

Constructing a Biosynthetic Pathway of Ginsenoside F1 to Achieve Its Synthesis in Tobacco

Qin Chen

Kunming University of Science and Technology

Jun Lei

Kunming University of Science and Technology

Xiaolei Li

Kunming University of Science and Technology

Jinyu Zhang

Yunnan Academy of Agricultural Sciences

Diqiu Liu

Kunming University of Science and Technology

Xiuming Cui

Kunming University of Science and Technology

Feng Ge (✉ panaxcordyceps@163.com)

Kunming University of Science and Technology <https://orcid.org/0000-0003-1127-7120>

Research Article

Keywords: Panax notoginseng, Synthetic biology, Metabolic engineering, Saponin

Posted Date: October 23rd, 2023

DOI: <https://doi.org/10.21203/rs.3.rs-3455673/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Ginsenoside F1 has high medicinal values, which is a kind of rare triterpene saponin isolated from *Panax* plants. The extremely low content of ginsenoside F1 in herbs has limited its research and application in medical field. In this work, we constructed a pathway in tobacco for the biosynthesis of ginsenoside F1 by metabolic engineering. Four enzyme genes (*PnDDS*, *CYP12H*, *CYP6H* and *UGT20*) isolated from *Panax notoginseng* were introduced into tobacco. Thus, a biosynthetic pathway for ginsenoside F1 synthesis was artificially constructed in tobacco cells; moreover, the four exogenous genes could be expressed in the roots, stems and leaves of transgenic plants. Consequently, ginsenoside F1 and its precursors were successfully synthesized in the transgenic tobacco, compared with *Panax* plants, the content of ginsenoside F1 in transgenic tobacco was doubled. Based on these results, a new approach was established to produce rare ginsenoside F1; meanwhile, such strategy could also be employed in plant hosts for the heterologous synthesis of other important or rare natural products.

INTRODUCTION

The plants of *Panax* genus including *Panax ginseng*, *Panax quinquefolius*, and *Panax notoginseng* are famous for their wide usage as dietary supplements, functional food and herbs for hundreds of years. Ginsenosides are considered as the crucial bioactive ingredients in *Panax* plants (Zu et al. 2018). The gap between supply and demand of ginsenosides is expanding because the herbs of *Panax* genus are perennial with long growth cycle (about 3–6 years) and low land utilization (Kim, 2018). In addition, the favorite growth condition of *Panax* herbs is that of shade and humidity. Such environments easily invite pests and diseases. Therefore, many pesticides, fungicides and bactericides must be used to prevent and control pests and diseases during the cultivation process, thus resulting in serious soil pollution and a potential risk to human health. The conventional cultivation of *Panax* herbs is considered as an unsustainable and unfriendly mode to environment (Yang et al. 2018). Therefore, new approach for ginsenosides production should be explored to solve the mentioned problems.

Dammarane-type saponins are the main ginsenosides in *Panax* species, which can be divided into protopanaxadiol (PPD) and protopanaxatriol (PPT) groups based on their aglycone structures, among which ginsenoside F1 is a kind of costly and rare PPT-type ginsenoside (Cao et al. 2020). Ginsenoside F1 exhibits anticancer, anti-aging and antioxidant activities (Yoo et al. 2011; Kim et al. 2015; Zhang et al. 2019; Kim et al. 2021; Wang et al. 2016). Unfortunately, the content of ginsenoside F1 is extremely low in *Panax* herbs and the price of ginsenoside F1 goes high, thus restricting the large-scale application of ginsenoside F1 in medicinal and healthcare industries. Recently, synthetic biology is considered as a sustainable and green solution to get important natural medicines by establishing heterologous synthesis system (Kai et al. 2018; Bi et al. 2022; Jiang et al. 2022). We consider that ginsenoside F1 could also be produced through heterologous expression system with green mode rather than extracted it from the plants of *Panax* species.

Using plants and microbes as hosts to realize heterologous production of rare natural products has been considered as a promising and valuable way to solve the deficiency of some natural medicine (Wei et al. 2015; Pouvreau et al. 2018; Pyne et al. 2019; Sun et al. 2019; Shang and Huang, 2020; Li et al. 2021). Microbes have fast reproduction rate, short growth cycle and mature fermentation process. Previous studies have shown that ginsenoside F1 could be synthesized in yeast through metabolic engineering (Wei et al. 2015; Wang et al. 2019). In some cases, microorganisms are not easy domesticating hosts for plant CYP450s and UGTs, thus making the instability or low activity of the heterologous enzymes in microorganisms (Seki et al. 2015). Furthermore, the CYP450s and UGTs derived from plants may be unstable and easy to be dismissed from the microbial hosts during long term of culture (Yu et al. 2017). In addition, plants can use atmospheric CO₂ as carbon source rather than consuming other exogenous carbon source, so plants may be another efficient choice for heterologous synthesis of plant secondary metabolites, especially in the production of plant-derived natural product (Chownk et al. 2018; Li et al. 2019b).

The biosynthetic pathway of ginsenoside F1 has been clarified (Hou et al. 2021). As shown in Fig. 1, ginsenoside F1 is biosynthesized through MVA pathway. The key steps include the following catalytic reactions: dammarenediol-II synthase (PnDDS) converts 2,3-oxidosqualene to dammarenediol- (DD) (Tansakul et al. 2006); DD is subsequently converted to protopanaxdiol (PPD) by hydroxylation of DD at C-12 (Han et al. 2011); and then PPD is hydroxylated to PPT by protopanaxadiol 6-hydroxylase; finally (Han et al. 2012), ginsenoside F1 is produced by glycosylation of C-20-OH of PPT by UDP-glycosyltransferase (UGT) (Yan et al. 2014).

In this study, the pathway of ginsenoside F1 biosynthesis were constructed in tobacco. Four enzyme genes (*PnDDS*, *CYP12H*, *CYP6H* and *UGT20*) derived from *P. notoginseng* were introduced into tobaccos (Fig. 1). The first enzyme PnDDS catalyzed the synthesis of dammarenediol- from 2,3-oxidosqualene; the second enzyme CYP12H hydroxylated dammarenediol- to produce protopanaxdiol; the third enzyme CYP6H hydroxylated protopanaxdiol to protopanaxatriol; the fourth enzyme UGT20 glycosylated C-20 position of protopanaxatriol to synthesize ginsenoside F1. Thus, ginsenoside F1 could be synthesized in tobacco, which would be a sustainable and green approach to produce ginsenoside F1. This study also presented a demonstration to synthesize natural products in heterologous plants.

MATERIALS AND METHODS

Construction of Expression Vectors and Genetic Transformation of Tobacco

The full length of *PnDDS* (GenBank accession number MZ288740), *CYP12H* (GenBank accession number MZ288741), *CYP6H* (GenBank accession number MZ27754) and *UGT20* (GenBank accession number MZ27755) used in this study were isolated from *P. notoginseng* by PCR. The primers of *PnDDS*, *CYP6H*, *CYP12H* and *UGT20* for cloning the open reading frames are as follows. The forward primer for

amplification of the *PnDDS* gene is 5'-ATG TGG AAG CTG AAG GTT GCT GCT CAA GGA-3' and the reverse primer is 5'-TTA AAT TTT GAG CTG CTG-3'; the forward primer for *CYP12H* gene is 5'-ATG GTG TTG TTT TTC TCC CTA TCT-3' and the reverse primer is 5'- TTA ATT GTG GGG ATG TAG ATG AAT-3'; the forward primer used for *CYP6H* is 5'- ATG GAT CTC TTT ATC TCA TCT CAA-3' and the reverse primer is 5'- TTA AAG CGT ACA AGG TGA TAGACG-3'; the forward primer of *UGT20* is 5'-ATG AAG TCA GAA TTG ATA TTC-3' and the reverse primer is 5'-TTA CAT AAT TTC CTC AAA TAG CTT CGC-3'. The open reading frame sequences of these genes were cloned into the pMD18-T vector and then ligated into the pCAMBIA2300s vector, where *PnDDS*, *CYP12H*, *CYP6H* and *UGT20* could be expressed under the 35S cauliflower mosaic virus (CaMV) promoter.

The vectors were transferred into *Agrobacterium tumefaciens* cells and transformed the *Nicotiana benthamiana* leaf discs. After the infection, the leaf discs were transferred to MS medium containing 25 mg/L acetosyringone and cultured in the dark for two days, and then transferred to MS medium containing 50 mg/L kanamycin and 100 mg/L cefotaxime sodium until the buds were differentiated. The buds were transferred to MS medium containing 50 mg/L kanamycin and 100 mg/L cefotaxime sodium.

Screening Positive Transgenic Tobaccos

Positive transgenic tobaccos were screened out on DNA and RNA levels. DNA were extracted from tobaccos by the CTAB method. Total RNA of tobaccos was extracted by Trizol reagent (Invitrogen) and cDNA was obtained by the Reverse Transcription System (Promega). The forward primer for amplification of *PnDDS* gene is 5'-ATG TGG AAG CTG AAG GTT GCT GCT CAA GGA-3' and the reverse primer is 5'-TTA AAT TTT GAG CTG CTG-3'; the forward primer for *CYP12H* gene is 5'-ATG GTG TTG TTT TTC TCC CTA TCT-3' and the reverse primer is 5'- TTA ATT GTG GGG ATG TAG ATG AAT-3'; the forward primer used for *CYP6H* is 5'- ATG GAT CTC TTT ATC TCA TCT CAA-3' and the reverse primer is 5'- TTA AAG CGT ACA AGG TGA TAGACG-3'; the forward primer of *UGT20* is 5'-ATG AAG TCA GAA TTG ATA TTC-3' and the reverse primer is 5'-TTA CAT AAT TTC CTC AAA TAG CTT CGC-3'. PCR was conducted under the conditions: 5.0 min at 94 °C; 30 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 90 s; and a final extension of 7.0 min at 72 °C.

Expression Analysis of Genes

Total RNA was extracted from different tissues (stems, leaves and roots) of transgenic tobacco by Trizol reagent (Invitrogen). The Reverse Transcription System (Promega) was used to synthesize the first strand cDNA. SYBR premix Taq II system was used for real-time quantitative PCR analysis. For RT-qPCR of *PnDDS* gene, the forward primer 5'-AAC ACA AGA AGG AAC TGA-3' and the reverse primer 5'- CTG GAT GAA AAG GAA AAC-3' were used. The forward primer used for *CYP12H* gene of RT-qPCR is 5'- ATT GGG AGG ATT TGA GGA – 3' and the reverse primer is 5'- AGC ATA GGT GAA GTC GGT – 3'. The forward primer used for *CYP6H* gene of RT-qPCR is 5'- AAT ACT CCC CTG ATG TCT – 3' and the reverse primer is 5'- GCT TTG CTT TTC TCT ACC-3'. The forward primer of *UGT20* for RT-qPCR is 5'- GTC CGT CTA TCC CTA AAG – 3' and the reverse primer is 5'- CCA CAC CAC ACA CTC TCT-3'. The forward primer of β -actin gene for RT-qPCR is 5'- CCA GAG CAA GAG AGG TAT – 3' and the reverse primer is 5'- AGG TCA TTG TAG AAG GTG –

3'. The RT-qPCR of GoTaq®2-Step RT-qPCR System was as follows: 95°C for 2.0 min; next 40 cycles of 95°C for 15 s, 57°C for 30 s, and 72°C for 30 s. The $2^{-\Delta\Delta CT}$ method was used to evaluate the relative expression levels of target gene. The *β-actin* gene was used as endogenous control, at least three independent experiments were conducted for each analysis.

Liquid Chromatography Mass Spectrometry Analysis of Ginsenosides

Milled powder (200 mg) from different tissues (stems, leaves and roots) of transgenic tobacco plants was soaked in 100% methanol and was sonicated at a frequency of 20 kHz. The supernatant was collected by centrifugation. After evaporation, the residue was dissolved in 100% methanol and then filtered with a Sep-Pak C-18 Cartridge (Waters, USA).

Liquid chromatography analyses were conducted on a HPLC system (Shimadzu, Kyoto, Japan) with an SPD-M20A photodiode array detector (Acquity UPLC Xevo TQD). The mobile phase (delivered at 0.5 mL/min) consisted of solvent A (H₂O) and solvent B (CH₃CN). A binary gradient elution was performed: 25–35% B for 5 min, 35–55% B for 10 min, 55–70% B for 15 min, 70–90% B for 20 min, 90–100% B for 30 min, 100 – 75% B for 35 min, 75 – 25% for 40 min, and 25% B for 50 min. Chromatographic separation was achieved on a YMC-Pack Pro C-18 RS column (150 ×2.0 mm; D, S-5 mm, 8 nm; YMC Co. Ltd, Japan) at 40 °C.

A liquid chromatography mass spectrometry ion trap/time of flight (LCMS-IT-TOF) mass spectrometer (Shimadzu, Kyoto, Japan) was equipped with an atmospheric pressure chemical ionization source in positive and negative ion mode. The optimized analytical methods were as follows: detector voltage, 1.60 kV; nebulizing gas (N₂) flow, 1.5 L/min; dry gas (N₂) flow, 50 kPa; pressure of TOF region, 1.5×10⁴ Pa; ion trap pressure, 1.7 × 10² Pa; ion accumulation time, 30 ms; precursor ion selected width, 3.0 amu. For qualitative analysis, scan ranges were set at m/z 100–1000 for MS1 and 100–500 for MS2; ultra-high-purity Ar was used as the cooling gas and the collision gas for Collision-induced dissociation (CID) experiments, and the collision energy was set at 50% for MS2. Authentic DD, PPD, PPT and ginsenoside F1 were handled under the same conditions.

Statistical analysis

Data are expressed as means ± SEM. Statistical comparisons were analyzed by one-way analysis of variance (ANOVA) followed by the Tukey-Kramer test to identify significant differences between groups.

RESULTS

Construction of Transgenic Tobaccos for Dammarenediol- β Synthesis

Tobacco can inherently synthesize 2,3-oxidosqualene which is the direct precursor of dammarenediol- biosynthesis. But there is no enzyme in tobacco that catalyzes the synthesis of dammarenediol- from 2,3-oxidosqualene. In this study, the gene *PnDDS* (MZ288740) from *P. notoginseng* was introduced into tobacco plants in order to get the transgenic tobacco which could synthesize dammarenediol- from 2,3-oxidosqualene, the similarity between *PnDDS* and *DDS* (GenBank accession number KJ939266.1) of *Panax ginseng* is 99.05%. PCR result confirmed that *PnDDS* was integrated into the genome of transgenic tobaccos (Fig. S1). Five transgenic tobacco plants were selected for *PnDDS* expression analysis, which showed that the expression level of *PnDDS* was the highest in the stems, followed by the leaves, and the lowest in the roots (Fig. 2A). Dammarenediol- was successfully detected in the transgenic tobaccos by LC/MS, which could not be synthesized in the wild type tobacco (Fig. S2). Such results indicated that *PnDDS* from *P. notoginseng* could be expressed in the transgenic tobaccos, and dammarenediol- synthetase encoded by *PnDDS* catalyzed the synthesis of dammarenediol- from 2,3-oxidosqualene in tobacco. Further analysis showed that the highest contents of dammarenediol- in roots, stems and leaves were about 4.08 µg/g DW, 5.10 µg/g DW and 4.85 µg/g DW in T1-2 plants, respectively (Fig. 2B).

Construction of Transgenic Tobaccos for Protopanaxadiol Synthesis

After the transgenic tobaccos that could synthesize dammarenediol- were achieved, we further constructed an artificial step to produce protopanaxadiol in tobaccos. An enzyme gene *CYP12H* (MZ288741) belonging to cytochrome P450 family was screened from *P. notoginseng*. Bioinformatic analysis of *CYP12H* showed that *CYP12H* had high homology with *CYP716A47*, the nucleic acid similarity between the two genes is 97.54%; and the enzyme encoded by *CYP716A47* could convert dammarenediol- to protopanaxadiol (Chun et al. 2015). So, we considered that *CYP12H* might have the same function as *CYP716A47*. In order to produce protopanaxadiol in tobaccos, *CYP12H* was introduced into the *PnDDS*-transgenic tobaccos that could synthesize dammarenediol-. PCR analysis showed that *CYP12H* was integrated into the genome of transgenic tobaccos (Fig. S3). Five transgenic tobacco plants were selected for expression analysis of *PnDDS* and *CYP12H*. *PnDDS* showed the highest expression level in the stems, followed by the leaves, and the lowest in the roots; and the expression level of *CYP12H* was the highest in the leaves, followed by the stems, and the lowest in the roots (Fig. 3).

Dammarenediol- and protopanaxadiol were detected by LC/MS in the transgenic tobaccos (Fig. S4), indicating that dammarenediol- and protopanaxadiol were successfully synthesized in the transgenic tobaccos. In addition, the content of dammarenediol- in the transgenic tobaccos showed organ specificity, which was mainly accumulated in the aerial parts (stems and leaves). The protopanaxadiol was accumulated the highest in the leaves, followed by the stems, and the lowest in the roots (Fig. 4).

Construction of Transgenic Tobaccos for Protopanaxatriol Synthesis

We have artificially constructed two steps in the transgenic tobaccos in order to synthesize protopanaxadiol. In this section, the third step was introduced into the transgenic tobaccos co-expressed *PnDDS* and *CYP12H* for protopanaxatriol synthesis. An enzyme gene *CYP6H* (MZ27754) belonging to cytochrome P450 family was screened out from *P. notoginseng*. Bioinformatic analysis of *CYP6H* showed that *CYP6H* had high homology with *CYP716A53V2*, the nucleic acid similarity between the two genes is 94.88%, and the enzyme encoded by *CYP716A53V2* could hydroxylated protopanaxadiol to protopanaxatriol (Gwak et al. 2019). So, *CYP6H* might has the similar function as *CYP716A53V2*. In order to produce protopanaxatriol in tobaccos, *CYP6H* was introduced into the transgenic tobaccos producing protopanaxadiol. PCR analysis confirmed that *CYP6H* was integrated into the genome of transgenic tobaccos (Fig. S5). Five transgenic tobacco plants were selected for expression analysis of *PnDDS*, *CYP12H* and *CYP6H*. *PnDDS* showed the highest expression level in the stems; the expression levels of *CYP12H* and *CYP6H* were both the highest in the leaves, followed by the stems and the roots (Fig. 5).

Dammarenediol- , protopanaxadiol and protopanaxatriol were detected by LC/MS in the transgenic tobaccos (Fig. S6), indicating that dammarenediol- , protopanaxadiol and protopanaxatriol were successfully produced in the transgenic tobaccos. Furthermore, the contents of such three compounds in the transgenic tobaccos showed organ specificity, the dammarenediol- was accumulated the highest in the stems (2.45 µg/g DW), followed by the leaves (1.85 µg/g DW), and the lowest in the roots (1.54 µg/g DW); the protopanaxadiol and protopanaxatriol were both accumulated the highest in the leaves, followed by the stems and the roots (Fig. 6).

Construction of Transgenic Tobacco for the Synthesis of Ginsenoside F1

Tobacco can inherently synthesize 2,3-oxidosqualene which is the indirect precursor of ginsenoside F1 biosynthesis. In order to synthesize ginsenoside F1 in tobacco, a biosynthetic pathway consisting of four catalytic steps could be constructed, which started from 2,3-oxidosqualene (Fig. 1). In this pathway, *PnDDS* cyclizes 2,3-oxidosqualene to synthesize dammarenediol- ; *CYP12H* hydroxylates dammarenediol- to produce protopanaxadiol; then *CYP6H* hydroxylates protopanaxadiol to produce protopanaxatriol; finally, glycosylates C-20 position of protopanaxatriol to synthesize ginsenoside F1.

In this study, a gene named *UGT20* (MZ27754) from *P. notoginseng* was further introduced into the transgenic tobacco plants co-expressing *PnDDS*, *CYP12H* and *CYP6H* in order to synthesize ginsenoside F1. The enzyme encoded by *UGT20* could glycosylate C-20 position of protopanaxatriol to synthesize ginsenoside F1 ((Wei et al. 2015), the similarity between *UGT20* and *UGTPg1* (GenBank accession number KF377585.1) of *Panax ginseng* is 93.53%. Based on the protopanaxatriol-producing transgenic tobaccos obtained above, five transgenic tobacco plants containing *PnDDS*, *CYP12H*, *CYP6H* and *UGT20* were screened out for gene expression analysis (Fig. 7). The results showed that *PnDDS* expressed the highest in the stems (Fig. 8A); *CYP12H* and *CYP6H* expressed the highest in the leaves (Fig. 8B and 8C); and the expression level of *UGT20* was the highest in the roots (Fig. 8D). These data indicated that the

glycosylation of protopanaxtriol to synthesize ginsenoside F1 might occur in the root of transgenic tobaccos.

Triterpenoid saponins in the transgenic tobacco were analyzed using LCMS-IT-TOF. There were four new compounds at the retention times of 11.80, 16.50, 25.25 and 27.45 min in the transgenic tobaccos compared with the wild type, which were in accordance with the retention times of ginsenoside F1, protopanaxtriol, protopanaxdiol and dammarenediol-, respectively (Fig. 9). The MS fragmentation pattern with peak retention times of 27.45 min included fragments with m/z ratios of 409 $[M-2H_2O + H]^+$ and 427 $[M-H_2O + H]^+$, which were the same as those for the dammarenediol- standard. The MS fragmentation pattern with peak retention times of 25.25 min included fragments with m/z ratios of 407 $[M-3H_2O + H]^+$, 425 $[M-2H_2O + H]^+$, and 443 $[M-H_2O + H]^+$, which were the same as those for the protopanaxdiol standard. The MS fragmentation pattern with peak retention times of 16.50 min in transgenic tobacco included fragments with m/z ratios of 423 $[M-3H_2O + H]^+$, 441 $[M-2H_2O + H]^+$, and 459 $[M-H_2O + H]^+$, which were the same as those for the protopanaxtriol standard. The MS fragmentation pattern with peak retention times of 11.80 min in transgenic tobacco included fragments with m/z ratios of 585 $[M-3H_2O + H]^+$, 603 $[M-2H_2O + H]^+$, and 621 $[M-H_2O + H]^+$, which were the same as those for the ginsenosides F1 standard (Fig. S7).

The contents of ginsenoside F1 and its precursors were determined in the roots, stems and leaves (Fig. 11). The accumulation of ginsenoside F1 was the highest in the roots (7.30 $\mu\text{g/g DW}$), followed by the leaves (4.85 $\mu\text{g/g DW}$) and the lowest in the stems (3.95 $\mu\text{g/g DW}$), compared with *Panax* plants, the content of ginsenoside F1 in transgenic tobacco was doubled. The accumulation of protopanaxtriol was the highest in the leaves (4.20 $\mu\text{g/g DW}$), followed by the stems (3.30 $\mu\text{g/g DW}$) and the lowest in the roots (3.00 $\mu\text{g/g DW}$). The accumulation of protopanaxdiol was the highest in the leaves (4.18 $\mu\text{g/g DW}$), followed by the stems (3.84 $\mu\text{g/g DW}$) and the lowest in the roots (3.58 $\mu\text{g/g DW}$). Moreover, the content of dammarenediol- was the highest in the stems (2.45 $\mu\text{g/g DW}$), followed by the leaves (1.85 $\mu\text{g/g DW}$) and the lowest in the roots (1.56 $\mu\text{g/g DW}$). According to the above results, we found that the precursor dammarenediol- for protopanaxdiol synthesis was mainly accumulated in stems and leaves, but the contents of protopanaxdiol and protopanaxtriol were higher in leaves compared with those in stems and roots. In addition, compared with wild-type tobacco, transgenic tobacco has thinner stems and leaves (Fig. S8).

DISCUSSION

Dammarenediol- is found only in a few plant species such as *Panax* and *Gynostemma* genera (Wang et al. 2014), which is the basic skeleton of ginsenosides. In *Panax* species, dammarenediol- was synthesized by cyclization of 2,3-oxidosqualene under the catalysis of dammarenediol- synthetase (Lee et al. 2012). Moreover, 2,3-oxidosqualene as the precursor of dammarenediol- exists in most of the eukaryotes (Chun et al. 2015). Therefore, it is possible to achieve heterologous biosynthesis of dammarenediol- by expressing dammarenediol- synthetase in the eukaryotes besides *Panax* species.

In this study, *PnDDS* was successfully expressed in tobaccos and synthesized the ginsenoside skeleton, dammarenediol-, using 2,3-oxidosqualene, which indicated that the tobacco might have the feasibility and potential of heterologous synthesis of ginsenosides. Although the expression of *PnDDS* in tobacco plants was driven by constitutive CaMV35 promoter, the expression level of *PnDDS* and the accumulation of dammarenediol- also showed organ specificity. The highest content of dammarenediol- was presented in the stem (Fig. 2). Dammarenediol-, as the skeleton of ginsenosides biosynthesis, cannot synthesize ginsenosides directly, which should be firstly converted to protopanaxadiol or protopanaxatriol, and then synthesized diverse ginsenosides. Therefore, after dammarenediol- produced, it is necessary to further construct pathway that can synthesize protopanaxadiol or protopanaxatriol in order to obtain ginsenosides in tobaccos.

Protopanaxadiol and protopanaxatriol are considered as the direct precursor of ginsenosides, which are generated by the hydroxylation of dammarenediol- (Kai et al. 2018; Zhan et al. 2018). In *Panax* plants, the amount of protopanaxadiol and protopanaxatriol are both too low and almost undetectable (Peng et al. 2018). This may be because the glycosyltransferases modifying protopanaxadiol or protopanaxatriol to synthesize ginsenosides have high activities and rapidly transform protopanaxadiol or protopanaxatriol to ginsenoside products, resulting in the difficulty of protopanaxadiol and protopanaxatriol accumulation in plants (Park et al. 2016). In this study, the wild tobacco has no enzyme to catalyze the synthesis of ginsenosides from protopanaxadiol or protopanaxatriol, so protopanaxadiol and protopanaxatriol could be accumulated and easily detected in the transgenic tobaccos.

Furthermore, the content of dammarenediol- was low in the root along with the low expression level of *PnDDS* compared with those in the leaf and stem. In addition, the contents of protopanaxadiol and protopanaxatriol in stems and leaves were significantly higher than those in roots. It is speculated that after dammarenediol- is produced in the stems and leaves, CYP12H and CYP6H continuously hydroxylate dammarenediol- to synthesize protopanaxatriol (Zeng et al. 2018). Then, part of the protopanaxatriol may be transported from leaves and stems to roots, and used to synthesize ginsenoside F1. Previous study has shown that the precursors for ginsenosides biosynthesis were mainly produced in the aerial parts of plants in *Panax* genus, and then such precursors were transported to the roots and glycosylated to synthesize ginsenosides (Rai et al. 2016; Li et al. 2019a). Our study indicated that the biosynthesis of ginsenosides in the transgenic tobaccos were similar to that in *Panax* plants. Such characters of ginsenoside biosynthesis should be common in plants, but more evidences need to support this (Liu et al. 2015; Tang et al. 2022).

Ginsenosides are composed of sapogenins linked to one or more glycosyls, thus UGTs play important roles in this process (Zhao et al. 2020; Li et al. 2022). Glycosylation changes the physicochemical properties and biological activities of triterpenoid saponins in plants (Lu et al. 2017). Although ginsenoside F1 has important biological activities, the content of which in *Panax* herb plants is extremely low. The heterologous biosynthesis of rare ginsenosides by synthetic biology and metabolic engineering is considered as a sustainable, green and effective approach (Ahmed et al. 2021). CYP450s and UGTs for ginsenoside biosynthesis are more suitable for expression in plants, so the heterologous biosynthesis of

ginsenoside F1 using plants as hosts may be a suitable choice. Tobacco is not only an industrial crop cultivated around the world, but also a model plant commonly used in plant research, which has a clear genetic background and is easy to carry out genetic transformation and regeneration (Xu et al. 2019). Our study showed tobacco could be a potential plant host for heterologous biosynthesis of ginsenoside F1.

Plants and microbes are usually used as hosts to realize heterologous synthesis of important natural products, which both have merits and demerits (Khakimov et al. 2015; Tatsis and O'Connor. 2016). As the host, microbes have the advantages of easy cultivation, fast growth and brief genetic operation, but some enzyme components obtained from plants for heterologous synthesis of compounds may present low expression level, low activity and easy loss in microbial cells, which is manifested as the incompatibility of biological components (Zhao and Li. 2018). For example, the construction of microbial cell factory for ginsenoside Rg3 synthesis led to poor compatibility between the key UGT components and microbial hosts, resulting in low expression levels of exogenous genes, low glycosylation efficiency and the limited yield of the target compound (Guo et al. 2020; Sokolov et al. 2022). Although ginsenoside F1 has been successfully synthesized in microbes (Wei et al. 2015; Wang et al. 2019), plant cells could be another choice as the host for heterologous synthesis of ginsenoside F1. So far, the production of ginsenoside F1 in plant hosts is too low to be used for application in industry. Compared with the long planting time and low biomass of *P. notoginseng*, during the same planting time, the content of ginsenoside F1 in transgenic tobacco was about 15 times that in *P. notoginseng*. In the next, several strategies for increasing the biosynthesis of ginsenoside F1 in tobaccos will be explored in order to realize a comparable yield to that in microbial hosts.

CONCLUSIONS

In this study, we reported a new approach to achieve heterologous biosynthesis of ginsenoside F1 in tobacco by metabolic engineering. Four enzyme genes (*PnDDS*, *CYP12H*, *CYP6H* and *UGT20*) isolated from *P. notoginseng* were introduced into tobacco, thus a biosynthetic pathway of ginsenoside F1 was constructed in tobacco. Ginsenoside F1 and its precursors, protopanaxatriol, protopanaxadiol and dammarenediol-II were successfully synthesized in the transgenic tobaccos, compared with *Panax* plants, the content of ginsenoside F1 in transgenic tobacco was doubled. This study established a novel way to produce ginsenoside F1 through an environmentally friendly mode, meanwhile, the strategy of this research could also be employed for the heterologous synthesis of other natural products in plant hosts.

Declarations

Author contributions

FG and XC designed the experiments; QC and JL performed the experiments; QC, DL, JZ and XL analyzed the data; QC wrote the paper; and all authors reviewed the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (82060692), the Yunnan Fundamental Research Projects, China (202101AS070024), and Yunnan Key Research and Development Program, China (2019ZF011-2). Feng Ge was supported by Yunnan Ten Thousand Talents Plan (Young & Elite Talents Project).

References

1. Ahmed MS, Lauersen KJ, Ikram S, Li C (2021) Efflux transporters' engineering and their application in microbial production of heterologous metabolites. *ACS Synth Biol* 10: 646–669
2. Bi H, Qu G, Wang S, Zhuang Y, Sun Z, Liu T, Ma Y (2022) Biosynthesis of a rosavin natural product in *Escherichia coli* by glycosyltransferase rational design and artificial pathway construction. *Metab Eng* 69: 15–25
3. Cao L, Wu H, Zhang H, Zhao Q, Yin X, Zheng D, Li C, Kim MJ, Kim P, Xue, Z., Wang Y, Li Y (2020) Highly efficient production of diverse rare ginsenosides using combinatorial biotechnology. *Biotechnol Bioeng* 117: 1615–1627
4. Chownk M, Thakur K, Yadav SK (2018) Retrospect and prospects of plant metabolic engineering. *J Plant Biochem Biotechnol* 28: 1–13
5. Chun JH, Adhikari PB, Park SB, Han JY, Choi YE (2015) Production of the dammarene sapogenin (protopanaxadiol) in transgenic tobacco plants and cultured cells by heterologous expression of *PgDDS* and *CYP716A47*. *Plant Cell Rep* 34: 1551–1560
6. Guo H, Wang H, Huo YX (2020) Engineering critical enzymes and pathways for improved triterpenoid biosynthesis in yeast. *ACS Synth Biol* 9: 2214–2227
7. Gwak YS, Han JY, Choi YE (2019) Production of ginsenoside aglycone (protopanaxatriol) and male sterility of transgenic tobacco co-overexpressing three *Panax ginseng* genes: *PgDDS*, *CYP716A47*, and *CYP716A53v2*. *J Ginseng Res* 43: 261–271
8. Han JY, Hwang HS, Choi SW, Kim HJ, Choi, YE (2012) Cytochrome P450 CYP716A53v2 catalyzes the formation of protopanaxatriol from protopanaxadiol during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol* 53: 1535–1545
9. Han JY, Kim HJ, Kwon YS, Choi YE (2011) The Cyt P450 enzyme CYP716A47 catalyzes the formation of protopanaxadiol from dammarenediol-II during ginsenoside biosynthesis in *Panax ginseng*. *Plant Cell Physiol*. 52: 2062–2073
10. Hou M, Wang R, Zhao S, Wang Z (2021) Ginsenosides in *Panax* genus and their biosynthesis. *Acta Pharm Sin B* 11: 1813–1834
11. Jiang Z, Gao H, Liu R, Xia M, Lu Y, Wang J, Chen X, Zhang Y, Li D, Tong Y, Liu P, Liu Y, Luo Y, Gao J, Yin Y, Huang L, Gao W (2022) Key glycosyltransferase genes of *Panax notoginseng*: identification

- and engineering yeast construction of rare ginsenosides. *ACS Synth Biol* 11: 2394–2404
12. Kai G, Zhou W, Wang Y, Hua Q, Shi M (2018) Research progress in terms of the biosynthesis and regulation of terpenoids from medicinal plants. *Sci Sin Vitae* 48: 352–364
 13. Khakimov B, Kuzina V, Erthmann PO, Fukushima EO, Augustin JM, Olsen CE, Scholtalbers J, Volpin H, Andersen SB, Hauser TP, Muranaka T, Bak S (2015) Identification and genome organization of saponin pathway genes from a wild crucifer, and their use for transient production of saponins in *Nicotiana benthamiana*. *Plant J* 84: 478–490
 14. Kim, JH (2018) Pharmacological and medical applications of *Panax ginseng* and ginsenosides: a review for use in cardiovascular diseases. *J Ginseng Res* 42: 264–269
 15. Kim JH, Baek EJ, Lee EJ, Yeom MH, Kang NJ (2015) Ginsenoside F1 attenuates hyperpigmentation in B16F10 melanoma cells by inducing dendrite retraction and activating Rho signalling. *Exp Dermatol* 24 (2): 150-152.
 16. Kim, SJ, Lee J, Choi WS, Kim HJ, Kim MY, Sun CK, Kim HS. (2021) Ginsenoside F1 attenuates eosinophilic inflammation in chronic rhinosinusitis by promoting NK cell function. *J Ginseng Res* 45 (6): 695-705
 17. Lee MH, Han JY, Kim HJ, Kim YS, Huh GH, Choi YE (2012) Dammarenediol-II production confers TMV tolerance in transgenic tobacco expressing *Panax ginseng* dammarenediol-II synthase. *Plant Cell Physiol* 53: 173–182
 18. Li J, Ma L, Zhang S, Zuo C, Song N, Zhu S, Wu J (2019a) Transcriptome analysis of 1-and 3-year-old *Panax notoginseng* roots and functional characterization of saponin biosynthetic genes *DS* and *CYP716A47*-like. *Planta* 249: 1229–1237
 19. Li J, Mutanda I, Wang K, Yang L, Wang J, Wang Y (2019b) Chloroplastic metabolic engineering coupled with isoprenoid pool enhancement for committed taxanes biosynthesis in *Nicotiana benthamiana*. *Nat Commun* 10: 4850.
 20. Li R, Wang K, Wang D, Xu L, Shi Y, Dai ZB, Zhang XL (2021) Production of plant volatile terpenoids (rose oil) by yeast cell factories. *Green Chem* 23: 5088–5096
 21. Li Y, Li J, Diao M, Peng L, Huang S, Xie N (2022) Characterization of a group of UDP-glycosyltransferases involved in the biosynthesis of triterpenoid saponins of *Panax notoginseng*. *ACS Synth Biol* 11: 770–779
 22. Liu MH, Yang BR, Cheung WF, Yang KY, Zhou HF, Kwok JS, Liu GC, Li XF, Zhong, S, Lee SM, Tsui SK (2015) Transcriptome analysis of leaves, roots and flowers of *Panax notoginseng* identifies genes involved in ginsenoside and alkaloid biosynthesis. *BMC Genomics* 16: 265
 23. Lu C, Zhao S, Wei G, Zhao H, Qu Q (2017) Functional regulation of ginsenoside biosynthesis by RNA interferences of a UDP-glycosyltransferase gene in *Panax ginseng* and *Panax quinquefolius*. *Plant Physiol Biochem* 111: 67–76
 24. Park SB, Chun JH, Ban YW, Han JY, Choi YE (2016) Alteration of *Panax ginseng* saponin composition by overexpression and RNA interference of the protopanaxadiol 6-hydroxylase gene (*CYP716A53v2*). *J Ginseng Res* 40: 47–54

25. Peng M, Yi YX, Zhang T, Ding Y, Le J (2018) Stereoisomers of saponins in *Panax notoginseng* (Sanqi): A review. *Front Pharmacol* 9: 188
26. Pouvreau B, Vanhercke T, Singh S (2018) From plant metabolic engineering to plant synthetic biology: The evolution of the design/build/test/learn cycle. *Plant Sci* 273: 3–12
27. Pyne ME, Narcross L, Martin VJJ (2019) Engineering plant secondary metabolism in microbial systems. *Plant Physiol* 179: 844–861
28. Rai A, Yamazaki M, Takahashi H, Nakamura M, Kojoma M, Suzuki H, Saito K (2016) RNA-seq Transcriptome analysis of *Panax japonicus*, and Its comparison with other *Panax* species to identify potential genes involved in the saponins biosynthesis. *Front Plant Sci* 7: 481
29. Seki H, Tamura K, Muranaka T (2015) P450s and UGTs: key players in the structural diversity of triterpenoid saponins. *Plant Cell Physiol* 56: 1463–1471
30. Shang Y, Huang S (2020) Engineering plant cytochrome P450s for enhanced synthesis of natural products: Past achievements and future perspectives. *Plant Commun* 1: 100012
31. Sokolov SS, Volynsky PE, Zangieva OT, Severin FF, Glagoleva ES, Knorre DA (2022) Cytostatic effects of structurally different ginsenosides on yeast cells with altered sterol biosynthesis and transport. *Biochim Biophys Acta Biomembr* 1864: 183993
32. Sun W, Qin L, Xue H, Yu Y, Ma Y, Wang Y, Li C (2019) Novel trends for producing plant triterpenoids in yeast. *Crit Rev Biotechnol* 39: 618–632
33. Tang Q, Lu Y, Zhang G, Tang J, Yang Z, Lin Y, Li Y, Song W, Li X, Chen J, Fan W, Yang S, Hao B (2022) Ginsenosides accumulation and related genes expression in different underground parts of *Panax notoginseng* during maturation stage. *Ind Crops Prod* 175: 114228
34. Tansakul P, Shibuya M, Kushiro T, Ebizuka YJ (2006) Dammarenediol-II synthase, the first dedicated enzyme for ginsenoside biosynthesis, in *Panax ginseng*. *FEBS Lett* 580(22): 5143–5149
35. Tatsis EC, O'Connor SE (2016) New developments in engineering plant metabolic pathways. *Curr Opin Biotechnol* 42: 126–132
36. Wang JH, Wang D, Li WX, Huang Y, Dai ZB, Zhang XL (2019) Optimization of UDP-glucose supply module and production of ginsenoside F1 in *Saccharomyces cerevisiae*. *China J Chin Mater Med* 44: 4596-4604.
37. Wang L, Zhao SJ, Cao HJ, Sun Y (2014) The isolation and characterization of dammarenediol synthase gene from *Panax quinquefolius* and its heterologous co-expression with cytochrome P450 gene *PqD12H* in yeast. *Funct Integr Genomics* 14: 545–557
38. Wang Y, Choi KD, Yu H, Jin F, Im WT (2016) Production of ginsenoside F1 using commercial enzyme Cellulase KN. *J Ginseng Res* 40 (2): 121-126
39. Wei W, Wang P, Wei Y, Liu Q, Yang C, Zhao G, Yue J, Yan X, Zhou Z (2015) Characterization of *Panax ginseng* UDP-Glycosyltransferases catalyzing protopanaxatriol and biosyntheses of bioactive ginsenosides F1 and Rh1 in metabolically engineered yeasts. *Mol Plant* 8 (9): 1412-1424

40. Xu H, Li W, Schilmiller AL, Eekelen H, Vos RCH, Jongsma MA, Pichersky E (2019) Pyrethric acid of natural pyrethrin insecticide: complete pathway elucidation and reconstitution in *Nicotiana benthamiana*. *New Phytol* 223: 751–765
41. Yan X, Fan Y, Wei W, Wang P, Liu Q, Wei Y, Zhang L, Zhao G, Yue J, Zhou Z (2014) Production of bioactive ginsenoside compound K in metabolically engineered yeast. *Cell Res* 24: 770–773
42. Yang JL, Hu ZF, Zhang TT, Gu AD, Gong T, Zhu P (2018) Progress on the studies of the key enzymes of ginsenoside biosynthesis. *Molecules* 23(3):589
43. Yoo DS, Yu T, Lee YG, Yang W, Hong S, Cho JY (2011) Ginsenoside F1 modulates cellular responses of skin melanoma cells. *J Ginseng Res* 35(1): 86-91
44. Yu T, Zhou YJ, Wenning L, Liu Q, Krivoruchko A, Siewers V, Nielsen J, David F (2017) Metabolic engineering of *Saccharomyces cerevisiae* for production of very long chain fatty acid-derived chemicals. *Nat Commun* 8: 15587
45. Zeng X, Luo T, Li J, Li G, Zhou D, Liu T, Zou X, Pandey A, Luo Z (2018) Transcriptomics-based identification and characterization of 11 CYP450 genes of *Panax ginseng* responsive to MeJA. *Acta Biochim Biophys Sin Shanghai* 50: 1094–1103
46. Zhan Y, Yin J, Xiao J, Yang J (2018) Advances in the function of cytochrome P450 in structural modifications of triterpenoid and sterol skeletons in plants. *Sci Sin Vitae* 48: 1065–1083
47. Zhang J, Liu M, Huang M, Chen M, Zhang DJ (2019) Ginsenoside F1 promotes angiogenesis by activating the IGF-1/IGF1R pathway. *Pharmacol Res* 144: 292-305
48. Zhao JN, Wang RF, Zhao SJ, Wang ZT (2020) Advance in glycosyltransferases, the important bioparts for production of diversified ginsenosides. *China J Nat Med* 18: 643–658
49. Zhao YJ, Li C (2018) Biosynthesis of plant triterpenoid saponins in microbial cell factories. *J Agric Food Chem* 66: 12155–12165
50. Zu Y, Li Z, Mei X, Wu J, Cheng S, Jiang Y, Li Y (2018) Transcriptome analysis of main roots of *Panax notoginseng* identifies genes involved in saponin biosynthesis under arsenic stress. *Plant Gene* 16: 1–7

Figures

Transgenic tobacco

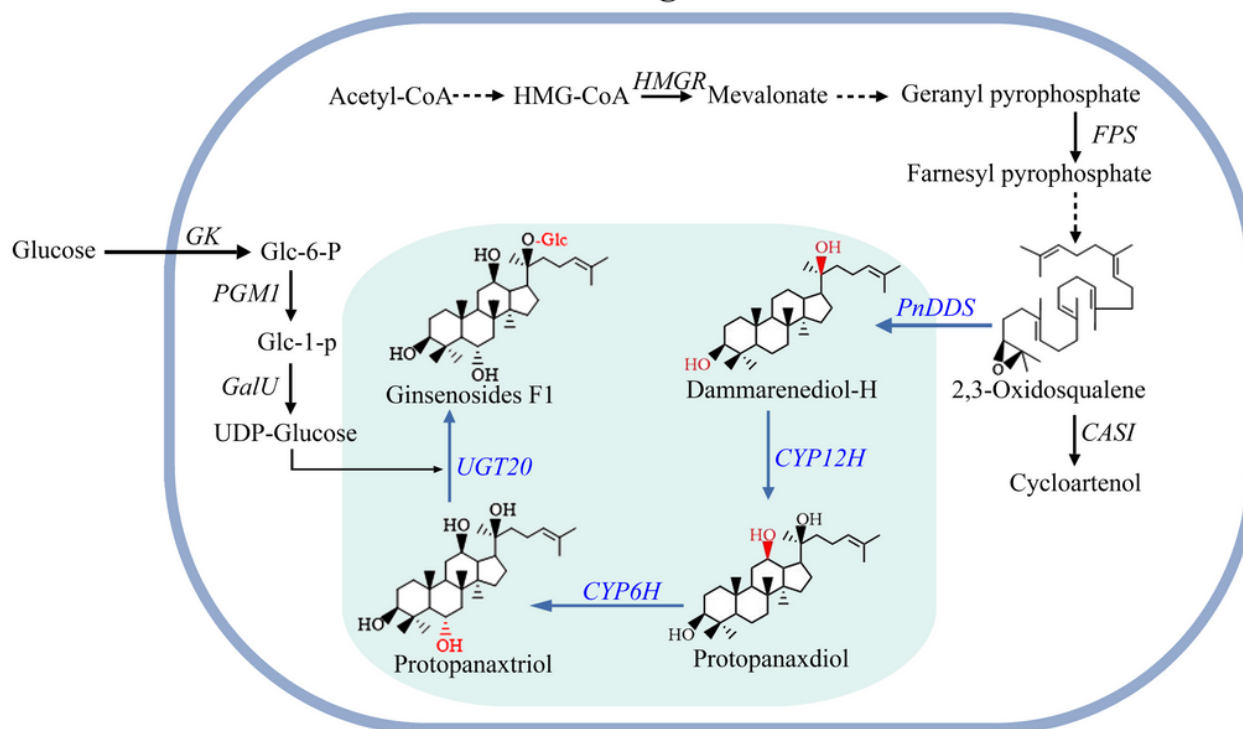


Figure 1

Designed biosynthetic pathway of ginsenoside F1 in tobacco. The black arrows indicate the connatural metabolic pathway in tobacco. The blue arrows represent the pathway constructed in tobacco for ginsenoside F1 biosynthesis. Genes from *P. notoginseng* are shown in blue. Red chemical groups indicate the modification sites. Multiple-step reactions are shown as dashed arrows. Glc-6-P, Glucose 6-phosphate; Glc-1-P, Glucose 1-phosphate.

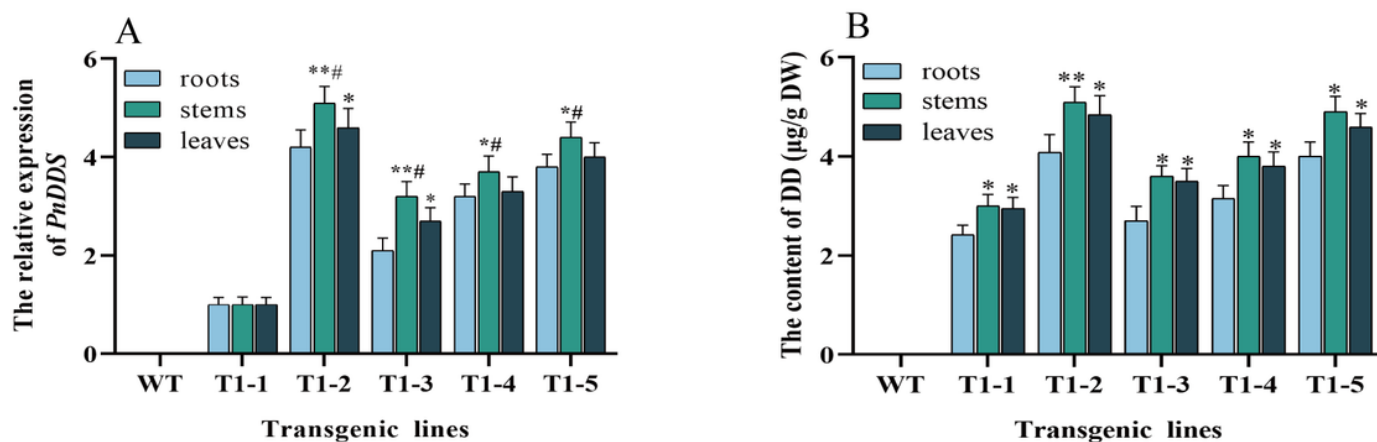


Figure 2

PnDDS expression level and dammarenediol- content in the *PnDDS*transgenic tobaccos. T1, *PnDDS*-transgenic plants; (A) *PnDDS* expression levels in transgenic tobaccos; (B) dammarenediol-II (DD) content in transgenic tobaccos. Compared with roots, * $P<0.05$, ** $P<0.01$. Compared with leaves, # $P<0.05$.

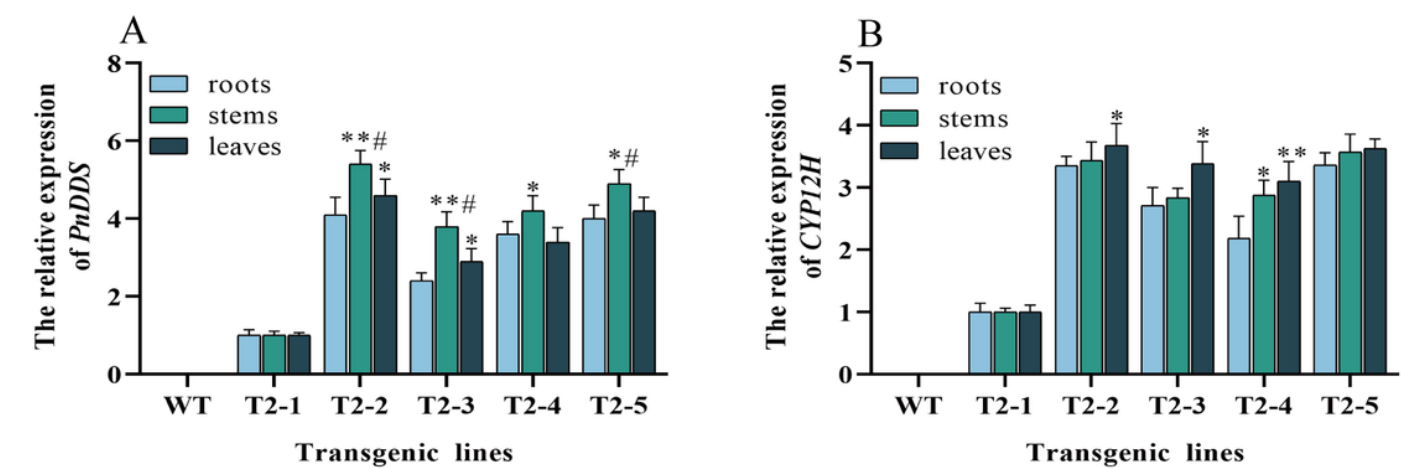


Figure 3

PnDDS and *CYP12H* expression levels in *PnDDS/ CYP12H*-transgenic tobaccos. T2, *PnDDS/ CYP12H*-transgenic plants; (A) *PnDDS* expression level in transgenic tobaccos; (B) *CYP12H* expression level in transgenic tobaccos. Compared with roots, * $P<0.05$, ** $P<0.01$. Compared with leaves, # $P<0.05$.

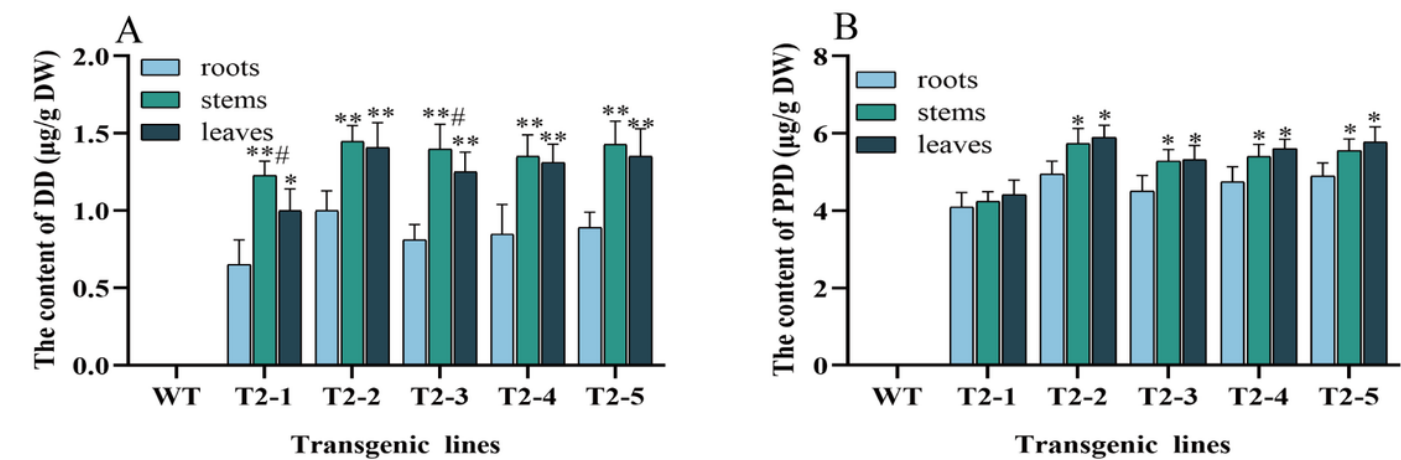


Figure 4

The contents of dammarenediol-II and protopanaxadiol in *PnDDS/ CYP12H*-transgenic tobaccos. T2, *PnDDS/ CYP12H*-transgenic plants; (A) dammarenediol-II content in transgenic tobaccos; (B) protopanaxadiol content in transgenic tobacco; DD, dammarenediol-II; PPD, protopanaxadiol. Compared with roots, * $P<0.05$, ** $P<0.01$. Compared with leaves, # $P<0.05$.

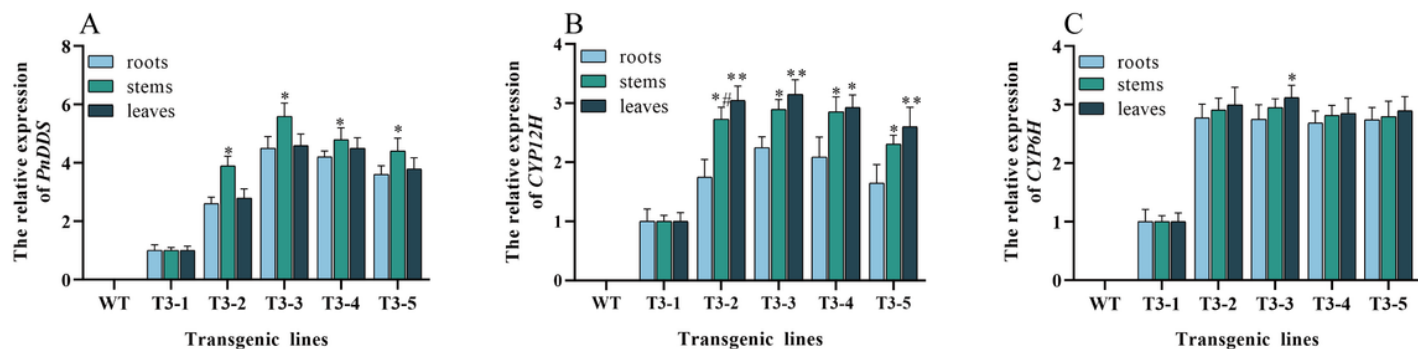


Figure 5

PnDDS, *CYP6H* and *CYP12H* expression levels in the *PnDDS*/*CYP12H*/*CYP6H*-transgenic tobaccos. T3, *PnDDS*/*CYP12H*/*CYP6H*-transgenic plants; (A) *PnDDS* expression level in transgenic tobaccos; (B) *CYP12H* expression level in transgenic tobaccos; (C) *CYP6H* expression level in transgenic tobaccos. Compared with roots, * $P < 0.05$, ** $P < 0.01$. Compared with leaves, # $P < 0.05$.

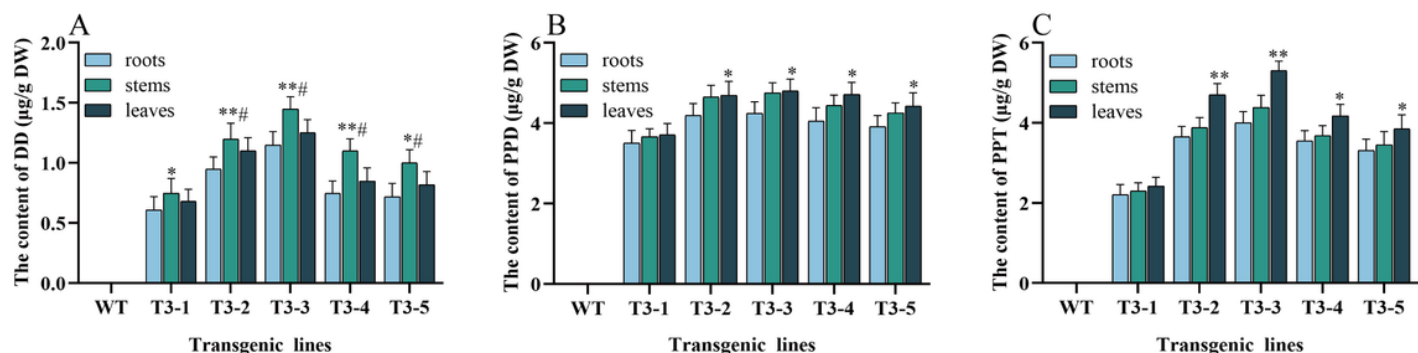


Figure 6

The contents of dammarenediol-II, protopanaxadiol and protopanaxatriol in *PnDDS*/*CYP12H*/*CYP6H*-transgenic tobacco. T3, *PnDDS*/*CYP12H*/*CYP6H*-transgenic plants; (A) dammarenediol-II content in transgenic tobaccos; (B) protopanaxadiol content in transgenic tobaccos; (C) protopanaxatriol content in transgenic tobaccos; DD, dammarenediol-II; PPD, protopanaxadiol; PPT, protopanaxatriol. Compared with roots, * $P < 0.05$, ** $P < 0.01$. Compared with leaves, # $P < 0.05$.

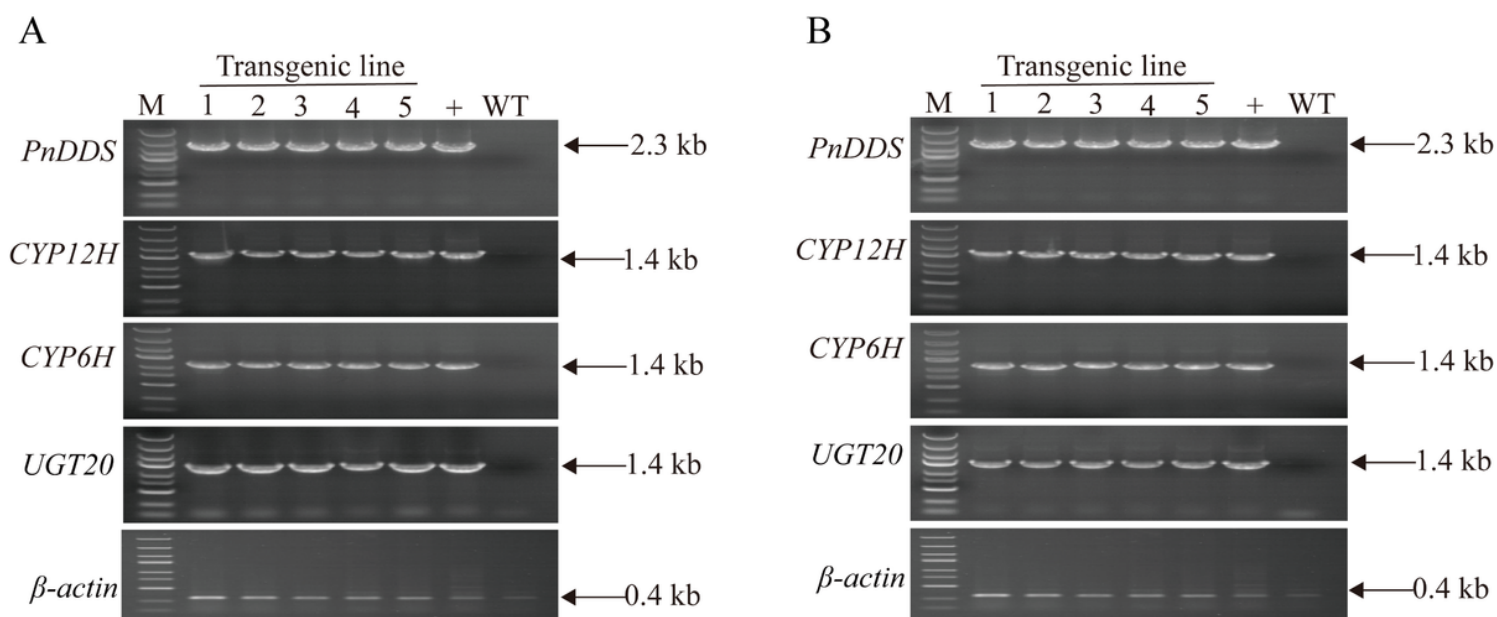


Figure 7

Detection and expression analysis of introduced genes in the transgenic tobacco co-expressing *PnDDS*, *CYP12H*, *CYP6H* and *UGT20*. (A) Genomic DNA PCR used for screening out positive transgenic tobacco lines. (B) RT-PCR analysis of the introduced *PnDDS*, *CYP12H*, *CYP6H* and *UGT20* genes in transgenic lines (T1, T2, T3, T4, and T5). WT, wild-type. +, positive control.

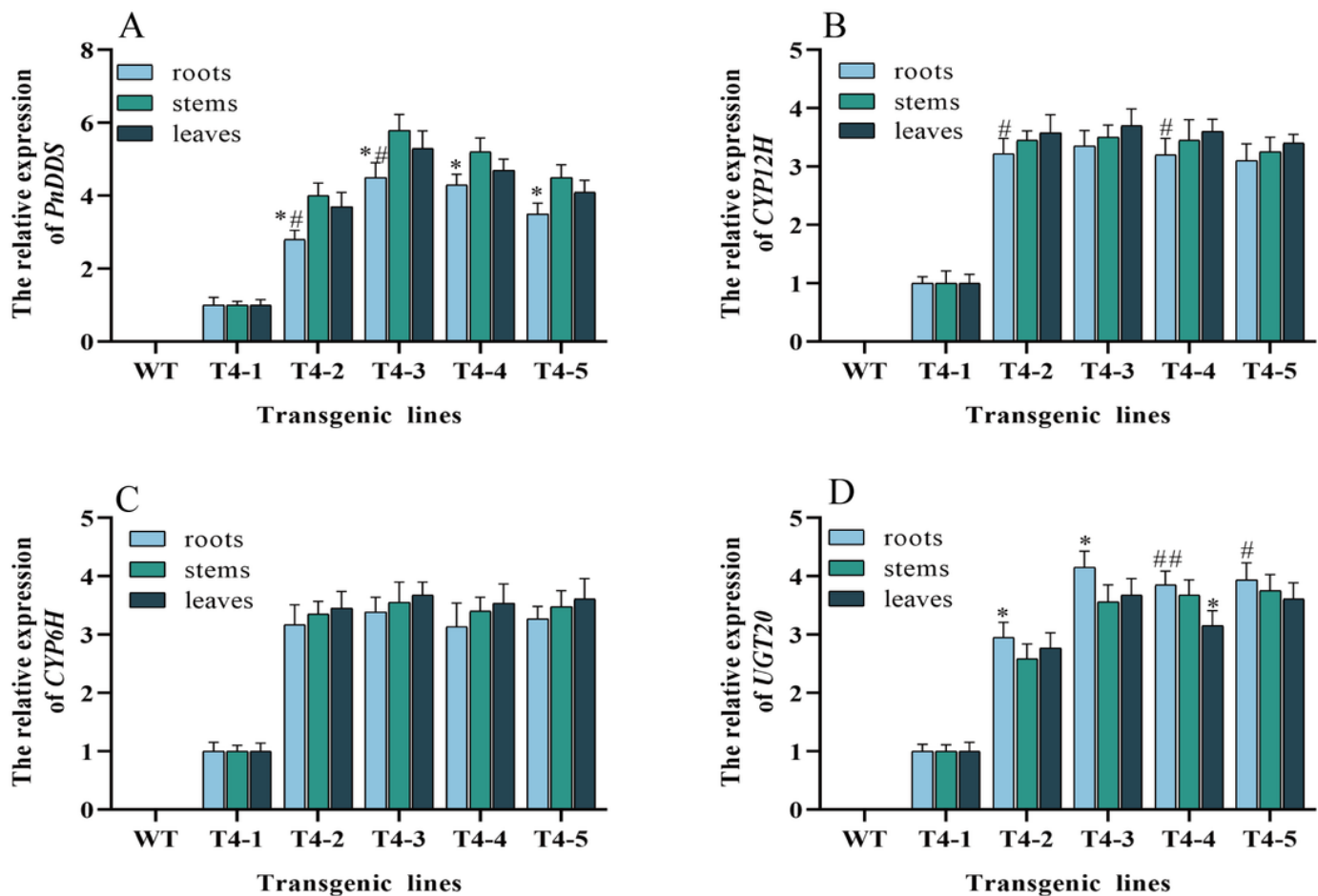


Figure 8

PnDDS, *CYP12H*, *CYP6H* and *UGT20* expression levels in *PnDDS*/*CYP12H*/*CYP6H*/*UGT20*-transgenic tobaccos. T4, *PnDDS*/*CYP12H*/*CYP6H*/*UGT20*-transgenic plants; (A) *PnDDS* expression level in transgenic tobaccos; (B) *CYP12H* expression level in transgenic tobaccos; (C) *CYP6H* expression level in transgenic tobaccos; (D) *UGT20* expression level in transgenic tobaccos. Compared with stems, * $P < 0.05$. Compared with leaves, # $P < 0.05$. ## $P < 0.01$.

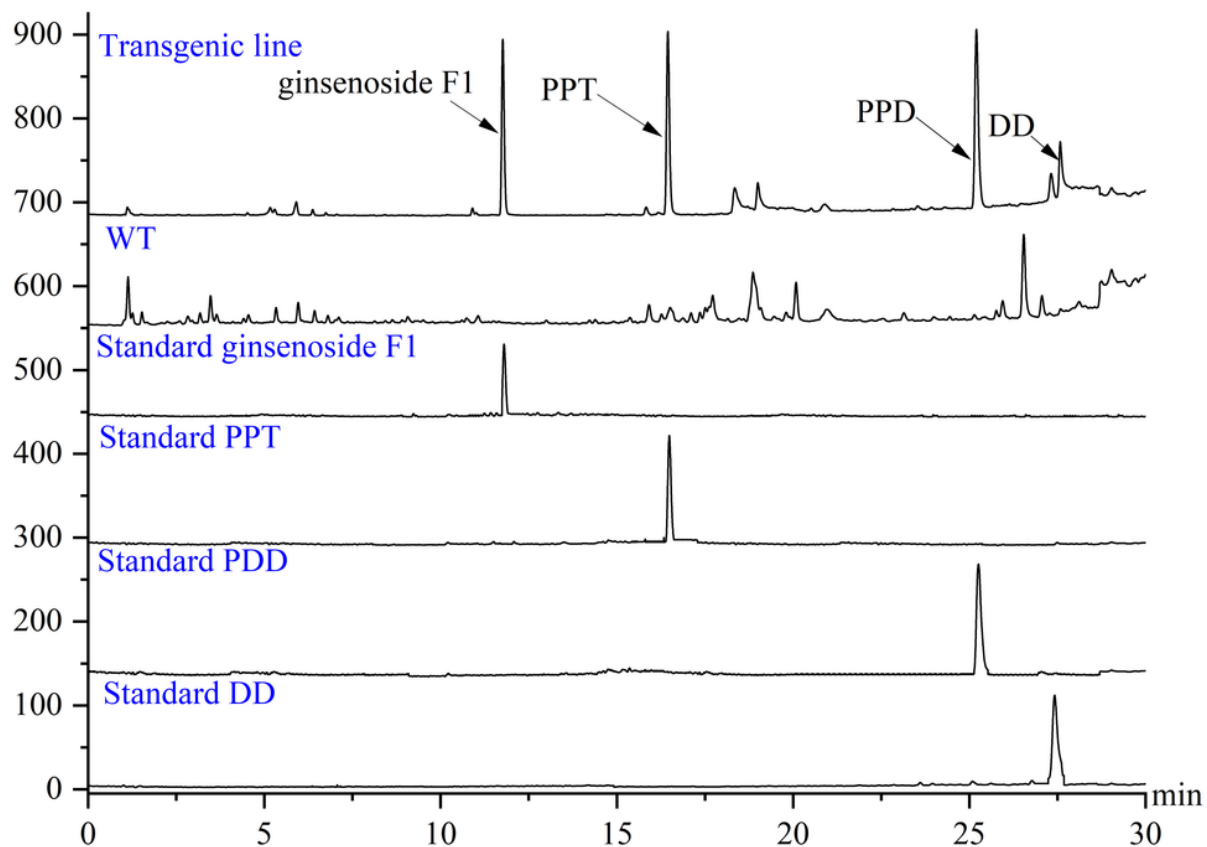


Figure 9

LC-MS analyses of dammarenediol-II, protopanaxadiol, protopanaxatriol and ginsenoside F1 in the transgenic tobacco co-expressing *PnDDS*, *CYP12H*, *CYP6H* and *UGT20*. DD, dammarenediol-II; PPD, protopanaxadiol; PPT, protopanaxatriol; F1, ginsenoside F1.

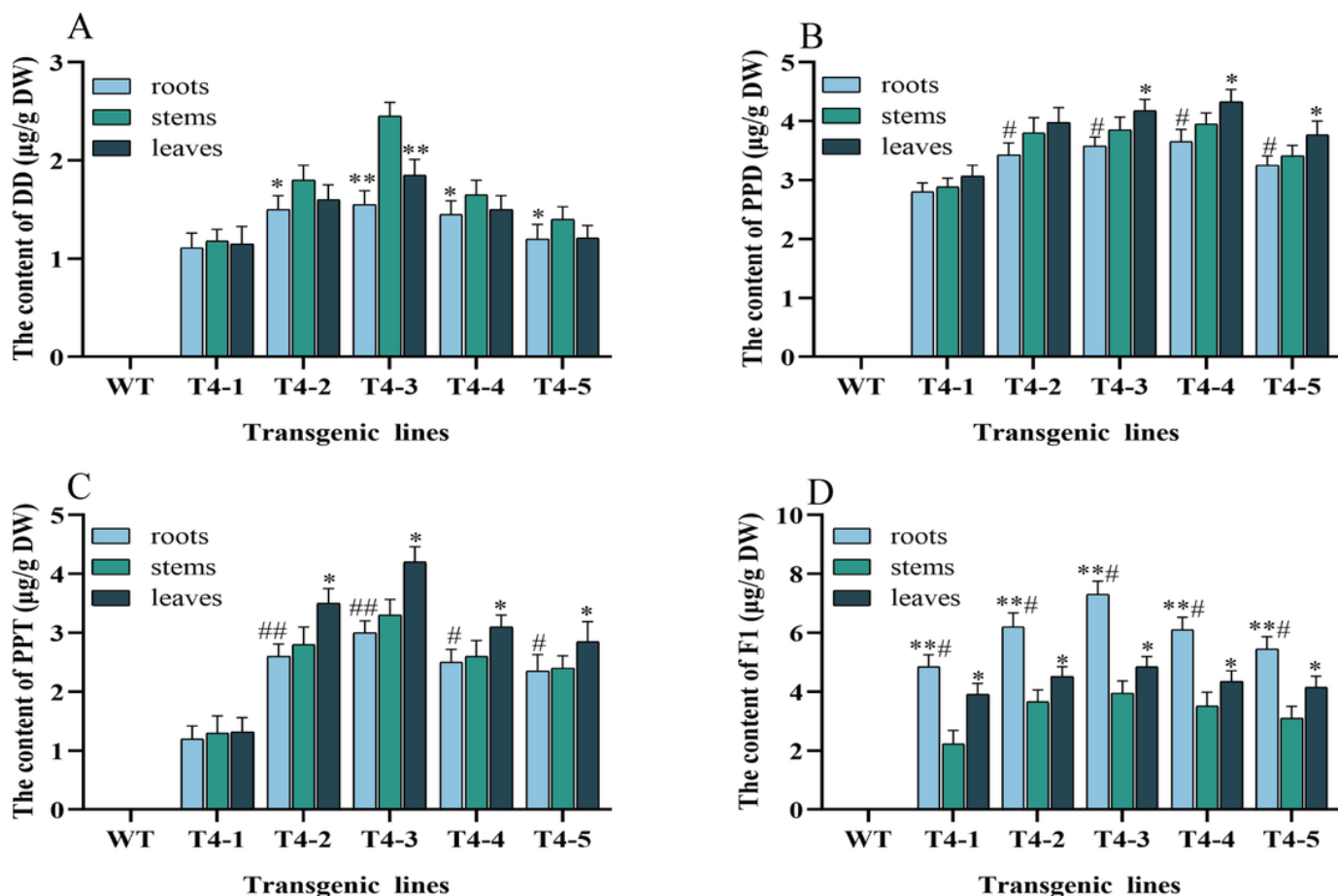


Figure 10

The contents of dammarenediol-II, protopanaxadiol, protopanaxatriol and ginsenoside F1 in *PnDDS/ CYP12H/ CYP6H/ UGT20*-transgenic tobaccos. T4, *PnDDS/ CYP12H/ CYP6H/ UGT20*-transgenic plants. (A) dammarenediol-II content in transgenic tobaccos; (B) protopanaxadiol content in transgenic tobaccos; (C) protopanaxatriol content in transgenic tobaccos; (D) ginsenoside F1 content in transgenic tobaccos; DD, dammarenediol-II; PPD, protopanaxadiol; PPT, protopanaxatriol; F1, ginsenoside F1. Compared with stems, * $P < 0.05$, ** $P < 0.01$. Compared with leaves, # $P < 0.05$, ## $P < 0.01$.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FIGURES1.docx](#)
- [FIGURES2.docx](#)
- [FIGURES3.docx](#)
- [FIGURES4.docx](#)
- [FIGURES5.docx](#)
- [FIGURES6.docx](#)

- [FIGURES7.docx](#)
- [FIGURES8.docx](#)
- [FIGURES9.docx](#)