Chinese Journal of Natural Medicines 24 (2026) 999-1005

Contents lists available at ScienceDirect

Chinese Journal of Natural Medicines

journal homepage: www.cjnmcpu.com/

Original article

Identification and characterization of a promiscuous pentacyclic triterpenoid *O*-glycosyltransferase from *Phytolacca americana*

Yi Zhang[a,Δ], Jinxin Wang[a,Δ], Yiqing Zhu[a], Yunheng Shen[a], Hongyuan Dong[a], Haiyang Yu[a], Zirui An[a], Weidong Zhang[a,b,*], Yu Zhang[b,*], Zhimin Hu[a,*]

[a] *Department of Phytochemistry, School of Pharmacy, Naval Medical University, Shanghai 200433, China*

[b] *Shanghai Frontiers Science Center for Chinese Medicine Chemical Biology, Institute of Interdisciplinary Integrative Medicine Research, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China*

ARTICLE INFO

Article history:
Received 9 October 2025
Revised 14 December 2025
Accepted 22 February 2026
Available online 20 August 2026

Keywords:
Phytolacca americana
Pentacyclic triterpenoid
Glycosylation
Glycosyltransferase

ABSTRACT

Glycosyltransferases (GTs) are key enzymes in the glycosylation of plant secondary metabolites, primarily catalyzing the transfer of a sugar moiety from an activated donor to a specific acceptor molecule. Triterpenoid saponins, an abundant and diverse group of natural products, are composed of triterpenoid aglycones and one or more sugar chains, under the catalysis of GTs. Phytolacca Radix is a traditional Chinese medicine containing over 40 triterpenoid saponins with pharmacological values. However, the identification of glycosyltransferases related to triterpenoid saponin synthesis from *Phytolacca* species remains scarce. In this study, a novel glycosyltransferase, PamUGT, was identified from *Phytolacca americana*. PamUGT exhibits catalytic activity toward 28-/30-COOH of pentacyclic triterpenoids, and it exhibits a preference for the C-30 position when both carboxyl groups are available. To our knowledge, PamUGT is the first glycosyltransferase identified in *Phytolacca* species that can catalyze triterpenoids. Further analysis of substrate specificity revealed that PamUGT catalyzes the glycosylation of a wide range of compounds, including triterpenoids, flavonoids, diterpenoids, alkaloids, and phenolic acids. Our findings identify a highly promiscuous glycosyltransferase, offering a valuable enzymatic tool for modifying diverse natural products.

1. Introduction

Glycosylation is widely recognized as the final step in the formation of glycosides and a key modification that significantly contributes to the structural diversity and complexity of pentacyclic triterpenoids [1, 2]. It can also improve pharmaceutical properties such as solubility, stability, and bioactivity, while reducing toxicity [2], as exemplified by compounds such as α-hederin [3], QS-21 [4], and saponariosides A and B [5]. Glycosyltransferases (GTs) catalyze the transfer of sugar moieties from activated sugar donors to various acceptor molecules. Among these, uridine diphosphate (UDP)-dependent glycosyltransferases (UGTs) play important roles in the biosynthesis of glycosides and typically catalyze the glycosylation at functional groups including carbon atoms, hydroxyl groups (–OH), carboxyl groups (–COOH), amino groups (–NH_2), and thiol groups (–SH) [6, 7].

Triterpenoid saponins are a class of medicinally valuable plant natural products consisting of a hydrophobic triterpenoid aglycone (sapogenin) and hydrophilic sugar moieties. The amphiphilic nature enables them to form water-soluble complexes with hydrophobic drug molecules, thereby enhancing the solubility and bioavailability of poorly soluble pharmaceuticals [8, 9]. Therefore, triterpenoid saponins exhibit diverse pharmacological activities [9], including anti-tumor [10], anti-bacterial [11], anti-viral [12], and anti-inflammatory [13] effects. However, the research and production of many such compounds are hindered by their complex chemical structures and limited plant availability [2, 14]. In recent years, biosynthesis has attracted increasing attention due to its environmental friendliness, high efficiency, product specificity, and mild reaction conditions. These advantages have greatly avoided a series of drawbacks in extraction and chemical synthesis [2, 15]. However, the biosynthetic pathways of many triterpenoid saponins remain constrained by the incomplete identification of key enzymes, such as GTs [6].

Phytolacca acinosa, as well as its relative *Phytolacca americana*, is the botanical origin of the traditional Chinese medicine known as Shanglu (Phytolacca Radix) with a long history of use. These plants produce various chemical constituents, including triterpenoids, flavonoids, and sterols [16]. Among these, triterpenoid saponins are characteristic compounds of Shanglu and one of the main bioactive constituents. They have been reported to exhibit anti-tumor [17], anti-inflammatory [13], and anti-viral [18]. To date, over 60 triterpenoids have been isolated from *Phytolacca* species. Triterpenoid saponins account for two-thirds of these compounds. Structurally, all of these triterpenoids are based on the oleanane-type skeleton [17]. These triterpenoid saponins are typically glycosylated at the C-3 hydroxyl and C-28 carboxyl positions. Their sugar moieties primarily consist of glucose and xylose, with some also containing rhamnose and galactose [19]. Although two

* Corresponding author.
E-mail addresses: wdzhangy@hotmail.com (W. Zhang); yzhang@shutcm.edu.cn (Y. Zhang); 13021941113@163.com (Z. Hu)
Δ These authors contributed equally to this work.

https://doi.org/10.1016/S1875-5364(26)61198-3

glycosyltransferases, PaGT2 and PaGT3, have been identified in *P. americana*, their catalytic activities appear to be restricted to flavonoids and phenolic compounds [20].

Herein, we report a functional GT from *P. americana*, designated PamUGT, which catalyzes the glycosylation at the C28-/C30-COOH of oleanane-type triterpenoids. Enzymatic property determination showed that it exhibited the optimal catalytic activity at 50 °C and pH 7.4. Investigations into substrate and sugar donor promiscuity revealed that PamUGT could utilize a variety of sugar donors and possess a broad substrate scope, including triterpenoids, flavonoids, diterpenoids, alkaloids, and phenolic acids. This study provides an alternative approach for the preparation of diverse *O*-glycosylated compounds.

2. Results

2.1. Gene screening and functional characterization

To identify GTs related to triterpenoid saponin biosynthesis in *Phytolacca*, the publicly available transcriptome data of three *Phytolacca* species were downloaded and analyzed using a combination of transcriptomic and bioinformatic approaches, with known glycosyltransferases involved in the triterpenoid saponin biopathway serving as reference [21]. A total of 36 candidate *GT* genes were screened out, and 31 of them were successfully cloned using gene-specific primers. After sequencing for confirmation, these candidate genes were cloned into the pET-28a(+) vector, and the recombinant plasmids were transformed into *Escherichia coli* transetta (DE3) strains and induced with isopropyl *β*-D-thiogalactoside (IPTG) for protein expression.

Functional characterization was subsequently performed using substrates **1**, **2** and **3** as substrates with UDP-Glc as the sugar donor in 50 mmol·L^{-1} Tris-HCl buffer (pH 7.4). The reactions were incubated overnight at 37 °C at 200 r·min^{-1}. UPLC/Q-TOF-MS analysis revealed that only PamUGT exhibited catalytic activity toward three sugar acceptors, with conversion rates ranging approximately from 30% to 40%. A new product was detected in each of the three reaction mixtures containing PamUGT with substrate **1**, **2**, or **3**. As shown in Fig. 1, the [M − H]$^-$ ion of the products exhibited an *m/z* 162 higher than that of the substrates, consistent with the mass of the corresponding monoglucosides. Further structural elucidation by NMR spectroscopy confirmed that product **1a** was phytolaccagenin 28-*O*-*β*-D-glucoside [22]. It can be initially inferred that **3a** is bayogenin 28-*O*-*β*-D-glucoside based on structural similarity. Therefore, PamUGT may be involved in the biosynthesis of esculentoside through glycosylation at the 28-COOH position. Unexpectedly, product **2a** was jaligonic acid 30-*O*-*β*-D-glucoside, determined by NMR spectroscopy [23], although compounds **2** and **1** are highly similar (Figs. 1A and 1B). The above results indicate that PamUGT can glycosylate the 28-/30-COOH groups of oleanene-type pentacyclic triterpenoids. Furthermore, when carboxyl groups are present at both the C-28 and C-30 positions, PamUGT exhibits a preferential selectivity toward the 30-COOH for glycosylation (compound **2**, Fig. 1B).

PamUGT comprises an open reading frame (ORF) of 1443 bp, encoding a protein of 480 amino acids. The molecular mass of the purified protein was approximately 60 kDa, consistent with the theoretically predicted molecular mass of recombinant PamUGT protein (Fig. S2). Multiple sequence alignment was performed between PamUGT and several known characterized glycosyltransferases from other species, including UGT74M1 (ABK76266.1), UGT74AG1 (OR902350), and UGTPg8 (AIE12488.1), which glycosylated the 28-COOH group, as well as UGT73F17 (AXS75258.1), which acted on the 30/29-COOH group. The results revealed that PamUGT exhibits low amino acid sequence similarity with the reported enzymes, with a maximum identity of only 33.06% (Table S5), though all sequences contained the conserved PSPG (plant secondary product glycosyltransferase) motif, which is characteristic of GT function (Fig. 2A). Phylogenetic analysis revealed that PamUGT clusters closely with GT known to glycosylate the C28-COOH group, which is consistent with the experimental demonstration that PamUGT catalyzes glycosylation at the C28-COOH position (Fig. 2B).

The enzymatic properties of PamUGT were investigated using compound **1** as the acceptor and UDP-Glc as the sugar donor. The optimal reaction temperature and pH for PamUGT were determined to be 50 °C and 7.4 (50 mmol·L^{-1} Tris-HCl buffer), respectively. Mg^{2+} could modestly enhance the enzymatic activity (Fig. S3). The apparent K_m and k_{cat} for substrates **1**, **2** and **3** are shown in Table 1 (Fig. S4).

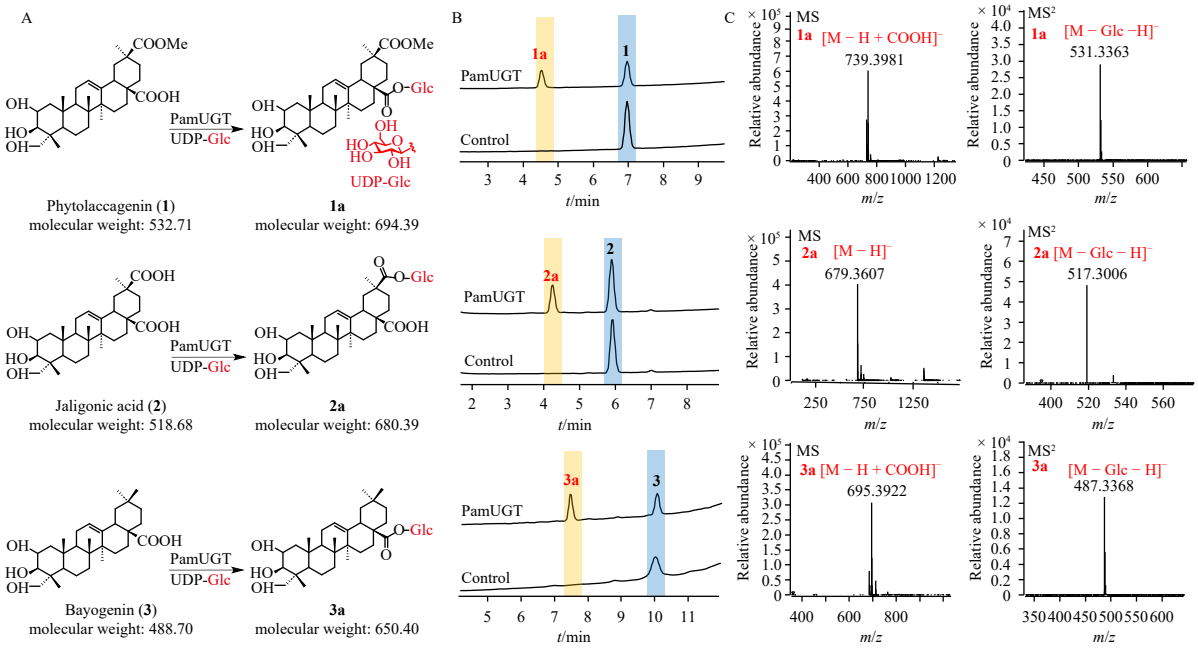

Fig. 1 Functional characterization of the glycosyltransferase PamUGT. (A) PamUGT catalyzes the *O*-glycosylation of substrates **1**, **2**, and **3**. (B) Total ion chromatograms (TICs) of the reaction mixtures of PamUGT with substrates **1**, **2**, and **3**. (C) MS/MS analysis of the products **1a**, **2a**, and **3a**.

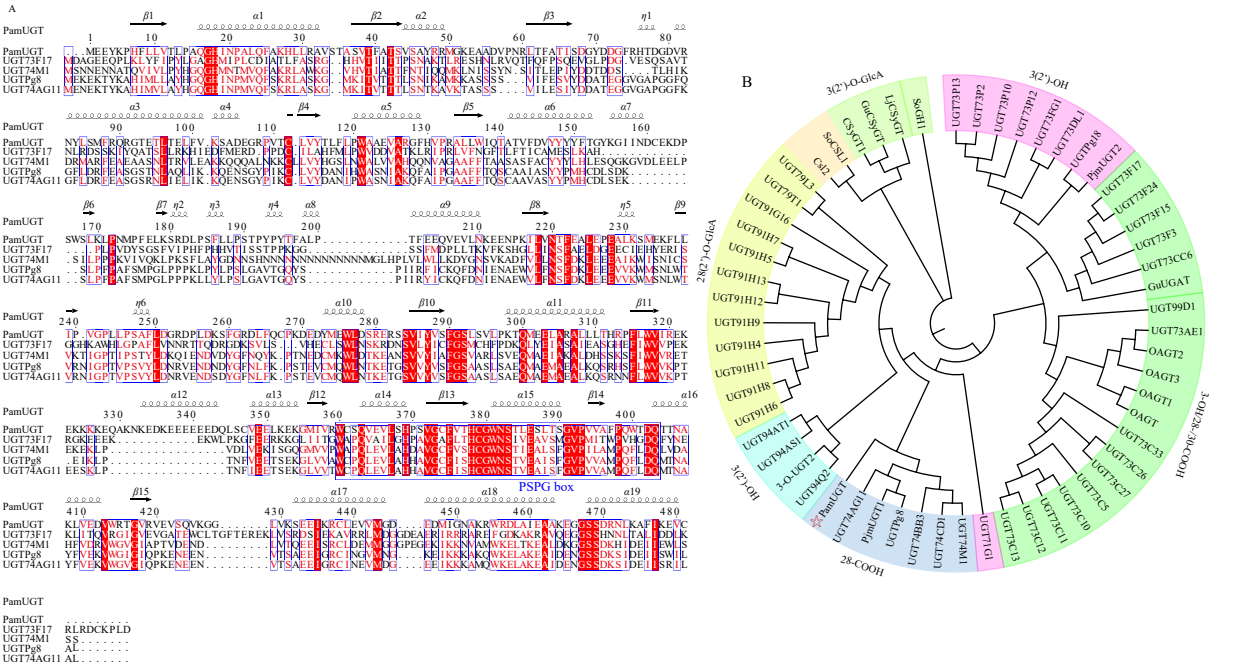

Fig. 2 Sequence alignment and phylogenetic analysis of PamUGT. (A) Multiple sequence alignment of PamUGT amino acid sequences with UGT74M1 (ABK76266.1), UGT74AG11 (OR902350), UGTPg8 (AIE12488.1), UGT73F17 (AXS75258.1). The blue box represents the PSPG box. (B) Phylogenetic analysis of PamUGT and UGT catalyzing oleanane-type triterpenoids from other plants. Enzyme was marked by "☆" in this study.

2.2. Probing the sugar donor specificity of PamUGT

To elucidate the sugar donor selectivity of PamUGT, five common UDP-sugars including UDP-glucose (UDP-Glc), UDP-xylose (UDP-Xyl), UDP-glucuronic acid (UDP-GlcA), and UDP-L-rhamnose (UDP-Rha), were evaluated as potential donors, using compounds **1**, **2** and **9** as acceptor substrates separately. As shown in Fig. 3, PamUGT displayed strict specificity for UDP-Glc when compound **1** was used as the acceptor. In contrast, with acceptors **2** and **9**, the enzyme exhibited catalytic promiscuity, utilizing multiple sugar donors including UDP-Glc. Notably, all five UDP-sugars could be accepted using compound **9** as substrate under the applied catalytic conditions. The above results indicate that PamUGT is a promiscuous GT that retains a strong inherent preference for UDP-Glc.

2.3. Exploring the substrate promiscuity of PamUGT

Most plant GTs possess a broad substrate scope and can serve as biocatalysts for the synthesis of glycosidic compounds [24, 25]. To investigate the substrate scope of PamUGT, 42 structurally diverse compounds, including triterpenoids (including saponins, **1**–**20**), diterpenoids (**21**), flavonoids (**22**–**38**), and others (**39**–**42**), were screened as potential glycosyl acceptors with UDP-Glc as the donor (Fig. 4B). PamUGT showed no activity toward tetracyclic triterpenoids but catalyzed the glycosylation of all tested pentacyclic triterpenoids (Fig. 4A). UPLC/Q-TOF-MS analysis indicated that all products exhibited an $[M + Glc - H]^-$ ion (or a corresponding adduct) with *m/z* 162 higher than their respective substrates, confirming the formation of monoglucosides (Fig. S38–S63).

Table 1 Kinetic parameters of PamUGT for representative substrates **1**–**3**, **9**, **21**, **24** and **34**.

Compound	K_m/(μmol·L^{-1})	k_{cat}/(s^{-1})	k_{cat}/K_m
1	169.56	5.85×10^{-5}	0.809
2	754.86	1.38×10^{-4}	0.183
3	320.71	1.17×10^{-4}	0.364
9	463.41	1.51×10^{-4}	0.325
21	1010.12	3.52×10^{-4}	0.348
24	790.42	3.05×10^{-4}	0.386
34	318.89	1.37×10^{-4}	0.430

Based on the structure of compound **1** and the sugar addition site of product **1a**, it is preliminarily speculated that products **3a**–**8a** were 28-*O*-glucoside derivatives (Fig. 4B). To further determine the catalytic capability of PamUGT at 30-COOH, we prepared the product of glycyrrhetinic acid (**9**). The NMR result showed that **9a** was glycyrrhetinic acid 30-*O*-*β*-D-glucoside (Figs. S20–S25) [23]. The apparent K_m and k_{cat} for **9** are shown in Table 1 (Fig. S4). It can therefore be inferred that compounds **10a** and **11a** are also glycosylation products at the 30-COOH position. This result is consistent with the glycosylation site specificity exhibited by PamUGT for substrate **2** (Fig. 1B). As shown in Fig. 4A, the conversion rates of the aglycones were higher than those of corresponding saponins (**1**, **5**–**7**, and **9**–**11**), suggesting that PamUGT preferentially utilizes free triterpenoid aglycones and that catalytic efficiency decreases with the increase of sugar (Fig. 4A). This may be attributed to the effect of steric hindrance, as a higher number of sugar groups leads to greater steric crowding, thereby resulting in a reduction in the catalytic rate. Furthermore, structural elucidation by NMR confirmed that the glycosylated products **12a** and **13a** are asiatic acid 28-*O*-glucoside and celastrol 30-*O*-glucoside, respectively. This result further demonstrates that PamUGT catalyzes the glycosylation of the 28-COOH or 30-COOH groups of pentacyclic triterpenoids (Figs. S26–S37).

PamUGT exhibited broad acceptance of various flavonoids, with conversion rates for most exceeding 50%, significantly higher than those observed with triterpenoids (Fig. 4A). Enzymatic reactions with compounds **22**, **24**, and **28** yielded multiple glycosylated products. Notably, products **22a**, **24a/b**, and **28a/b** exhibited an *m/z* increase of 324 Da, identifying them as diglucosides. Comparison with standards confirmed **22b**, **24b**, **24c**, and **28c** as apigenin-7-*O*-glucoside, luteolin-3',7-di-*O*-glucoside, luteolin-7-*O*-glucoside, and quercetin-7-*O*-glucoside, respectively (Figs. S49, S50, S53). Furthermore, the glycosylated products of isoflavone **33** and flavanone **35** were both identified as 7-*O*-glucosides, further supporting the high regioselectivity of PamUGT

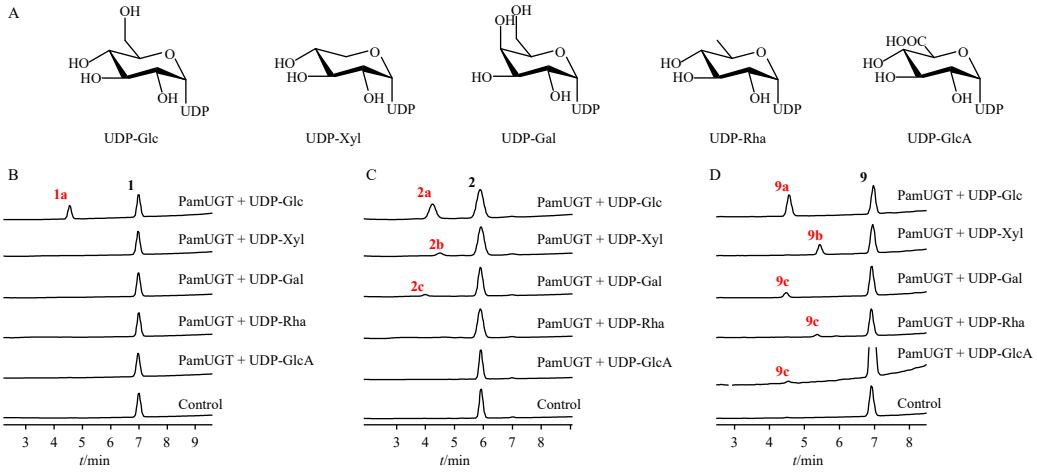

Fig. 3 Sugar donor specificity of PamUGT. (A) The structures of sugar donors. (B) TICs of compound **1** as a sugar acceptor. (C) TICs of compound **2** as a sugar acceptor. (D) TICs of compound **9** as a sugar acceptor.

for the 7-OH position of flavonoids (Figs. S56, S58). Based on the catalytic activity of PamUGT on luteolin-7-*O*-glucoside, quercetin-7-*O*-glucoside, and quercetin-4'-*O*-glucoside, we proposed that **24a** and **28a** are the 7,4'-di-*O*-glucoside derivatives, whereas **24d** and **28d** are the 3'-*O*-glucoside derivatives. The high yields of **24b** and **25b** indicated strong regioselectivity of PamUGT for the 3'-OH and 7-OH positions (Fig. 4A).

PamUGT also demonstrated superior catalytic efficiency for substrate **21**, yielding 68%, which was about 1.7 times higher than that for substrate **1** (Fig. 4A). The product was confirmed as steviol-19-*O*-glucoside by comparison with a chemical standard (Fig. S48). The enzyme also exhibited low activity (6%–15% conversion) against phenolic, alkaloid and phenylpropanoid compounds, generating a single glycosylated product (Figs. 4A, S61–S63), although these compounds do not exist in *P. americana*. These results confirm that PamUGT is a versatile GT with a broad substrate scope and high efficiency for most flavonoids.

3. Discussion

The amphiphilic structure of saponins, comprising a hydrophobic aglycone and hydrophilic sugar chains, is fundamental to their structure-activity relationship, critically determining their biological effects [9]. GTs play a crucial role in the metabolism of triterpenoid saponins. It is closely associated with the structural diversity of triterpenoids and confers multiple biological activities to these compounds [26]. Pentacyclic triterpenoids commonly form glycosidic bonds at the C3 and C28 positions, with a few occurring 30-COOH. GTs catalyzing the glycosylation of the C3-OH and C28-COOH in pentacyclic triterpenoids have been extensively studied in various plants, however, most of them exhibit strict regioselectivity [2]. *P. americana* is a traditional Chinese medicine in which triterpenoids constitute the primary chemical constituents. However, investigations into the relevant GTs remain scarce.

A 28-/30-COOH *O*-glycosyltransferase, designated PamUGT, was identified from *P. americana*. To our knowledge, PamUGT represents the first GT discovered in *Phytolacca* species capable of catalyzing pentacyclic triterpenoids. Although PamUGT clusters with those GTs catalyzing 28-COOH, PamUGT shares a low degree of similarity to GTs catalyzing the glycosylation at 28-COOH (30.22%–33.06%) and 30-COOH (23.22%–28.07%) (Fig. 2, Table S5). PamUGT is capable of catalyzing phytolaccagenin (**1**) and its glycosides at the 28-COOH position, leading to the formation of corresponding esculentoside. Therefore, it can be preliminarily speculated that PamUGT participates in esculentoside biosynthesis in *P. americana* through glycosylation at the 28-COOH group (Fig. S5). Generally, GTs that catalyze at different positions on the same substrate can produce multiple products if the corresponding sites on the substrate are available. For example, UGT73C33 [27], a glycosyltransferase from *Glycyrrhiza uralensis*, can glycosylate the 3-OH and 28-COOH positions of glycyrrhetinic acid separately or simultaneously, resulting in three products. However, when both 28-COOH and 30-COOH are present, PamUGT exclusively glycosylates the 30-COOH position, forming a monoglycoside. The absence of PamUGT activity against tetracyclic triterpenoids is likely due to the conformational flexibility of their C17 single bond. This flexibility introduces dynamic steric and electronic constraints that hinder the substrate from achieving a reactive orientation. Consistent with this premise, compound **42**, which also possesses a rotatable bond at the analogous position, was not accepted as a substrate. This parallel observation validates the proposed mechanism that conformational lability at this key position is detrimental to catalysis by PamUGT.

Further phylogenetic analysis revealed that PamUGT clustered closely with UGT75L4 from *Maclura pomifera* [28], which glycosylates the 7-OH position of flavonoids, and UGT75L20 from *Rubus suavissimus* [29], which glycosylates the 19-COOH position of diterpenoids (Fig. S7). This phylogenetic proximity suggested that PamUGT may also catalyze the glycosylation at the 7-OH of flavonoids and the 19-COOH of diterpenoids. Enzymatic assays *in vitro* demonstrated that PamUGT exhibits significantly higher catalytic efficiency toward diterpenoids and flavonoids compared to triterpenoid substrates. This activity profile aligns with its closer phylogenetic relationship to glycosyltransferases known to act on flavonoids and diterpenoids (Fig. 4B). Although no clear univariate structure-activity relationship was observed among flavonoid substrates, PamUGT showed markedly higher catalytic activity toward flavones and flavonols, efficiently generating multiple glycosylated products from these compounds. This suggests that the enzyme possesses a broad substrate adaptability and notable regioselectivity toward this class of structures. However, phytolaccagenin (**1**) exhibited the lowest K_m value, suggesting it is a potential native substrate for PamUGT. Notably, diterpenoid **21** and flavonoids **24** and **34** also displayed substantial catalytic efficiencies, with k_{cat}/K_m values within the same order of magnitude as those of the triterpenoid substrates. These results provide evidence for the broad substrate promiscuity of PamUGT (Table 1). PaGT2 and PaGT3, two flavonoid-glycosylating UGTs isolated from *P. americana* previously, share 24.24% and 25.65% amino acid sequence similarities with PamUGT, respectively [30]. Both enzymes exhibit substrate promiscuity toward diverse flavonoids. Notably, PaGT3 efficiently catalyzes quercetin glycosylation to generate four products, all of which are monoglucosides [30]. Additionally, PaGT3 can glycosylate the alkaloid capsaicin. Although several promiscuous GTs, such as MhGT1 [31], OleD [25], and UGT734AN3 [24], have been reported, enzymes cap-

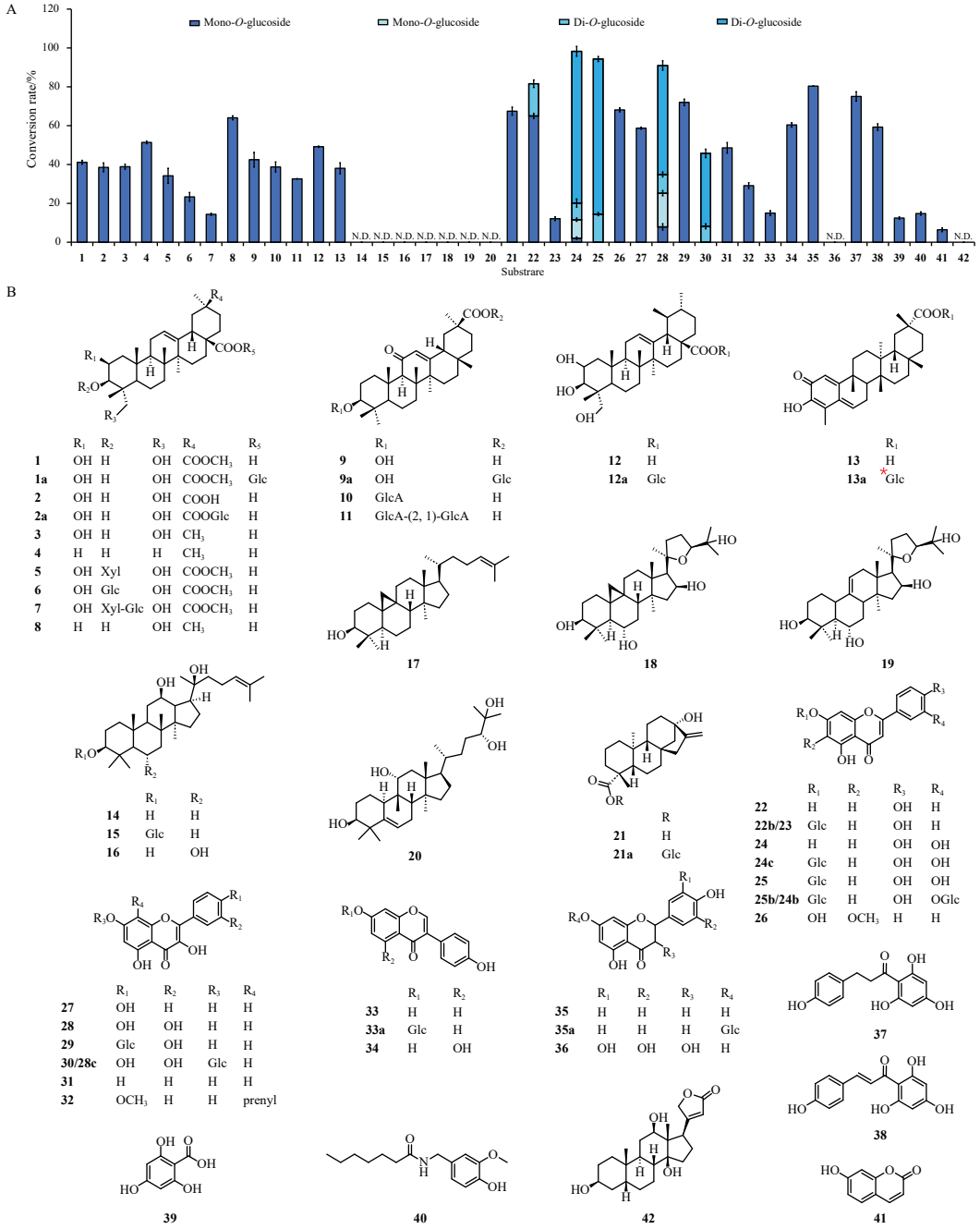

Fig. 4 Substrate promiscuity of PamUGT. (A) Conversion rate percentages of glycosylation products for different substrates (**1**–**42**) using UDP-Glc as the sugar donor. (B) Structures of the compounds (**1**–**42**) catalyzed by PamUGT. Products **1a**, **2a**, **9a**, **12a** and **13a** were identified by NMR. Products **21a**, **22b**, **24b**, **24c**, **25b**, **28c**, **33a** and **35a** were consistent with the authentic standards. N.D. denotes "not detected". Error bars represent the standard deviation of three biological replicates. * represents new compound.

able of efficiently glycosylating such a broad range of substrates, including flavonoids, diterpenoids, and diverse triterpenoids, are relatively uncommon.

4. Conclusion

In this study, we identify a glycosyltransferase that catalyzes the glycosylation of pentacyclic triterpenoids at the 28-COOH or 30-COOH positions. Notably, PamUGT showed a preference for the 30-COOH position to yield a monoglycosylated product when both 28-COOH and 30-COOH are present. Substrate promiscuity assays further revealed the broad catalytic scope of PamUGT, including triterpenoids, flavonoids, diterpenoids, and phenolic compounds. The unique preference and unexpected catalytic promiscuity of PamUGT distinguish it from previously described glycosyltransferases. These distinctive features will provide novel insights into the functional evolution and adaptive mechanisms of plant UGTs.

5. Experimental

5.1. Materials and chemicals

P. americana plants were collected from Shanghai, China, in September 2023. The fresh tissues were immediately frozen in liquid nitrogen, stored at −80 °C, and subsequently used for RNA and DNA extraction. The plasmid of pET-28a(+) (Invitrogen, USA) was preserved in our laboratory. *E. coli* DH5α (Tsingke Biotech, China) and transetta (DE3) (TransGen Biotech, China) competent cells were employed for vector construction and protein expression, respectively.

Substrates phytolaccagenin (**1**), jaligonic acid (**2**), **7**, **11**, **13**–**16**, **20**–**22**, **24**–**42** were purchased from Bio CuiRan (Hubei, China). Substrates bayogenin (**3**), **4**, **10** were purchased from NatureStandard (Shanghai, China). Substrates **6**, **9**, **12**, **17**–**19**, **23** were purchased from Desite (Chengdu, China). Substrates **5** were purchased from BAICAOYUAN (Jiangxi, China). Substrate **8** was isolated by our group from *Verbena officinalis* previously. Sugar donors UDP-Xyl, UDP-Glc and UDP-Gal were purchased from CuiYuan (Hubei, China). Sugar donors UDP-GluA and UDP-Rha were purchased from Yuanye (Shanghai, China).

5.2. Gene screening and plasmid construction

The original transcriptome data for three *Phytolacca* species (*P. acinosa*, *P. americana*, and *P. icosandra*) [21] were downloaded from the NCBI database (https://www.ncbi.nlm.nih.gov/) and assembled using the Trinity software (Table S1). Putative glycosyltransferase sequences were identified by performing a local BLAST analysis (using BioEdit) with 18 reported plant oleanane-type glycosyltransferases as reference queries (E-value ≤ 10^{-20}, ORF length > 1000 bp; Table S2). Homologous sequences conserved across all three species were identified as the final set of candidate genes. Among these, genes from *P. americana* were chosen for subsequent experimental validation.

Genomic DNA and total RNA were extracted from the leaves of *P. americana* using the DNAquick Plant System (TIANGEN, China) and the TransZol up kit (TransGen Biotech, China), respectively. The samples were genetically identified and classified into *P. americana* through DNA barcoding [32] (Fig. S1). The extracted RNA was then reverse-transcribed into cDNA using the TransScript® Uni One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen Biotech, China) according to the instructions. The full-length sequences of candidate genes were amplified with the primers listed in Table S3, using the 2 × TransTaq® High Fidelity (HiFi) PCR SuperMix I (TransGen Biotech, China). The resulting PCR products were subcloned into the pET-28a(+) vector using the pEASY®-Uni Seamless Cloning and Assembly kit (TransGen Biotech, China) and then transformed into *E. coli* DH5α competent cells to confirm whether the recombinant plasmids have been successfully constructed.

5.3. Protein expression and purification

After sequence verification, the recombinant plasmids were transformed into *E. coli* transetta (DE3) competent cells for heterologous expression. The transformed strains were cultured in Luria-Bertani (LB) medium containing 50 μg·mL^{-1} kanamycin at 37 °C until the OD_{600} reached 0.6–0.8. Protein expression was induced by adding 0.1 mmol·L^{-1} IPTG, followed by incubation at 16 °C for 18–20 h.

Bacterial cells were collected by centrifugation at 12 000 r·min^{-1} for 3 min and resuspended in lysis buffer (1 g/8 mL, pH 7.4, 50 mmol·L^{-1} Tris-HCl, 10 mmol·L^{-1} imidazole, 300 mmol·L^{-1} NaCl), and disrupted by ultrasonication (30%) on ice for 15 min. The lysate was centrifuged at 12 000 r·min^{-1} for 50 min at 4 °C to remove cell debris. Then, 1 mL of Ni-NTA Resin (TransGen Biotech, China) was added to the supernatant, and the mixture was incubated with shaking for 1 h at 4 °C. Subsequently, the mixture was transferred into an affinity chromatography column (BioSharp, China) and purified by lysis buffer with different concentrations of imidazole (from 20 to 250 mmol·L^{-1}). The elution fractions containing the target protein were merged. The purified protein was concentrated using Amicon® Ultra-15 Centrifugal Filters 30K (Millipore, USA). Finally, the concentrate was washed three times with storage buffer containing 20% glycerol (pH 7.4, 50 mmol·L^{-1} Tris-HCl, 300 mmol·L^{-1} NaCl), and stored at −80 °C.

5.4. Functional characterization *in vitro* and phylogenetic analysis

To identify the function of the candidate enzymes, phytolaccagenin (**1**), jaligonic acid (**2**) and bayogenin (**3**) was used as substrates. The reaction mixture was prepared containing 100 μg of purified enzyme, 0.5 mmol·L^{-1} UDP-Glc, 0.2 mmol·L^{-1} aglycone, respectively, and 50 mmol·L^{-1} Tris-HCl buffer (pH 7.4) in a total volume of 100 μL. The reaction was carried out at 37 °C for 12 h and subsequently terminated by adding 200 μL of ice-cold methanol. The mixture was then centrifuged at 12 000 r·min^{-1} for 20 min, and the supernatant was analyzed by ultra-performance liquid chromatography/quadrupole time-of-flight mass spectrometry (UPLC/Q-TOF-MS, Agilent, USA).

The amino acid sequence of PamUGT was aligned with those of glycosyltransferases known to catalyze the 28-COOH position [UGT74M1 (ABK76266.1), UGT74AG11 (OR902350), and UGT-Pg8 (AIE12488.1)] and the 30-COOH position [UGT73F17 (AXS75258.1)] using MEGA. A neighbor-joining tree was constructed with bootstrap values from 1000 replicates in MEGA to illustrate the relationship between PamUGT and other glycosyltransferases that catalyze oleanane-type triterpenes from various plants (Table S6).

5.5. Biochemical properties investigation

The optimal temperature for PamUGT was determined by conducting reactions for 2 h at temperatures ranging from 4 to 60 °C. The reaction mixture (100 μL) contained 50 μg of purified enzyme, 0.2 mmol·L^{-1} phytolaccagenin (**1**), 0.5 mmol·L^{-1} UDP-Glc, and 50 mmol·L^{-1} Tris-HCl buffer (pH 7.4). After incubation, the mixture was centrifuged at 12 000 r·min^{-1} for 20 min, and the supernatant was analyzed by UPLC/Q-TOF-MS.

Using the optimal temperature, the optimal pH of PamUGT was determined over pH range of 4.0 to 11.0 using the following buffer systems, citric acid-sodium citrate buffer (pH 4.0–6.0), Na_2HPO_4–NaH_2PO_4 buffer (pH 6.0–8.0), Tris-HCl buffer (pH 7.0–9.0), and $NaHCO_3$–Na_2CO_3 buffer (pH 9.0–11.0).

To test the divalent cation dependence of PamUGT, reactions were performed under optimal conditions in the presence of 10 mmol·L^{-1} of various additives: Ba^{2+}, Ca^{2+}, Co^{2+}, Zn^{2+}, Mn^{2+}, Fe^{2+}, Cu^{2+}, Mg^{2+}, Ni^{2+}, and EDTA.

5.6. Kinetic parameters determination

The kinetic parameters of PamUGT were determined in a 100 μL reaction system containing 50 μg of purified enzyme, 1 mmol·L^{-1} UDP-Glc, 50 mmol·L^{-1} Tris-HCl buffer (pH 7.4), and substrates at different concentrations (ranging from 20 to 1000 μmol·L^{-1}). The reactions were conducted at 50 °C for 2 h. The conversion rate was calculated, and the K_m value of PamUGT for substrates was subsequently determined using the Lineweaver-Burk plot.

5.7. Glycosyltransferase activity assays

To explore the substrate promiscuity of PamUGT, various sugar acceptors were tested, including triterpenoids (**1**–**20**), diterpenoid (**21**), flavonoids (**22**–**38**), phenolic compound (**39**), alkaloid (**40**), coumarin (**41**), and steroid (**42**). The reactions were individually carried out in a total volume of 100 μL containing 50 μg of purified PamUGT, 0.2 mmol·L^{-1} aglycone, 0.5 mmol·L^{-1} UDP-Glc, and 50 mmol·L^{-1} Tris-HCl buffer (pH 7.4). All reactions were incubated at 50 °C overnight and terminated by adding 200 μL of ice-cold methanol. The mixtures were centrifuged at 12 000 r·min^{-1} for 20 min, and the supernatants were analyzed using UPLC/Q-TOF-MS.

To investigate the sugar donor selectivity of PamUGT, five

common sugar donors (UDP-Glc, UDP-Xyl, UDP-GlcA, UDP-Rha, and UDP-Gal) were evaluated using compounds **1**, **2**, and **9** as acceptors. The reactions were performed in a total volume of 100 μL containing 50 μg of purified PamUGT, 0.2 mmol·L^{-1} aglycone, 0.5 mmol·L^{-1} UDP-sugar, and 50 mmol·L^{-1} Tris-HCl buffer (pH 7.4). All reactions were incubated at 50 °C overnight, terminated with 200 μL of ice-cold methanol, and analyzed by UPLC/Q-TOF-MS.

5.8. UPLC-MS analysis

Analysis of the reaction mixtures was conducted *via* UPLC/Q-TOF-MS employing an Agilent ZORBAX SB-C_{18} column (2.1 mm × 150 mm, 1.8 μm; Agilent, USA). The separation was achieved at a flow rate of 0.3 mL·min^{-1} with a mobile phase composed of 0.1% (*v/v*) formic acid in water (A) and methanol (B). Mass spectrometric detection was operated in full scan mode under negative electrospray ionization (ESI). The mobile phase gradient program is detailed in Table S4.

5.9. Large-scale preparation and structural characterization of glycosylated products

The substrate was dissolved in DMSO to a final concentration of 20 mmol·L^{-1}, while UDP-Glc was dissolved in deionized water to a final concentration of 50 mmol·L^{-1}. The reaction system contained the substrate at a final concentration of 0.2 mmol·L^{-1}, UDP-Glc at 0.5 mmol·L^{-1}, and 500 μg of purified enzyme; the total volume was adjusted to 1 mL with buffer (50 mmol·L^{-1} Tris-HCl buffer, pH 7.4), and the reaction was conducted at 50 °C for 6 h. The reaction was terminated by adding 2 mL of ice-cold methanol, followed by centrifugation at 12 000 r·min^{-1} for 20 min. The supernatant was collected and concentrated to a final volume of 1 mL in methanol. Subsequently, the glycosylated products were isolated *via* reversed-phase semi-preparative liquid chromatography (SHIMADZU, China) with an Agilent ZORBAX SB-C_{18} column (9.4 mm × 250 mm, 5 μm; Agilent, USA). and verified by UPLC/Q-TOF-MS and NMR analyses.

Funding

This study was supported by the National Natural Science Foundation of China (No. 82304667) and the Science and Technology Commission of Shanghai Municipality (No. Shanghai Sailing Program 23YF1457800).

Supporting information

The supporting information for this article, including NMR spectra of products **1a**, **2a**, and **9a**, TIC spectra of the PamUGT-catalyzed reaction of compounds **3**–**42**, and UPLC/Q-TOF-MS data of **4a**–**42a**, can be obtained by contacting the corresponding authors *via* E-mail.

Declaration of competing interest

All authors declare that there are no conflicts of interest.

References

1 Sirirungruang S, Barnum CR, Tang SN, et al. Plant glycosyltransferases for expanding bioactive glycoside diversity. *Nat Prod Rep*. 2023;40(7):1170-1180. https://doi.org/10.1039/d2np00077f.
2 Li Y, Wang J, Li L, et al. Natural products of pentacyclic triterpenoids: from discovery to heterologous biosynthesis. *Nat Prod Rep*. 2023;40(8):1303-1353. https://doi.org/10.1039/d2np00063f.
3 Yao W, Niu T, Pan J, et al. Discovery and functional characterization of two novel glycosyltransferases associated with the biosynthesis of α-hederin in *Dipsacus asperoides*. *Plant Physiol Biochem*. 2024;217:109273. https://doi.org/10.1016/j.plaphy.2024.109273.
4 Liu Y, Zhao X, Gan F, et al. Complete biosynthesis of QS-21 in engineered yeast. *Nature*. 2024;629(8013):937-944. https://doi.org/10.1038/s41586-024-07345-9.
5 Jo S, El-Demerdash A, Owen C, et al. Unlocking saponin biosynthesis in soapwort. *Nat Chem Biol*. 2025;21(2):215-226. https://doi.org/10.1038/s41589-024-01681-7.
6 Zhang P, Zhang Z, Zhang L, et al. Glycosyltransferase GT1 family: phylogenetic distribution, substrates coverage, and representative structural features. *Comput Struct Biotechnol J*. 2020;18:1383-1390. https://doi.org/10.1016/j.csbj.2020.06.003.
7 Wen C, Huang W, Zhu XL, et al. UGT74AN1, a permissive glycosyltransferase from *Asclepias curassavica* for the regiospecific steroid 3-*O*-glycosylation. *Org Lett*. 2018;20(3):534-537. https://doi.org/10.1021/acs.orglett.7b03619.
8 Lacchini E, Venegas-Molina J, Goossens A. Structural and functional diversity in plant specialized metabolism signals and products: the case of oxylipins and triterpenes. *Curr Opin in Plant Biol*. 2023;74:102371. https://doi.org/10.1016/j.pbi.2023.102371.
9 Kundu P, Thorat SA, Kiran A, et al. Plant-derived triterpenoid saponins: multifaceted roles and bioengineering prospects. *Plant Cell Rep*. 2025;44(8):172. https://doi.org/10.1007/s00299-025-03483-0.
10 Yang C, Qian C, Zheng W, et al. Ginsenoside Rh2 enhances immune surveillance of natural killer (NK) cells *via* inhibition of ERp5 in breast cancer. *Phytomedicine*. 2024;123:155180. https://doi.org/10.1016/j.phymed.2023.155180.
11 Zhang L, Hu Z, Yang L, et al. Saikosaponin a promotes neutrophil extracellular trap formation and bactericidal activity. *Nat Prod Res*. 2025;39(16):4592-4599. https://doi.org/10.1080/14786419.2024.2343918.
12 Li J, Xue M, Xu C, et al. Saikosaponin C as an antiviral agent inhibits the largemouth bass ranavirus *in vitro* and *in vivo*. *Aquaculture*. 2024;585:740632. https://doi.org/10.1016/j.aquaculture.2024.740632.
13 Yang H, Chen Y, Yu L, et al. Esculentoside A exerts anti-inflammatory activity in microglial cells. *Int Immunopharmacol*. 2017;51:148-157. https://doi.org/10.1016/j.intimp.2017.08.014.
14 Atanasov AG, Zotchev SB, Dirsch VM, et al. Natural products in drug discovery: advances and opportunities. *Nat Rev Drug Discov*. 2021;20(3):200-216. https://doi.org/10.1038/s41573-020-00114-z.
15 Shoda SI, Uyama H, Kadokawa JI, et al. Enzymes as green catalysts for precision macromolecular synthesis. *Chem Rev*. 2016;116(4):2307-2413. https://doi.org/10.1021/acs.chemrev.5b00472.
16 Bailly C. Medicinal properties and anti-inflammatory components of *Phytolacca* (Shanglu). *Digit Chin Med*. 2021;4(3):159-169. https://doi.org/10.1016/j.dcmed.2021.09.001.
17 Saleri FD, Chen G, Li X, et al. Comparative analysis of saponins from different *Phytolaccaceae* species and their antiproliferative activities. *Molecules*. 2017;22(7):1077. https://doi.org/10.3390/molecules22071077.
18 Zeng MS, Yu WD, Wang HX, et al. A potential antiviral activity of esculentoside A against binding interactions of SARS-COV-2 spike protein and angiotensin converting enzyme 2 (ACE2). *Int J Biol Macromol*. 2021;183:2248-2261. https://doi.org/10.1016/j.ijbiomac.2021.06.017.
19 Bailly C, Vergoten G. Esculentosides: insights into the potential health benefits, mechanisms of action and molecular targets. *Phytomedicine*. 2020;79:153343. https://doi.org/10.1016/j.phymed.2020.153343.
20 Iwakiri T, Mase S, Murakami T, et al. Glucosylation of hydroxyflavones by glucosyltransferases from *Phytolacca americana*. *J Mol Cata B Enzym*. 2013;90:61-65. https://doi.org/10.1016/j.molcatb.2013.01.016.
21 Liu D, Chen L, Chen C, et al. Full-length transcriptome analysis of *Phytolacca americana* and its congener *P. icosandra* and gene expression normalization in three *Phytolaccaceae* species. *BMC Plant Biol*. 2020;20(1):396. https://doi.org/10.1186/s12870-020-02608-9.
22 Zhu YY, Qian LW, Zhang J, et al. New approaches to the structural modification of olean-type pentacylic triterpenes *via* microbial oxidation and glycosylation. *Tetrahedron*. 2011;67(23):4206-4211. https://doi.org/10.1016/j.tet.2011.04.055.
23 He J, Chen K, Hu ZM, et al. UGT73F17, a new glycosyltransferase from *Glycyrrhiza uralensis*, catalyzes the regiospecific glycosylation of pentacyclic triterpenoids. *Chem Commun (Camb)*. 2018;54(62):8594-8597. https://doi.org/10.1039/c8cc04215b.
24 Huang W, Zhang X, Li J, et al. Substrate promiscuity, crystal structure, and application of a plant UDP-glycosyltransferase UGT74AN3. *ACS Catal*. 2024;14(1):475-488. https://doi.org/10.1021/acscatal.3c05309.
25 Gantt RW, Goff RD, Williams GJ, et al. Probing the aglycon promiscuity of an engineered glycosyltransferase. *Angew Chem Int Ed Engl*. 2008;47(46):8889-8892. https://doi.org/10.1002/anie.200803508.
26 Shibuya M, Nishimura K, Yasuyama N, et al. Identification and characterization of glycosyltransferases involved in the biosynthesis of soyasaponin I in *Glycine max*. *FEBS Lett*. 2010;584(11):2258-2264. https://doi.org/10.1016/j.febslet.2010.03.037.
27 Zhang L, Ren S, Liu X, et al. Mining of UDP-glucosyltransferases in licorice for controllable glycosylation of pentacyclic triterpenoids. *Biotechnol Bioeng*. 2020;117(12):3651-3663. https://doi.org/10.1002/bit.27518.
28 Tian L, Blount JW, Dixon RA. Phenylpropanoid glycosyltransferases from osage orange (*Maclura pomifera*) fruit. *FEBS Lett*. 2006;580(30):6915-6920. https://doi.org/10.1016/j.febslet.2006.11.056.
29 Huang J, Lei T, Zhou Q, et al. Comparative metabolome and transcriptome analysis revealed the accumulative mechanism of rubusoside in Chinese sweet tea. *J Agric Food Chem*. 2024;72(44):24539-24551. https://doi.org/10.1021/acs.jafc.4c07127.
30 Noguchi A, Kunikane S, Homma H, et al. Identification of an inducible glucosyltransferase from *Phytolacca americana* L. cells that are capable of glucosylating capsaicin. *Plant Biotechnol J*. 2009;26(3):285-292. https://doi.org/10.5511/plantbiotechnology.26.285.
31 Feng J, Zhang P, Cui Y, et al. Regio- and stereospecific *O*-glycosylation of phenolic compounds catalyzed by a fungal glycosyltransferase from *Mucor hiemalis*. *Adv Synth Catal*. 2017;359(6):995-1006. https://doi.org/10.1002/adsc.201601317.
32 Chen S, Yao H, Han J, et al. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One*. 2010;5(1):e8613. https://doi.org/10.1371/journal.pone.0008613.