCH_3$

F
Cycloartenol

G
Lanosterol

H
Thalianol

Fig. 1 Triterpenoid saponins most studied in medicinal plant with diverse structures skeletons of *α*-amyrin (A), *β*-amyrin (B), dammarenediol (C), lupeol (D), cucurbitadienol (E), cycloartenol (F), lanosterol (G) and thalianol (H), respectively

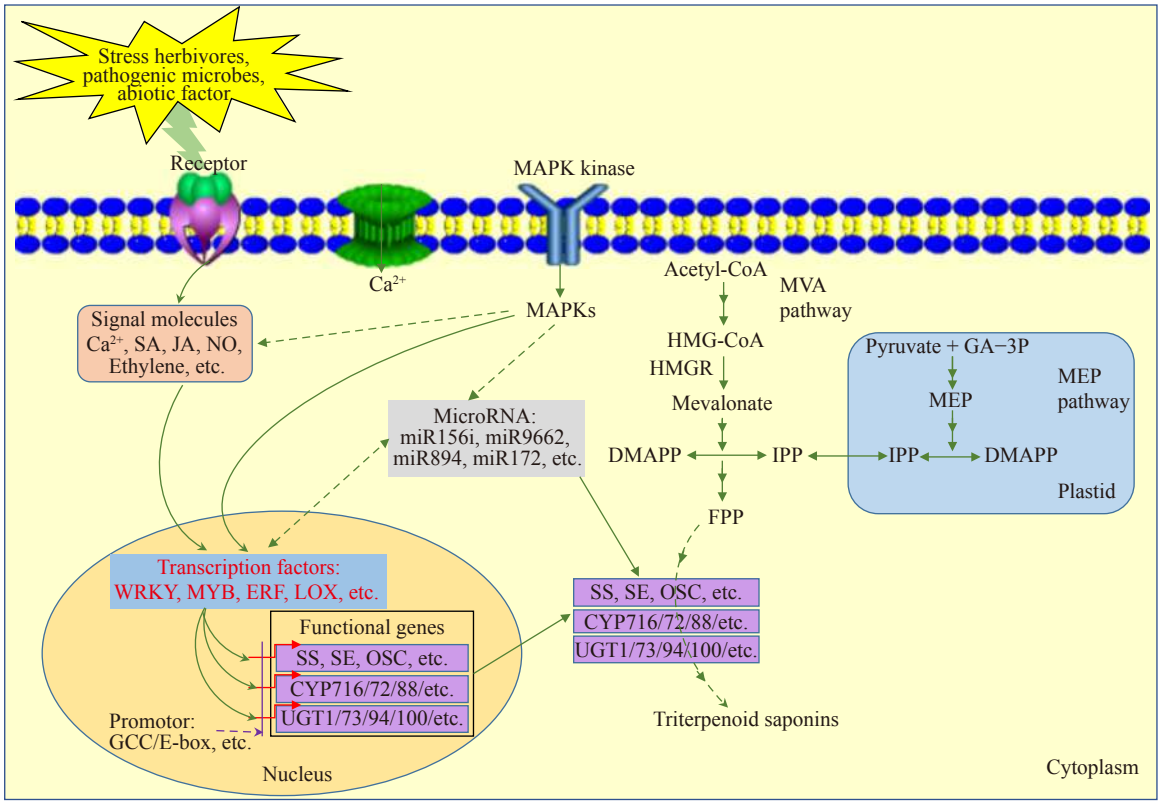


Fig. 2 The regulation of triterpenoid saponins biosynthesis in medicinal plant. Triterpenoid saponins are synthesized *via* the cytosol MVA pathway and plastid MEP pathway. Different stress affects the change of signal molecules and MAPKs, then actives the expression of transcription factors, microRNA. Transcription factors and microRNA further regulates the functional genes expression. HMG-CoA, 3-hydroxy-3-methylglutaryl-CoA; MEP, 2-C-methyl-Derythritol 4-phosphate; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; SS, squalene synthase; SE, squalone epoxidase; OSC, oxidosqualene cyclase; UGT, UDP-glycosyltransferases. Dotted lines with multiple arrows represent multiple enzymatic catalyzed steps

sources are studied, such as *Artemisia annua*, *Panax ginseng*, *Medicago sativa*. Catalytic sites and moieties of various CYP450s and UGTs also have been studied, including 3-carbonyl, 6-hydroxyl, 23-hydroxyl, 28-carboxyl, 3-glucose, 3-galactose, etc. In dammarane-type saponins biosynthesis, protopanaxadiol is synthesized from hydroxylation C-12 of dammarenediol-II by CYP716A47 [44]. Pq3-O-UGT1 from *P. quinquefolius* was identified to transfer a glucose moiety into the C-3 glucose in ginsenoside Rh_2 [45]. However, there is no report on the biosynthesis of notoginsenosides catalyzed by uridine diphosphate xylosyltransferase. TFs that participate in the biosynthesis of triterpenoid saponins usually belong to the family of bHLH [46], ERF [41], MYB [39, 40], and WRKY [43]. The mechanisms of TFs on triterpenoid saponins biosynthesis still need further study.

This review summarizes the recent progress in triterpenoid saponins biosynthesis in medicinal plants including their sources, key biosynthesis genes, etc. Gene mining, gene screening, gene function verification, catalyzing mechanism of key enzymes, and the regulation of triterpenoid saponins biosynthesis are also summarized. Perspectives on further discovery of medicinal plants biosynthesis and synthetic biology strategies for microbial production of triterpenoid saponins derived from plant are also discussed.

Gene Mining by Omics-based Sequencing

Omics-based research provides crucial insights into individual biological components (such as gene expression, protein modification, metabolite profiles) levels. Through statistical tools and data mining, the integration of omics datasets provides a blueprint of relationships between different biological components. These analyses assist in devising strategies to produce metabolites of interest by generating bacterial or plant-based systems [47].

In recent years, some important medicinal plants have been subjected to omics sequencing, such as *A. annua* [48], *Gastrodia elata* [49], *G. uralensis* [50], *P. notoginseng* [51], which draws the blueprint of medicinal plants and realizes sustainable production of pharmacological triterpenoid saponins. In *P. notoginseng*, genes potentially related to ginsenoside biosynthesis were discovered by a draft genome and comparative transcriptome, which provides the foundation and makes it possible to genetic improvement of *P. notoginseng* [52]. These potential key genes in triterpenoid saponins biosynthesis need to be further screened by genes screening in order to discover the key one.

Gene Screening

In recent years, many triterpenoid saponins such as gin-

Table 1 Overview of medicinal plant CYP450s, UGTs and transcription factors in triterpenoid saponins biosynthesis

	Gene	Position	Plant sources	Reference
CYP450s	CYP716A14v2	3-carbonyl	*Artemisia annua*	[18]
	CYP716A53v2	6-hydroxyl	*Panax ginseng*	[19]
	CYP716A47	12-hydroxyl	*Panax ginseng*	[19]
	CYP716A52v2	28-carboxyl	*Panax ginseng*	[20]
	CYP716A12	28-carboxyl	*Medicago sativa*	[21]
	CYP88D6	11-carbonyl	*Glycyrrhiza uralensis*	[22]
	CYP72A154	30-hydroxyl	*Glycyrrhiza uralensis*	[23]
	CYP716A94	28-carboxyl	*Kalopanax septemlobus*	[24]
	CYP72A397	23-hydroxyl		
UGTs	UGT73C10	3-glucose	*Barbarea vulgaris*	[25]
	UGT73C11			
	UGT74AE2	3-glucose	*Panax ginseng*	[26]
	UGT94Q2	3-glucose(2-1) glucose	*Panax ginseng*	[26]
	UGT74AC1	3-glucose	*Siraitia grosvenorii*	[27]
	GuUGAT	3-glucuronic acid	*Glycyrrhiza uralensis*	[28]
	UGT71A27	20-glucose	*Panax ginseng*	[29]
	UGTPg1	20-glucose	*Panax ginseng*	[8]
	UGTPg100	6-glucose	*Panax ginseng*	[8]
	UGTPg101	6-glucose	*Panax ginseng*	[30]
	UGTPg100	28-glucose	*Panax ginseng*	[8]
	UGTPg101	28-glucose	*Panax ginseng*	[30]
	UGT74M1	28-glucose	*Saponaria vaccaria*	[31]
	UGT73AD1	28-glucose	*Centella asiatica*	[32]
	UGT73P2	3-galactose	*Glycine max*	[33]
	UGT91H4	3-rhamnose	*Glycine max*	[33]
	UGT73P12	3-glucuronic acid	*Glycyrrhiza uralensis*	[34]
	UGT73F4	22-xylose	*Glycine max*	[35]
	UGT73F2	22-xylose	*Glycine max*	[35]
Transcription factors	PgWRKY1-9		*Panax ginseng*	[36]
	PgWRKY1		*Panax ginseng*	[37]
	PgLOX6		*Panax ginseng*	[38]
	PgMYB1		*Panax ginseng*	[39]
	PgMYB1-5		*Panax ginseng*	[40]
	PnERF1		*Panax notoginseng*	[41]
	PqWRKY1		*Panax quinquefolium*	[42]
	WsWRKY1		*Withania somnifera*	[43]

senoside, notoginsenoside and glycyrrhizin are mostly extracted from native roots and tissue cultures of medicinal plant [28, 53, 54]. The related biosynthesis genes have been discovered [55, 56], but lots of key genes have not been explored. For example, xylosyltransferase genes in notoginsenoside biosynthesis of *P. notoginseng* have not been clearly studied.

A UGT gene, *Pq3-O-UGT2* was obtained from *P. quinquefolius* by genome-wide searching. The BLASTP search results revealed its deduced amino acid sequence had 99.55% similarity with that of PgUGT94Q2 in *P. ginseng*. The gene of *Pq3-O-UGT2* also showed a close evolutionary relationship with ginseng UGTs by phylogenetic analysis [55]. In *W. somnifera*, by phylogenetic analysis, WsWRKY1 showed close similarity to AaWRKY1, CrWRKY1 from *A. annua*, *C. roseus*, respectively. WsWRKY1 was further characterized as WRKY protein, which was involved in regulating stress tolerance of *W. somnifera* [43]. By genetic screening, key genes in the biosynthesis of triterpenoid saponins could be further selected, which makes it possible to explore gene function and improve the content of bioactive compounds by

genetic manipulation.

Verification of Gene Function

Verification of gene function in vivo or in vitro

Overexpression and RNA interference (RNAi) are common methods to verify gene function *in vivo* and can perform direct verification of gene function. In *P. quinquefolius* transgenic hairy roots, overexpression of *Pq3-O-UGT1* gene led to a higher level of protopanaxadiol-group ginsenosides, which is consistent with enzymatic assay *in vitro* that Pq3-O-UGT1 could glycosylate protopanaxadiol to produce ginsenoside Rh_2 [45]. However, it is difficult to induce their genetically modified system for some plants. Thus, verification of gene function *in vitro* was very meaningful.

Verification of gene function *in vitro* was usually studied by model plant (such as *A. thaliana*, tobacco) or engineered strains (such as *S. cerevisiae*, *Escherichia coli*) transferred specific vectors with target genes. Overexpression of *PgHMGR1* from *P. ginseng* resulted in higher production of triterpenes and sterols in Arabidopsis [56]. In *S. cerevisiae*, the function of CYP716A252 and CYP716A253 from *Ocimum basilicum* was also determined [57]. Through gene function verification, key genes could be introduced in other expression system to yield corresponding triterpenoid saponins. In *P. japonicas*, *β*-Amyrin synthase was recognized as the first key enzyme in oleanane-type saponins biosynthesis. *P. notoginseng* cells transferred into *β*-Amyrin synthase gene from *P. japonicus* (*PjβAS*) produced two oleanane-type saponins, chikusetsusaponin IV and chikusetsusaponin IVa. Both saponins were only contained in *P. japonicas*, and were first discovered in transgenic *P. notoginseng* cells [58].

When gene function is verified, these genes could be introduced into engineering strains and their corresponding product could be synthesized by synthetic biology strategies. Furthermore, gene mutations and gene editing technology could vary gene expression causing a change in active product content. By repurposing UGT51, an inherently promiscuous glycosyltransferase, about 300 $mg \cdot L^{-1}$ of ginsenoside Rh_2 was obtained from *S. cerevisiae* [10]. Metabolic switching of astringent and beneficial triterpenoid saponins in soybean could be achieved by a loss-of-function mutation in cytochrome P450 72A69 [59]. Verification of gene function is the significant step, and it provides a basis for synthetic biology to obtain expected compounds.

Subcellular localization

Subcellular localization was commonly used to reveal the expression and distribution sites of expressed protein [43, 60]. PgCYP736B was discovered to localize in the plasma membrane, which indicated PgCYP736B could express and function on endoplasmic reticulum [61]. UGTs are generally localized in the cytosol of medicinal plants. Subcellular localization of GFP-tagged UGT71C5 showed UGT71C5 distributed throughout the protoplast [62], similar to the cytosolic localization of the other reported UGTs [63]. A MYB gene, *PgMYB1* was cloned from *P. ginseng* C.A. Meyer. PgMYB1-mGFP5 fusion protein was discovered to specifically localize in the nucleus by subcellular localization [39]. Overall, subcellular localization of CYP450s, UGTs and TFs contributes to exploring their functional site and makes it possible to obtain the corresponding protein by directed transformation.

Catalyzing Mechanism of Key Enzymes

The discovery of catalyzing mechanism of key enzymes helps further verify gene function. It promotes the exploration of specific substrate, key catalytic sites of key enzymes [26, 28, 45]. Furthermore, it contributes to regulate triterpenoid saponins biosynthesis in transformed cell lines or engineered strains with recombinant vectors constructed by specifically optimized key enzymes and efficient substrate in large scale. The catalytic function of GuUGAT was determined by an *in vitro* enzyme assay, which showed GuUGAT could catalyse glycyrrhetinic acid to produce glycyrrhizin. Site-directed mutagenesis revealed key catalytic sites of glucuronosylation as Gln-352, and Gln-392, Glu-375, His-22, Trp-370, respectively [28].

With the development of gene engineering, gene mutation makes it possible to obtain specific functional genes in triterpenoid saponins biosynthesis. Gene mutation is essential to discover the key amino acids of catalyzing enzyme, and help to improve enzyme activity. By structural modeling and site-directed mutagenesis, Wei *et al*. found key amino acids of UGTPg1, UGTPg100 and UGTPg102 that could play essential roles in determining their bioactivities [30]. The discovery of catalyzing mechanism of key enzymes makes it possible to further change enzyme activity on their key catalytic sites.

Gene Regulation of Triterpenoid Saponins Biosynthesis in Medicinal Plants

The biosynthesis of triterpenoid saponins is affected by gene regulation in medicinal plants. Gene expression varies in different plants. Even in the same genus, genes expression is different. Both *P. ginseng* and *P. notoginseng* belong to Araliaceae family, but their saponins composition and concentration show a big difference [64, 65]. This difference may finally attribute to the various gene regulations of triterpenoid saponins biosynthesis genes between *P. ginseng* and *P. notoginseng*. Artificial changes in gene regulation by increasing the gene copy number or knocking out negative branch pathway help to synthesize expected compounds.

Key genes

Overexpression is a common method that could help verify and regulate key genes in triterpenoid saponins biosynthesis. In *P. quinquefolius* transgenic hairy roots, *Pq3-O-UGT1* overexpression induced Pq3-O-UGT1 mRNA accumulation and a higher level of protopanaxadiol-group ginsenosides, which is consistent with enzymatic assay *in vitro* that Pq3-O-UGT1 could catalyze protopanaxadiol to produce ginsenoside Rh_2 [45]. RNAi also is performed to regulate the spe-

cific functional genes in triterpenoid saponins biosynthesis. In *I. indigotica*, transcription factor Ii049 was characterized to play important role in lignan/lignin biosynthesis regulation. In order to verify its function, Ii049 was knocked down by RNAi approach, which caused a remarkable reduction of expression levels of genes involved in lignan/lignin biosynthesis and lignan/lignin contents. The result suggested Ii049 positively regulated the biosynthesis of lignan [60]. These results from overexpression and RNAi could directly reflect the regulation of genes on triterpenoid saponins biosynthesis.

Transcription factors

Transcription factors (TFs), also known as trans-acting factors, are DNA-binding proteins. They specifically interact with cis-acting elements in the promoter region of eukaryotic genes, inducing the expression of other related proteins, or inhibiting downstream genes transcription. They play an essential role in secondary metabolites biosynthesis [46, 66].

In *P. ginseng* C.A. Meyer, eight WRKY genes were cloned, characterized and their encoding WRKY proteins were assigned to WRKY Group II by phylogenetic analysis [36]. In PnbHLH1 transgenic cells, the expression levels of key triterpenoid saponins biosynthesis genes were higher than those in control. Similarly, the total saponin contents were increased compared with the control. These results suggested that in *P. notoginseng*, the PnbHLH1 is a positive regulator in triterpenoid saponins biosynthesis [46].

MicroRNA

MicroRNA (miRNA), a class of endogenous small noncoding RNA, have been discovered to play an important role in regulating gene expressions in plants [67]. In *P. ginseng*, transcriptome analysis discovered potential miRNAs could regulate ginsenosides biosynthesis by function on their target of TFs and other proteins [68]. There were four miRNAs (miR5021, miR5163, miR5293 and novel_miR_27) predicted to target genes involved in the terpenoid backbone biosynthesis pathway in *P. notoginseng* [69].

Target genes and target proteins exploration

In recent years, accumulating evidence suggests that certain regulatory factor, such as TFs, microRNA regulate secondary metabolites production by regulating genes involved in metabolic pathway [43, 69-71]. Yeast one-hybrid (Y1H) assay, electrophoreticmobility shift assay (EMSA) and chromatin immunoprecipitation assay were performed to discover the interaction between protein and DNA [43, 72]. The yeast one-hybrid assay revealed that AaGSW1 of *A. annua* only bound to one of the W-box motifs in CYP71AV1 promoter. AaGSW1 positively regulates CYP71AV1 expression. Moreover, overexpression of AaGSW1 significantly improves artemisinin content. These results revealed AaGSW1 was a positive regulator in the biosynthetic pathway of artemisinin [53]. The EMSA demonstrated that PnERF1 might bind to the promoter regions which contained GCC-box and regulate corresponding genes expression [41]. The binding interaction between PgMYB2 protein and dammarenediol synthase (PgDDS) promoter was found by Y1H assay. Moreover, EMSA identified their binding site. Based on the key role of PgDDS gene in ginsenoside synthesis, it is reasonable to believe that PgMYB2 has the potential to improve the ginsenoside production through genetic and metabolic engineering [73]. The protein-protein interactions were explored by yeast two-hybrid assay, GST pull-down, dual-luciferase assay, coimmunoprecipitation, bimolecular fluorescence complementation. The exploration of target genes and target proteins was meaningful to regulate triterpenoid saponins biosynthesis by synthetic biology strategies, gene mutations and gene editing technology in order to further improve the yield of triterpenoid saponins.

Future Perspectives and Conclusions

In this review, we systematically summarize the identification and verification of key genes in triterpenoid saponins biosynthesis in medicinal plants including their plant sources, catalytic sites and moieties. Although several triterpenoid saponins have been successfully synthesized and their yield have great improvement by biomanufacturing [10, 54, 74], there are still huge challenges in triterpenoid saponins industrial production. For some essential triterpenoid saponins (such as notoginsenoside R1), their biosynthesis catalyzed by uridine diphosphate xylosyltransferase is needed. Furthermore, the biosynthesis pathway of triterpenoid saponins is not clear and their regulation mechanism is not very clear. Future study can focus on clarification the biosynthetic pathway of triterpenoid saponins and discovering their regulatory mechanism in order to regulate it. In addition, high-yield mutant strain/plant reconstructed by synthetic biology strategies, gene mutations and gene editing technology was meaningful to further increase triterpenoid saponins production.

References

[1] Ye Y, Ding Y, Jiang Q, *et al*. The role of receptor-like protein kinases (RLKs) in abiotic stress response in plants [J]. *Plant Cell Rep*, 2017, **36**(2): 235-242.

[2] Song X, Wu H, Yin Z, *et al*. Endophytic bacteria isolated from *Panax ginseng* improves ginsenoside accumulation in adventitious ginseng root culture [J]. *Molecules*, 2017, **22**(6): E837-849.

[3] Duan L, Xiong XJ, Hu JY, *et al*. *Panax notoginseng* saponins for treating coronary artery disease: a functional and mechanistic overview [J]. *Front Pharmacol*, 2017, **8**: 702-720.

[4] Zeng X, Li Y, Ling H, *et al*. Transcriptomic analyses reveal clathrin-mediated endocytosis involved in symbiotic seed germination of *Gastrodia elata* [J]. *Bot Stud*, 2017, **58**(1): 31-42.

[5] Dai LH, Li J, Yang JG, *et al*. Use of a promiscuous glycosyltransferase from *Bacillus subtilis* 168 for the enzymatic synthesis of novel protopanaxatriol-type ginsenosides [J]. *J Agric Food Chem*, 2018, **66**(4): 943-949.

[6] Liu CX. Plant tissue culture and biosynthesis provide a fast way to produce active constituents of traditional Chinese medicines [J]. *Chin Herb Med*, 2017, **9**(2): 99-100.

[7] Wang P, Wei Y, Fan Y, *et al*. Production of bioactive ginsenosides Rh2 and Rg3 by metabolically engineered yeasts [J]. *Metab Eng*, 2015, **29**: 97-105.

[8] Yan X, Fan Y, Wei W, *et al*. Production of bioactive ginsenoside compound K in metabolically engineered yeast [J]. *Cell Res*, 2014, **24**(6): 770-773.

[9] Zhao Y, Fan J, Wang C, *et al*. Enhancing oleanolic acid pro-

duction in engineered *Saccharomyces cerevisiae* [J]. *Bioresour Technol*, 2018, **257**: 339-343.
[10] Zhuang Y, Yang GY, Chen X, *et al*. Biosynthesis of plant-derived ginsenoside Rh2 in yeast *via* repurposing a key promiscuous microbial enzyme [J]. *Metab Eng*, 2017, **42**: 25-32.
[11] Zhu M, Wang C, Sun W, *et al*. Boosting 11-oxo-beta-amyrin and glycyrrhetinic acid synthesis in *Saccharomyces cerevisiae via* pairing novel oxidation and reduction system from legume plants [J]. *Metab Eng*, 2018, **45**: 43-50.
[12] Mangas S, Moyano E, Osuna L, *et al*. Triterpenoid saponin content and the expression level of some related genes in calli of *Centella asiatica* [J]. *Biotechnol Lett*, 2008, **30**(10): 1853-1859.
[13] Seki H, Tamura K, Muranaka T. P450s and UGTs: key players in the structural diversity of triterpenoid saponins [J]. *Plant Cell Physiol*, 2015, **56**(8): 1463-1471.
[14] Luo ZL, Zhang KL, Ma XJ, *et al*. Research progress in synthetic biology of triterpene saponins [J]. *Chinese Traditional and Herbal Drugs*, 2016, **47**(10): 1806-1814.
[15] Szakiel A, Pączkowski C, Henry M. Influence of environmental abiotic factors on the content of saponins in plants [J]. *Phytochem Rev*, 2010, **10**(4): 471-491.
[16] Shibata S. Chemistry and cancer preventing activities of ginseng saponins and some related triterpenoid compounds [J]. *J Korean Med Sci*, 2001, **16**: S28-37.
[17] Murthy HN, Dandin VS, Park SY, *et al*. Quality, safety and efficacy profiling of ginseng adventitious roots produced *in vitro* [J]. *Appl Microbiol Biotechnol*, 2018, **102**(17): 7309-7317.
[18] Moses T, Pollier J, Shen Q, *et al*. OSC2 and CYP716A14v2 catalyze the biosynthesis of triterpenoids for the cuticle of aerial organs of *Artemisia annua* [J]. *Plant Cell*, 2015, **27**(1): 286-301.
[19] Han JY, Hwang HS, Choi SW, *et al*. Cytochrome P450 CYP716A53v2 catalyzes the formation of protopanaxatriol from protopanaxadiol during ginsenoside biosynthesis in *Panax ginseng* [J]. *Plant Cell Physiol*, 2012, **53**(9): 1535-1545.
[20] Han JY, Kim MJ, Ban YW, *et al*. The involvement of beta-amyrin 28-oxidase (CYP716A52v2) in oleanane-type ginsenoside biosynthesis in *Panax ginseng* [J]. *Plant Cell Physiol*, 2013, **54**(12): 2034-2046.
[21] Moses TPJ, Almagro L, Buyst D, *et al*. Combinatorial biosynthesis of sapogenins and saponins in *Saccharomyces cerevisiae* using a C-16α hydroxylase from *Bupleurum falcatum* [J]. *Proc Natl Acad Sci U S A*, 2014, **111**(4): 1634-1639.
[22] Seki H, Ohyama K, Sawai S, *et al*. Licorice beta-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin [J]. *Proc Natl Acad Sci U S A*, 2008, **105**(37): 14204-14209.
[23] Seki H, Sawai S, Ohyama K, *et al*. Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin [J]. *Plant Cell*, 2011, **23**(11): 4112-4123.
[24] Han JY, Chun JH, Oh SA, *et al*. Transcriptomic analysis of *Kalopanax septemlobus* and characterization of KsBAS, CYP716A94 and CYP72A397 genes involved in hederagenin saponin biosynthesis [J]. *Plant Cell Physiol*, 2018, **59**(2): 319-330.
[25] Augustin JM, Drok S, Shinoda T, *et al*. UDP-glycosyltransferases from the UGT73C subfamily in *Barbarea vulgaris* catalyze sapogenin 3-*O*-glucosylation in saponin-mediated insect resistance [J]. *Plant Physiol*, 2012, **160**(4): 1881-1895.
[26] Jung SC, Kim W, Park SC, *et al*. Two ginseng UDP-glycosyltransferases synthesize ginsenoside Rg3 and Rd [J]. *Plant Cell Physiol*, 2014, **55**(12): 2177-2188.
[27] Dai LH, Liu C, Zhu YM, *et al*. Functional characterization of cucurbitadienol synthase and triterpene glycosyltransferase involved in biosynthesis of mogrosides from *Siraitia grosvenorii* [J]. *Plant Cell Physiol*, 2015, **56**(6): 1172-1182.
[28] Xu G, Cai W, Gao W, *et al*. A novel glucuronosyltransferase has an unprecedented ability to catalyse continuous two-step glucuronosylation of glycyrrhetinic acid to yield glycyrrhizin [J]. *New Phytol*, 2016, **212**(1): 123-135.
[29] Lu J, Yao L, Li JX, *et al*. Characterization of UDP-glycosyltransferase involved in biosynthesis of ginsenosides Rh1 and Rd and identification of critical conserved amino acid residues for its function [J]. *J Agric Food Chem*, 2018, **66**(36): 9446-9455.
[30] Wei W, Wang PP, Wei YJ, *et al*. Characterization of *Panax ginseng* UDP-glycosyltransferases catalyzing protopanaxatriol and biosyntheses of bioactive ginsenosides F1 and Rh1 in metabolically engineered yeasts [J]. *Mol Plant*, 2015, **8**(9): 1412-1424.
[31] Meesapyodsuk D, Balsevich J, Reed DW, *et al*. Saponin biosynthesis in *Saponaria vaccaria*. cDNAs encoding beta-amyrin synthase and a triterpene carboxylic acid glucosyltransferase [J]. *Plant Physiol*, 2007, **143**(2): 959-969.
[32] de Costa F, Barber CJS, Kim YB, *et al*. Molecular cloning of an ester-forming triterpenoid: UDP-glucose 28-*O*-glucosyltransferase involved in saponin biosynthesis from the medicinal plant *Centella asiatica* [J]. *Plant Sci*, 2017, **262**: 9-17.
[33] Shibuya M, Nishimura K, Yasuyama N, *et al*. Identification and characterization of glycosyltransferases involved in the biosynthesis of soyasaponin I in *Glycine max* [J]. *FEBS Lett*, 2010, **584**(11): 2258-2264.
[34] Nomura Y, Seki H, Suzuki T, *et al*. Functional specialization of UDP-glycosyltransferase 73P12 in licorice to produce a sweet triterpenoid saponin, glycyrrhizin [J]. *Plant J*, 2019, **99**(6): 1127-1143.
[35] Sayama T, Ono E, Takagi K, *et al*. The Sg-1 glycosyltransferase locus regulates structural diversity of triterpenoid saponins of soybean [J]. *Plant Cell*, 2012, **24**(5): 2123-2138.
[36] Xiu H, Nuruzzaman M, Guo XQ, *et al*. Molecular cloning and expression analysis of eight PgWRKY genes in *Panax ginseng* responsive to salt and hormones [J]. *Int J Mol Sci*, 2016, **17**(3): 319-333.
[37] Nuruzzaman M, Cao H, Xiu H, *et al*. Transcriptomics-based identification of WRKY genes and characterization of a salt and hormone-responsive PgWRKY1 gene in *Panax ginseng* [J]. *Acta Biochim Biophys Sin (Shanghai)*, 2016, **48**(2): 117-131.
[38] Rahimi S, Kim YJ, Sukweenadhi J, *et al*. PgLOX6 encoding a lipoxygenase contributes to jasmonic acid biosynthesis and ginsenoside production in *Panax ginseng* [J]. *J Exp Bot*, 2016, **67**(21): 6007-6019.
[39] Afrin S, Zhu J, Cao HZ, *et al*. Molecular cloning and expression profile of an abiotic stress and hormone responsive MYB transcription factor gene from *Panax ginseng* [J]. *Acta Bioch Bioph Sin*, 2015, **47**(4): 267-277.
[40] Choi JY, Abbai R, Kim YJ, *et al*. Molecular characterization of MYB transcription factor genes from *Panax ginseng* [J]. *Russ J Plant Physiol*, 2017, **64**(3): 398-409.
[41] Deng B, Huang Z, Ge F, *et al*. An AP2/ERF family transcription factor PnERF1 raised the biosynthesis of saponins in *Panax notoginseng* [J]. *J Plant Growth Regul*, 2017, **36**(3): 691-701.
[42] Sun YZ, Niu YY, Xu J, *et al*. Discovery of WRKY transcription factors through transcriptome analysis and characterization of a novel methyl jasmonate inducible PqWRKY1 gene from *Panax quinquefolius* [J]. *Plant Cell Tiss Organ Cult*, 2013, **114**: 269-277.
[43] Singh A, Kumar S, Dwivedi V, *et al*. A WRKY transcription factor from *Withania somnifera* regulates triterpenoid withanolide accumulation and biotic stress tolerance through modulation of phytosterol and defense pathways [J]. *New Phytol*, 2017, **215**(3): 1115-1131.
[44] Han JY, Kim HJ, Kwon YS, *et al*. The Cyt P450 enzyme

CYP716A47 catalyzes the formation of protopanaxadiol from dammarenediol-II during ginsenoside biosynthesis in *Panax ginseng* [J]. *Plant Cell Physiol*, 2011, **52**(12): 2062-2073.
[45] Lu C, Zhao SJ, Wang XS. Functional regulation of a UDP-glucosyltransferase gene (Pq3-O-UGT1) by RNA interference and overexpression in *Panax quinquefolius* [J]. *Plant Cell Tiss Organ Cult*, 2017, **129**(3): 445-456.
[46] Zhang X, Ge F, Deng B, *et al*. Molecular Cloning and Characterization of PnbHLH1 Transcription Factor in *Panax notoginseng* [J]. *Molecules*, 2017, **22**(8): E1268-1280.
[47] Rai A, Saito K, Yamazaki M. Integrated omics analysis of specialized metabolism in medicinal plants [J]. *Plant J*, 2017, **90**(4): 764-787.
[48] Shen Q, Zhang LD, Liao ZH, *et al*. The genome of *Artemisia annua* provides insight into the evolution of Asteraceae family and artemisinin biosynthesis [J]. *Mol Plant*, 2018, **11**(6): 776-788.
[49] Yuan Y, Jin XH, Liu J, *et al*. The Gastrodia elata genome provides insights into plant adaptation to heterotrophy [J]. *Nat Commun*, 2018, **9**(1): 1615-1626.
[50] Mochida K, Sakurai T, Seki H, *et al*. Draft genome assembly and annotation of *Glycyrrhiza uralensis*, a medicinal legume [J]. *Plant J*, 2017, **89**(2): 181-194.
[51] Chen W, Kui L, Zhang GH, *et al*. Whole-genome sequencing and analysis of the chinese herbal plant *Panax notoginseng* [J]. *Mol Plant*, 2017, **10**(6): 899-902.
[52] Zhang D, Li W, Xia EH, *et al*. The medicinal herb *Panax notoginseng* genome provides insights into ginsenoside biosynthesis and genome evolution [J]. *Mol Plant*, 2017, **10**(6): 903-907.
[53] Zheng Y, Chen K, Xu Z, *et al*. Small RNA profiles from *Panax notoginseng* roots differing in sizes reveal correlation between miR156 abundances and root biomass levels [J]. *Sci Rep*, 2017, **7**(1): 9418-9433.
[54] Thanh NT, Murthy HN, Paek KY. Optimization of ginseng cell culture in airlift bioreactors and developing the large-scale production system [J]. *Ind Crop Prod*, 2014, **60**: 343-348.
[55] Lu C, Zhao S, Wei G, *et al*. Functional regulation of ginsenoside biosynthesis by RNA interferences of a UDP-glycosyltransferase gene in *Panax ginseng* and *Panax quinquefolius* [J]. *Plant Physiol Biochem*, 2017, **111**: 67-76.
[56] Kim YJ, Lee OR, Oh JY, *et al*. Functional analysis of 3-hydroxy-3-methylglutaryl coenzyme a reductase encoding genes in triterpene saponin-producing ginseng [J]. *Plant Physiol*, 2014, **165**(1): 373-387.
[57] Misra R, Sharma S, Sandeep, *et al*. Two CYP716A subfamily cytochrome P450 monooxygenases of sweet basil play similar but nonredundant roles in ursane- and oleanane-type pentacyclic triterpene biosynthesis [J]. *New Phytol*, 2017, **214**(2): 706-720.
[58] Zhang X, Yu Y, Jiang S, *et al*. Oleanane-type saponins biosynthesis in *Panax notoginseng via* transformation of beta-amyrin synthase gene from *Panax japonicus* [J]. *J Agric Food Chem*, 2019, **67**(7): 1982-1989.
[59] Yano R, Takagi K, Takada Y, *et al*. Metabolic switching of astringent and beneficial triterpenoid saponins in soybean is achieved by a loss-of-function mutation in cytochrome P450 72A69 [J]. *Plant J*, 2017, **89**(3): 527-539.
[60] Ma R, Xiao Y, Lv Z, *et al*. AP2/ERF transcription factor, Ii049, positively regulates lignan biosynthesis in *Isatis indigotica* through activating salicylic acid signaling and lignan/lignin pathway genes [J]. *Front Plant Sci.*, 2017, **8**: 1361-1377.
[61] Devi BSR, Rahimi S, Yang DC. Characterization of squalene-induced PgCYP736B involved in salt tolerance by modulating key genes of abscisic acid biosynthesis [J]. *Int J Biol Macromol*, 2018, **121**: 796-805.
[62] Liu Z, Yan JP, Li D, *et al*. UDP-glucosyltransferase71c5, a major glucosyltransferase, mediates abscisic acid homeostasis in Arabidopsis [J]. *Plant Physiol*, 2015, **167**(4): 1659-1670.
[63] Lim E-K, Doucet CJ, Hou B, *et al*. Resolution of (+)-abscisic acid using an Arabidopsis glycosyltransferase [J]. *Tetrahedron: Asymmetry*, 2005, **16**(1): 143-147.
[64] Li F, Lv CN, Li Q, *et al*. Chemical and bioactive comparison of flowers of *Panax ginseng* Meyer, *Panax quinquefolius* L., and *Panax notoginseng* Burk [J]. *J Ginseng Res*, 2017, **41**(4): 487-495.
[65] Wang JR, Yau LF, Gao WN, *et al*. Quantitative comparison and metabolite profiling of saponins in different parts of the root of *Panax notoginseng* [J]. *J Agric Food Chem*, 2014, **62**(36): 9024-9034.
[66] Liu J, Gao F, Ren J, *et al*. A novel AP2/ERF transcription factor CR1 regulates the accumulation of vindoline and serpentine in *Catharanthus roseus* [J]. *Front Plant Sci*, 2017, **8**: 2082-2093.
[67] Bartel DP. MicroRNAs: target recognition and regulatory functions [J]. *Cell*, 2009, **136**(2): 215-233.
[68] Li CF, Zhu YJ, Guo X, *et al*. Transcriptome analysis reveals ginsenosides biosynthetic genes, microRNAs and simple sequence repeats in *Panax ginseng* C. A. Meyer [J]. *BMC Genomics*, 2013, **14**: 245-256.
[69] Wei R, Qiu D, Wilson IW, *et al*. Identification of novel and conserved microRNAs in *Panax notoginseng* roots by high-throughput sequencing [J]. *BMC Genomics*, 2015, **16**: 835-845.
[70] Tamura K, Yoshida K, Hiraoka Y, *et al*. The basic Helix-Loop-Helix transcription factor GubHLH3 positively regulates soyasaponin biosynthetic genes in *Glycyrrhiza uralensis* [J]. *Plant Cell Physiol*, 2018, **59**(4): 778-791.
[71] Nejat N, Ramalingam A, Mantri N. Advances in transcriptomics of plants [J]. *Adv Biochem Eng Biotechnol*, 2018, **164**: 161-185.
[72] Mertens J, Pollier J, Vanden Bossche R, *et al*. The bHLH transcription factors TSAR1 and TSAR2 regulate triterpene saponin biosynthesis in *Medicago truncatula* [J]. *Plant Physiol*, 2016, **170**(1): 194-210.
[73] Liu T, Luo T, Guo X, *et al*. PgMYB2, a MeJA-responsive transcription factor, positively regulates the dammarenediol synthase gene expression in *Panax Ginseng* [J]. *Int J Mol Sci*, 2019, **20**(9): E2219-2234.
[74] Johnson EE, Jetter R, Wasteneys G. Rapid induction of the triterpenoid pathway in *Arabidopsis thaliana* mesophyll protoplasts [J]. *Biotechnol Lett*, 2014, **36**(4): 855-858.

Cite this article as: YAO Lu, LU Jun, WANG Juan, GAO Wen-Yuan. Advances in biosynthesis of triterpenoid saponins in medicinal plants [J]. *Chin J Nat Med*, 2020, **18**(6): 417-424.