Identification of α-Amyrin 28-Carboxylase and Glycosyltransferase From *Ilex asprella* and Production of Ursolic Acid 28-O-β-D-Glucopyranoside in Engineered Yeast

Xiaoyu Ji\(^1\,^2\,^3\,^4\), Shumin Lin\(^1\,^2\,^3\), Yuanyuan Chen\(^1\,^2\,^3\), Jiawei Liu\(^1\,^2\,^3\), Xiaoyun Yun\(^1\,^2\,^3\), Tiancheng Wang\(^1\,^2\,^3\), Jialiang Qin\(^1\,^2\,^3\), Chaoquan Luo\(^1\,^2\,^3\), Kui Wang\(^1\,^2\,^3\), Zhongxiang Zhao\(^5\), Ruoting Zhan\(^1\,^2\,^3\,*\) and Hui Xu\(^1\,^2\,^3\,*\)

\(^1\) Research Center of Chinese Herbal Resource Science and Engineering, Guangzhou University of Chinese Medicine, Guangzhou, China, \(^2\) Key Laboratory of Chinese Medicinal Resource from Lingnan (Guangzhou University of Chinese Medicine), Ministry of Education, Guangzhou, China, \(^3\) Joint Laboratory of National Engineering Research Center for the Pharmaceutics of Traditional Chinese Medicines, Guangzhou, China, \(^4\) Department of Pharmacology, Shantou University Medical College, Shantou, China, \(^5\) School of Pharmaceutical Sciences, Guangzhou University of Chinese Medicine, Guangzhou, China

*Ilex asprella* is a medicinal plant that is used extensively in southern China. The plant contains ursane-type triterpenoids and triterpenoid saponins which are known to be responsible for its pharmacological activities. Previously, a transcriptomic analysis of *I. asprella* was carried out and the gene *IaAS1*, which is important in the formation of the core structure α-amyrin, was identified. However, the genes related to the subsequent derivatization of the core structures of the triterpenoid remain largely unknown. Herein, we describe the cloning and functional characterization of an amyrin 28-carboxylase *IaAO1* (designated as *IaCYP716A210*) and a glycosyltransferase *IaAU1* (designated as *UGT74AG5*), based on transcriptomic data. The expression of *IaAO1* in an α-amyrin producing yeast strain led to the accumulation of ursolic acid. An enzyme assay using recombinant protein *IaAU1* purified from *E. coli* revealed that *IaAU1* can catalyze the conversion of ursolic acid to ursolic acid 28-O-β-D-glucopyranoside. *IaAU1* has regiospecificity for catalyzing the 28-O-glucosylation of ursane-/oleanane-type triterpene acids, as it can also catalyze the conversion of oleanolic acid, hederagenin, and ilexgenin A to their corresponding glycosyl compounds. Moreover, co-expression of *IaAO1* and *IaAU1* in the α-amyrin-producing yeast strain led to the production of ursolic acid 28-O-β-D-glucopyranoside, although in relatively low amounts. Our study reveals that *IaAO1* and *IaAU1* might play a role in the biosynthesis of pentacyclic triterpenoid saponins in *I. asprella* and provides insights into the potential application of metabolic engineering to produce ursane-type triterpene glycosides.

**Keywords**: *Ilex asprella*, Triterpenoid saponins, Biosynthesis, CYP, UGT
INTRODUCTION

Triterpenoids and triterpenoid glycosides constitute a major class of plant secondary metabolites, which are thought to be involved in defense against pathogens and pests (Singh and Sharma, 2015). Compounds such as ginsenosides and glycyrrhizic acid have also been shown to possess health benefits in humans (Kim et al., 2017). However, access to these compounds is limited due to their low levels in plants and difficulties in their purification and chemical synthesis. Unraveling the biosynthetic pathways used for their production might provide the possibility of improving their availability through synthetic biology (Zhou et al., 2015).

Over the past several decades, tremendous interest and progress in the understanding of the biosynthesis of triterpenoids and triterpenoid glycosides have been observed. Generally, triterpenoid glycoside is assembled from six isoprene units followed by cyclization and scaffold modifications. The cyclization reaction mediated by oxidosqualene cyclases (OSC) is the first diversifying step in the biosynthetic pathway. So far, more than 100 triterpene scaffolds have been reported, primarily including lupane, dammarane, oleanane (derived from α-amyrin), and ursane (derived from α-amyrin) (Shang and Huang, 2019). Recently, a novel triterpene oysatinol was identified, which even widen the potential scope of triterpene scaffolds that could exist in nature (Xue et al., 2018; Stephenson et al., 2019). The subsequent site-specific oxidation and glycosylation of the cyclic scaffold are catalyzed by cytochrome P450 monooxygenases (CYPs) and UDP-dependent glycosyltransferases (UGTs), respectively, conferring further structural and functional diversity.

Both CYPs and UGTs belong to multigene families and are involved in numerous metabolic processes including those related to triterpenoid saponins. To date, several plant CYPs have been functionally characterized and their diverse roles in triterpene scaffold modification have been reviewed (Miettinen et al., 2017). Members from different classes (e.g., CYP51, CYP71, CYP72, CYP85) have been shown to be associated with triterpene scaffold oxidation. Moreover, the reactions catalyzed by CYPs are extremely diverse, including desaturation, oxidation, and C-C bond cleavage. In the case of UGTs, a few enzymes have been identified, including members within the UGT71, UGT73, UGT74, UGT85, UGT91, and UGT94 families (Rahimi et al., 2019). They catalyze versatile glycosylation reactions that result in variations in the number of sugar chains, composition, and position on the triterpene scaffold (Seki et al., 2015; Xu et al., 2016; de Costa et al., 2017; He et al., 2018). Considering the huge diversity of triterpene-related CYPs and UGTs, as well as the fact that the majority of plant triterpene compounds are biosynthesized in species-specific manner, it is of interest to isolate and characterize more triterpene tailoring enzymes CYPs and UGTs from various plant species to extend our knowledge of triterpene metabolism and for utilization of these enzymes.

Ilex asprella is a medicinal plant that originates from southern China and its root is usually used to treat influenza and pharyngitis. I. asprella contains a wide range of triterpenoids and related saponins which possess various bioactivities such as anti-inflammatory, anticancer, and antiviral activities. Most Ilex triterpenoids are of the ursane-type and are derived from multiple modifications of α-amyrin (Figure 1). Oxidative modification occurs most commonly at positions C-19, C-24, and C-28 of the ursane skeleton, while glycosylation occurs at positions C-3 and C-28 (Zhou et al., 2012; Peng et al., 2016; Wen et al., 2017). Although the pharmaceutical and physiological effects of these triterpenoids and related saponins are well-known, our understanding of their biosynthesis in Ilex asprella remains limited.

Previously, we have obtained the transcriptome of I. asprella using RNA-sequencing (GenBank accession number SRP035767). Analysis of the transcriptome revealed several OSC, CYP, and UGT genes that could be potentially involved in triterpenoid biosynthesis pathway. Among these, two triterpene cyclases, IaAS1 and IaAS2, which catalyze the cyclization of 2,3-oxidosqualene to form α- and β-amyrin in different ratios, have been identified (Zheng et al., 2015). However, the enzymes involved in the oxidation and glycosylation steps of triterpene biosynthesis in I. asprella remain largely unknown. In this study, we report the identification of IaAO1 (named as IaCYP716A210) which can catalyze the C-28 carboxylation of α-amyrin, and a UDP-glycosyltransferase IaAU1 (named as UGT74AG5) that has regiospecificity for catalyzing the 28-O-glucosylation of ursane-/oleanane-type triterpene acid. Furthermore, we successfully co-expressed IaAO1 and IaAU1 in yeast carrying IaAS1, resulting in the production of an unusual glycoside ursolic acid 28-O-β-D-glucopyranoside.

MATERIALS AND METHODS

Sequence Analysis

Complete amino acid sequences of CYPs and UGTs known were collected from NCBI1 for analysis. Multiple sequence alignments were performed using the software Clustal Omega2. The phylogenetic tree was constructed using the maximum likelihood method with Molecular Evolutionary Genetics Analysis Program (MEGA7.0) (Sudhir et al., 2016). A bootstrap analysis with 1,000 replicates was used to assess the strength of the nodes in the tree (Felsenstein, 1985).

cDNA Preparation and Cloning of IaAO1 and IaAU1

Total RNA was extracted from the leaves of 2-year-old I. asprella using a HiPure Plant RNA Mini Kit (Magen, Guangzhou, China). The polyadenylated RNA was reverse transcribed into cDNA using a TransScript II All-in-One First-Strand cDNA Synthesis SuperMix (TransGen Biotech, Beijing, China) according to the manufacturer’s protocol. Using cDNA as the template, the coding regions of IaAO1 and IaAU1 were amplified using the Primer STAR high-fidelity DNA polymerase (Takara, Dalian, China).

---

1https://www.ncbi.nlm.nih.gov/nuccore/
2https://www.ebi.ac.uk/Tools/msa/clustalo/
The PCR products obtained were purified and ligated into the vector pEASY-T5 and subsequently recombinant plasmids were used to transform E. coli Trans1-T1 competent cells using a pEASY-T5 Zero Cloning Kit (TransGen Biotech, Beijing, China). Both of the recombinant plasmids were verified by sequencing. All the sequences of the primers used in this study are shown in Supplementary Table S1. Detail information about strains and plasmids is listed in Supplementary Table S2.

**Heterologous Expression of IaAO1 in Yeast**

Two pairs of In-Fusion primers were designed based on the sequences of IaAO1 and the yeast expression vector pESC-TRP. After amplification, the coding region of IaAO1 was ligated into the vector pESC-TRP at the EcoRI and SpeI sites using the In-Fusion HD Cloning Kit (Takara, Dalian, China). The plasmid obtained was named pTlAOI. Using a standard lithium acetate protocol, pTlAOI was transformed into an α-amyrin-producing strain of Saccharomyces cerevisiae WAT11fAX (a WAT11-derived yeast strain with an integration of IaAS1 in the genome created using CRISPR/Cas9, unpublished data). After 16-h of growth in SC-T media containing 2% glucose at 30°C, the transformed yeast cells were washed three times with sterile water, re-suspended in SC-T media containing 2% galactose and allowed to grow for 48 h. An equivalent number of yeast cells were harvested at different time points and extracted for total protein. The target protein was identified using an anti-His mouse monoclonal antibody (TransGen Biotech, Beijing, China) by western blotting. Yeast transformed with the empty vector pESC-TRP was used as a negative control. The positive transformants were selected and incubated for 7-day. Cell metabolites were extracted, derivatized and analyzed by GC-MS according to the method established previously (Qin et al., 2019).

**Expression and Purification of Recombinant IaAU1**

Similarly, IaAU1 was ligated into the Escherichia coli expression vector pET32a (+) at the EcoRV and SacI sites using In-Fusion cloning. The pET-IaAU1 plasmid obtained was used to transform E. coli Rosetta (DE3) cells. Transformants were cultured in LB media with appropriate antibiotics and induced with 0.1 mM IPTG. After harvesting the cells, total proteins were extracted and the recombinant protein was purified by Ni²⁺-NTA chromatography (Qiagen, Germany). SDS-PAGE was performed to assess the expression levels and purity of the recombinant IaAU1.

**IaAU1 Enzyme Assay**

IaAU1 activity was measured in a final volume of 200 µL of a buffer consisting of 50 mM Tris (pH 7.5), 10 mM MgCl₂, 1 mM DTT, 20–60 µg purified IaAU1, 250 µM sugar donor UDP-glucose (UDP-Glc), and 1 mM sugar acceptor (Meesapyodsuk et al., 2007; Naoumkina et al., 2010; de Costa et al., 2017). Six different triterpene sapogenins, ursolic acid (Urs), ilexgenin A (Ilex), oleanolic acid (Ole), hederagenin (Hed), glycyrrhetic acid (Gly), and soyasapogenol B (Soy), were used as sugar acceptors. The reaction was carried out at 30°C for 30 min, and then stopped by adding two volumes of ethyl acetate. The ethyl acetate phase was removed, evaporated to a volume of about 100 µL and analyzed by thin-layer chromatography (TLC). The TLC plate was developed using C₆H₁₄:CH₃COOC₂H₅:CH₃COOH (1:12:0.5) as the mobile phase and visualized by spraying with 20% sulfuric acid in absolute ethyl alcohol followed by heating at 105°C for 3 min. The exact mass to charge ratio (m/z) was measured with a high-resolution mass spectrometer (Orbitrap Fusion™ Tribird™, Thermo Fisher, San Jose, CA, United States). Furthermore, the assay with Urs as the sugar donor was performed under the same conditions to determine the specificity of IaAU1.
acceptor was carried out at a preparative scale and the product generated was isolated using a Sephadex™ LH-20 column. The purified product was structurally characterized using MS and NMR analyses.

**Co-expression of IaAO1 and IaAU1 in Yeast**

Two separate yeast strains carrying both IaAO1 and IaAU1 genes were constructed. First, IaAO1 was ligated into the yeast expression vector pESC-URA using In-Fusion Cloning. The resulting plasmid pUlaAO1 was used to transform the abovementioned yeast strain, WAT11fAX expressing IaAO1, to produce the strain WAT11S1. Alternatively, IaAO1 and IaAU1 were sub-cloned into the yeast expression vector p426GPD to create the expression cassettes P\_GAP\_IaAO1-T\_CYC1 and P\_GAP\_IaAU1-T\_CYC1, respectively. The P\_GAP\_IaAO1-T\_CYC1 cassette was then integrated into the ade2 locus of strain WAT11fAX via CRISPR/Cas9 while P\_GAP\_IaAU1-T\_CYC1 was subsequently integrated into the bts1 locus, resulting in strain WAT11S2. WAT11S1 and WAT11S2 were cultured in SC-U-T and YPD media, respectively, at 30°C for 7 days. The cells were then harvested for metabolite extraction. Cell pellets were suspended in 20 mL of sterile water and disrupted using a high pressure homogenizer (20,000 psi, 50 s) (D-6L, PHD Technology LCC, Saint Paul, MN, United States). After extraction with the same volume of \textit{n}-butanol twice, the organic phases were concentrated by evaporation and the residues were reconstituted in methanol and subjected to an LC-MS/MS analysis.

**LC-MS/MS Analysis**

HPLC separation was performed using a Surveyor plus HPLC system (Thermo Fisher, Singapore) equipped with a Hypersil Golden C\textsubscript{18} column (100 mm \times 2.1 mm, 3 \textmu m particles) at 20°C. The mobile phase consisted of water (component A) and methanol (component B) and the flow rate was 0.2 mL min\textsuperscript{-1}. The gradient program was: 0–2.0 min, 30% B; 2.0–5.0 min, 30–80% B; 5.0–13.0 min, 80% B; 13.0–15.0 min, 80–90% B; 15.0–25.0 min, 90% B. HESI-MS/MS data were acquired with a triple quadrupole mass spectrometer (TSQ Quantum Access, Thermo Fisher, United States) coupled with an electrospray source in positive ionization mode. The MS parameters were set as follows: vaporizer temperature 375°C, capillary temperature 20°C, sheath gas 40 (arbitrary units), aux gas 5 (arbitrary units) and spray voltage 3.5 kV. For qualitative and semi-quantitative analysis of metabolites, selected reaction monitoring...
FIGURE 3 | In vivo enzymatic activity assays in yeast. (A) Western blot analysis of IaAO1 expressed in yeast transformants. Empty vector containing yeast was indicated as blank. (B) Total ion chromatograms of ursolic acid (red line), oleanolic acid (green line), WAT11 tfAX-pT IaAO1 (blue line), and blank control WAT11 tfAX-pT (black line). Major peaks were numbered. (C) Mass spectra of peak 1 from the ursolic acid standard. The chemical structure shown was TMS-derivatized ursolic acid. Characteristic fragment ions were highlighted with red frames. (D) Mass spectra of peak 3 from WAT11 tfAX-pT IaAO1.

(SRM) mode was used. The contents of ursolic acid 28-O-β-D-glucopyranoside were determined by calculating relative peak areas using the product prepared from IaAU1 enzymatic activity assay as standard.

RESULTS

Screening of the Candidate Genes IaAO1 and IaAU1 for Triterpene Oxidation and Glycosylation at C-28

In the transcriptomic data of I. asprella, more than 200 transcripts were annotated as CYPs. A phylogenetic analysis revealed a putative CYP gene (IaAO1) that was closely clustered with CYP716AL1 (Huang et al., 2012), which was identified from Catharanthus roseus as a multifunctional C-28 oxidase capable of converting α-amyrin, β-amyrin, and lupeol to ursolic, oleanolic, and betulinic acids, respectively (Figure 2A). Sequence analysis of IaAO1 revealed an open reading frame of 1,443 bp, encoding a protein of 481 aa. The sequence similarity between IaAO1 and CYP716AL1 was high, being 82%. Therefore, IaAO1 is very likely a triterpene C-28 oxidase. A similar process was applied to screen for triterpene-related UGT candidates. A putative UGT gene (IaAU1) was closely clustered with the C-28 glycosyltransferase UGT74M1 from Saponaria vaccaria (Figure 2B). IaAU1 contains an open reading frame of 1,368 bp, encoding a 459 aa protein with a predicted molecular mass of 50.5 kDa. In addition, IaAU1 contains the plant secondary product glycosyltransferase (PSPG) domain and had a 58.75% similarity with UGT74M1. Therefore, it is plausible that IaAU1 encodes a triterpene C-28 glycosyltransferase.

Functional Characterization of Amyrin C-28 Oxidase IaAO1

In order to elucidate the function of IaAO1, it was cloned and expressed in the α-amyrin-producing yeast strain WAT11 tfAX.
Western blot analysis of the total protein extracted from the cells showed that IaAO1 was successfully expressed after induction with 2% galactose (Figure 3A).

Culture extracts from the WAT11tfAX expressing IaAO1 yeast were derivatized with trimethylsilylating agents and submitted to GC-MS analysis (Figure 3B). Authentic ursolic acid showed a dominant peak at 17.25 min. A peak with the corresponding retention time was observed in the total ion chromatogram of WAT11tfAX expressing IaAO1 yeast, and mass spectra confirmed it was ursolic acid (Figures 3C,D). These results demonstrate that IaAO1 from I. asprella catalyzes oxidation at the C-28 position of α-amyrin to yield ursolic acid. The sequence data of IaAO1 have been submitted to GenBank with the accession number of MK994507.

**In vitro Functional Characterization of C-28 Glycosyltransferase IaAU1**

To determine the function of IaAU1, the gene was expressed in E. coli Rosetta (DE3). SDS-PAGE analysis of total protein, soluble protein, and the purified protein showed a unique band at 68 kDa corresponding to the predicted size of recombinant protein (618 amino acids containing multiple
tags from vector) (Supplementary Figure S1). Enzyme activity assays were carried out with the sugar donor UDP-Glc and six triterpene sapogenins, respectively. Primary detection using TLC showed new products were formed with ursane-type triterpenoids (ursolic acid, ilexgenin A) and oleanane-type ones (oleanolic acid, hederagenin), but not with oleanane-type ones without a carboxyl group at C-28 position (glycyrrhetic acid and soyasapogenol B) (Figure 4). The exact mass to charge ratio of assay product was consistent with the expected monoglucosylation molecular ion, i.e., m/z 641.4023 (C_{36}H_{58}O_{8}Na^{+}, [Urs-Glc]), m/z 641.4014 (C_{36}H_{58}O_{8}Na^{+}, [Ole-Glc]), m/z 657.3968 (C_{38}H_{58}O_{9}Na^{+}, [Hed-Glc]), m/z 687.3724 (C_{38}H_{56}O_{11}Na^{+}, [Ilex-Glc]) (Supplementary Figure S2). Therefore, IaAU1 is deduced to be a glycosyltransferase that can transfer a glucosyl group to the C-28 carboxyl moiety of ursane- or oleanane-type triterpene acids to produce an ester.

FIGURE 5 | Mass spectra of IaAU1 assay product of ursolic acid. Chemical structures of monoglycoside (m/z 641.4026), aglycone (m/z 479.3497) and sugar (m/z 185.0522) were shown upon the corresponding MS fragments.

FIGURE 6 | LC-MS/MS SRM mode analysis of ursolic acid 28-O-β-D-glucopyranoside. Purified ursolic acid 28-O-β-D-glucopyranoside together with the crude enzyme reaction mixture of IaAU1 was used as control. SRM chromatograms of characteristic transitions m/z 641.2→m/z 479.0 was colored in red while m/z 641.2→m/z 185.0 was in green.
To further ascertain the regioselectivity of IaAU1, the product derived from ursolic acid was chosen for structural elucidation. It was isolated in preparative amounts and subjected to MS/MS, and $^1$H- and $^{13}$C-NMR analyses (Figure 5, Supplementary Figure S3, and Table S3). The MS/MS spectrum clearly demonstrated the characteristic fragment ions of ursolic saponin, namely an aglycone at $m/z$ 479.3497 (C$_{30}$H$_{48}$O$_{3}$Na$^+$) and a sugar at $m/z$ 185.0422 (C$_6$H$_{10}$O$_5$Na$^-$). Compared to the parent compound, the NMR spectra of the product showed additional signals for a glucose moiety, especially $\delta$H 5.36 (1H, d, J = 8.0Hz) and $\delta$C 95.3, indicating the substitution of a glucose at position C-28. As a result, the product was determined to be ursolic acid 28-O-β-D-glucopyranoside, which confirmed the proposed function of IaAU1. IaAU1 therefore catalyzes the glycosylation of ursane- or oleanane-type triterpene acids at the C-28 position and has been designated as UGT74AG5 by the UDP-glycosyltransferase (UGT) Nomenclature Committee. The sequence data of IaAU1 can be accessed in GenBank under the accession number MK994508.

**DISCUSSION**

CYPs and UGTs have been demonstrated to be two classes of key enzymes responsible for structural diversity in their biosynthesis. In this study, we identified a cytochrome P450 IaAO1 and a glycosyltransferase UGT74AG5 from *I. asprella*. The biochemical functions of IaAO1 and UGT74AG5 indicated the possible involvement of both genes in the biosynthesis of ursane-type triterpenoids and triterpenoid saponins.

---

**Co-expression of IaAO1 and IaAU1 and Production of Ursolic Acid 28-O-β-D-Glucopyranoside**

To confirm the activity of IaAO1 and IaAU1, both genes were expressed simultaneously in *S. cerevisiae* WAT11tfAX and the metabolites were analyzed to look for the expected product ursolic acid 28-O-β-D-glucopyranoside. Two different yeast strains (WAT11S1 and WAT11S2) were constructed by either transformation of two separate expression plasmids or by integration of two genes into the genome. Analysis of yeast metabolites by LC-MS/MS in selected reaction monitoring (SRM) mode revealed two characteristic transitions ($m/z$ 641.2 [Urs-Glc + Na]$^+$ $\rightarrow$ $m/z$ 479.0 [Urs + Na]$^+$, $m/z$ 641.2 [Urs-Glc + Na]$^+$ $\rightarrow$ $m/z$ 185.0 [Glc + Na-H$_2$O$^-$]) that were clearly present in both of the engineered strains, which was consistent with assay product ursolic acid 28-O-β-D-glucopyranoside (Figure 6). However, while WAT11S1 produced up to 27.4 g/L of Urs-Glc, only a trace amount of Urs-Glc (0.5 g/L) was detected in WAT11S2. In addition, an unknown peak was observed in the transition $m/z$ 641.2 $\rightarrow$ $m/z$ 479.0 of WAT11S2 at about 14 min, indicating the production of some other unexpected compound.

---

*https://prime.vetmed.wsu.edu/resources/udp-glucuronsyltransferase-homepage*
Phylogenetically, IaAO1 belongs to the CYP716A subfamily. Most of the characterized CYP716As catalyze the C-28 oxidation of pentacyclic triterpene scaffolds and a large number of them show activity on multiple substrates (Miettinen et al., 2017). IaAO1 was proven to be an ordinary CYP716A member, catalyzing the oxidation of α-amyrin at C-28. But its substrate specificity has not been investigated herein. Very recently, we isolated the gene encoding IpAO1 (an ortholog of CYP716A210, D. Nelson, personal communication, July 11, 2018) from I. pubescens, which has been shown to catalyze the oxidation of both α- and β-amyrin to give ursolic acid and oleanolic acid, respectively (Qin et al., 2019). Since IaAO1 shares 98% sequence identity and 99% sequence similarity with IpAO1, IaAO1 should also be an ortholog of CYP716A210 and must also catalyze the oxidation of β-amyrin at C-28. In accordance, IaAO1 was named as IaCYP716A210.

Glycosylation commonly occurs at the C-3 or C-28 sites in many bioactive pentacyclic triterpene skeletons. To date, only a handful of triterpene UGTs that modify the carboxyl group at C-28 have been characterized, such as UGT73AD1 from Centella asiatica (de Costa et al., 2017), UGT73F3 from Medicago truncatula (Naoumkina et al., 2010), UGT74M1 from Saponaria vaccaria (Meesapyodskul et al., 2007), UGT73C12 and UGT73C13 from Barbarea vulgaris (Augustin et al., 2012). The newly discovered UGT74AG5 provides insight into glycosylation of pentacyclic triterpenoid saponin biosynthesis. An in vitro enzymatic assay implied that UGT74AG5 has low substrate specificity but high regiospecificity, which has also been observed with many other UGTs (Rahimi et al., 2019). UGT74AG5 can catalyze the conversion of various ursane-type and oleanane-type carboxylic acids, but it cannot accept substrates without a carboxyl group at C-28. Moreover, MS and NMR analyses of the enzyme product confirmed the substitution of glucose at C-28 in the β-configuration.

After functional characterization of IaCYP716A210 and UGT74AG5, they were co-expressed in a yeast strain that had been pre-transformed with the mixed amyrin synthase IaAS1, which led to the production of ursolic acid 28-β-D-glucopyranoside as the natural substrate of UGT74AG5 or if ursolic acid 28-O-β-D-glucopyranoside is a biosynthetic intermediate and subject to further structural modification in the plant cell. To answer these questions, further studies should be carried out.

**DATA AVAILABILITY STATEMENT**

The sequences of both genes identified in this study can be retrieved in GenBank with the accession Nos. MK994507 and MK994508.

**AUTHOR CONTRIBUTIONS**

HX and XJ designed the experimental. XJ and HX wrote the manuscript. XJ, SL, YC, and TW performed the major experiments including gene cloning, enzyme expression and functional characterization. JL contributed to the NMR analysis. ZZ provided the substrate ilexgenin A and helpful guidance. KW, XY, JQ, and CL constructed the yeast strain WAT11tfAX. HX and RZ supervised the entire project. All authors have read and approved the manuscript.

**FUNDING**

This work was funded by Natural Science Foundation of Guangdong Province, China (Project No. 2015A030313347) and Guangdong Key Research and Development Program (Program No. 2020B020221001).

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00612/full#supplementary-material

---

**REFERENCES**

Augustin, J. M., Drok, S., Shinoda, T., Sanmiya, K., Nielsen, J. K., Khakimov, B., et al. (2012). UDP-glycosyltransferases from the UGT73C subfamily in Barbarea vulgaris catalyze sapogenin 3-O-glycosylation in saponin-mediated insect resistance. Plant Physiol. 60, 1881–1895. doi: 10.1104/pp.112.202747

de Costa, F., Barber, C. J. S., Kim, Y. B., Reed, D. W., Zhang, H., Fett-Neto, A. G., et al. (2017). Molecular cloning of an ester-forming triterpenoid: UDP-glucose 28-O-glucosyltransferase involved in saponin biosynthesis from the medicinal plant Centella asiatica. Plant Sci. 262, 9–17. doi: 10.1016/j.plantsci.2017.05.009

Du, B. Z., Yang, X. Y., Feng, X., Yin, X., and Chai, X. Y. (2017). A phytochemical and pharmacological advance on ilex asprella. China J. Chinese Mater. Med. 42, 20–28. doi: 10.19540/j.cnki.cjcmm.20161222.025

Felsenstein, J. E. (1985). Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791. doi: 10.1111/j.1558-5646.1985.tb0420x

He, J., Chen, K., Hu, Z. M., Li, K., Song, W., Yu, L. Y., et al. (2018). UGT73F17, a new glycosyltransferase from Glycyrhiza uralensis, catalyzes the regiospecific glycosylation of pentacyclic triterpenoids. Chem. Commun. 54, 8594–8597. doi: 10.1039/C8CC04215B

Huang, L., Li, J., Ye, H., Li, C., Wang, H., Liu, B., et al. (2012). Molecular characterization of the pentacyclic triterpenoid biosynthetic pathway incanthanathus roseus. Planta 236, 1571–1581. doi: 10.1007/s00425-012-1712-0

Kim, J. H., Yi, Y. S., Kim, M. Y., and Cho, J. Y. (2017). Role of ginsenosides, the main active components of Panax ginseng, in inflammatory responses and diseases. J. Ginseng. Res. 41, 433–443. doi: 10.1016/j.jgr.2016.08.004
Li, C., Li, Z. L., Wang, T., and Qian, S. H. (2014). Chemical constituents of the aerial parts of *Lycopus lucidus* var. hirtus. *Chem. Nat. Compd.* 50, 288–290. doi: 10.1007/s10600-014-0933-9

Meesapyodsuk, D., Balsevich, J., Reed, D. W., and Covello, P. S. (2007). Saponin biosynthesis in *Saponaria vaccaria*. cDNAs encoding β-amyrin synthase and a triterpene carboxylic acid glucosyltransferase. *Plant Physiol.* 143, 959–969. doi: 10.1104/pp.106.088484

Miettinen, K., Pollier, J., Buyst, D., Arendt, P., Csuk, R., Sommerwerk, S., et al. (2017). The ancient CYP716 family is a major contributor to the diversification of eudicot triterpenoid biosynthesis. *Nat. Commun.* 8, 14153–14153. doi: 10.1038/ncomms14153

Naoumkina, M. A., Modolo, L. V., Huhman, D. V., Urbanczyk-Wochniak, E., Tang, Y., Sumner, L. W., et al. (2010). Genomic and coexpression analyses predict multiple genes involved in triterpene saponin biosynthesis in *Medicago truncatula*. *Plant Cell* 22, 850–866. doi: 10.1105/tpc.109.073270

Peng, M. H., Dai, W. P., Liu, S. J., Yu, L. W., Wu, Y. N., Liu, R., et al. (2016). Bioactive glycosides from the roots of *Ilex asprella*. *Pharm. Biol.* 54, 2127–2134. doi: 10.3109/13880209.2016.1146779

Qin, J. L., Yun, X. Y., Ji, X. Y., Zhan, R. T., and Xu, H. (2019). Identification and functional analysis of a C-28 oxidase involved in the biosynthesis of triterpenoid saponins from *Ilex asprella*. *Lishizhen Med. Mater. Med. Res.* 30, 53–57.

Rahimi, S., Kim, J., Mijakovic, I., Jung, K., Choi, G., Kim, S. C., et al. (2019). Triterpenoid-biosynthetic UDP-glycosyltransferases from plants. *Biotechnol. Adv.* 37:107394. doi: 10.1016/j.biotechadv.2019.04.016

Seki, H., Tamura, K., and Muranaka, T. (2015). P450s and UGTs: key players in the structural diversity of triterpenoid saponins. *Plant Cell Physiol.* 56, 1463–1471. doi: 10.1093/pcp/pct062

Shang, Y., and Huang, S. (2019). Multi-omics data-driven investigations of metabolic diversity of plant triterpenoids. *Plant J.* 97, 101–111. doi: 10.1111/tpj.14132

Singh, B., and Sharma, R. A. (2015). Plant terpenes: defense responses, phylogenetic analysis, regulation and clinical applications. *J Biotech* 5, 129–151. doi: 10.1007/s13205-014-0220-2

Stephenson, M. J., Field, R. A., and Osbourn, A. (2019). The protosteryl and dammareryl cation dichotomy in polycyclic triterpene biosynthesis revisited: has this ‘rule’ finally been broken? *Nat. Prod. Rep.* 36, 1044–1052. doi: 10.1039/c8np0096d

Sudhir, K., Glen, S., and Koichiro, T. (2016). MEGAT: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874. doi: 10.1093/molbev/msw054

Wen, Q., Lu, Y., Chao, Z., and Chen, D. F. (2017). Anticomplement triterpenoids from the roots of *Ilex asprella*. *Bioorg. Med. Chem. Lett.* 27, 880–886. doi: 10.1016/j.bmcl.2017.01.007

Xu, G., Cai, W., Gao, W., and Liu, C. (2016). A novel glucuronosyltransferase has an unprecedented ability to catalyse continuous two-step glucuronosylation of glycyrrhetinic acid to yield glycyrrhizin. *New Phytol.* 212, 123–135. doi: 10.1111/nph.14039

Zhang, X., Luo, X., Ye, G., Chen, Y., Ji, X., Wen, L., et al. (2015). Characterization of two oxidosqualene cyclases responsible for triterpenoid biosynthesis in *Ilex asprella*. *Int. J. Mol. Sci.* 16, 3564–3578. doi: 10.3390/ijms16023564

Zhou, K., Qiao, K., Edgar, S., and Stephanopoulos, G. J. (2015). Distributing a metabolic pathway among a microbial consortium enhances production of natural products. *Nat. Biotechnol.* 33, 377–383. doi: 10.1038/nbt.3095

Zhou, M., Xu, M., Ma, X. X., Zheng, K., Yang, K., Yang, C. R., et al. (2012). Antiviral triterpenoid saponins from the roots of *Ilex asprella*. *Planta Med.* 78, 1702–1705. doi: 10.1055/s-0032-1315209

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Ji, Lin, Chen, Liu, Yun, Wang, Qin, Luo, Wang, Zhao, Zhan and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.