Cloning and Expression Analysis of the ZaGGPPS Gene from Zanthoxylum armatum (Postprint)

Authors: Guan Shuwen, Wang Yi, Hao Jiabo, Yuan Xiaolong, Lu Bin, Li Xianzhong

Date: 2019-08-26T00:00:00+00:00

Abstract

To elucidate the molecular mechanism of terpenoid metabolism in Zanthoxylum armatum and the influence of grafting on its flavor, this study designed specific primers based on transcriptome data and employed RT-PCR to clone the full-length cDNA sequence of a novel geranylgeranyl pyrophosphate synthase (GGPPS) gene from Zanthoxylum armatum, designated as ZaGGPPS. Bioinformatics analysis of the ZaGGPPS gene was conducted using NCBI, ProParam, SignalP 4.1 server, DNAMAN, and MEGA 7.0 software, and its expression levels were compared between grafted and seedling trees. The results demonstrated that ZaGGPPS contains a complete cDNA open reading frame (ORF) of 1,086 bp, encoding 361 amino acids. The relative molecular weight of the protein is 39,079.14 Da, with a theoretical isoelectric point (pI) of 6.38. Blast alignment revealed that this protein belongs to the GGPPS family, containing two aspartate-rich motifs characteristic of GGPPS proteins, namely "DDXXXXD" and "DDXXD" , as well as five characteristic functional domains. Phylogenetic analysis indicated that Zanthoxylum armatum is closely related to Rutaceae species such as sweet orange (Citrus sinensis), clementine (C. clementina), and pomelo (C. maxima). Quantitative real-time PCR analysis revealed that the expression levels of ZaGGPPS in Zanthoxylum armatum, in descending order, were: leaves of seedling trees, leaves of grafted trees, stems of seedling trees, and stems of grafted trees. Geranylgeranyl pyrophosphate synthase is a key enzyme in the terpenoid biosynthetic pathway in Zanthoxylum armatum, and grafting can affect the expression level of ZaGGPPS in leaves and stems. The cloning and analysis of ZaGGPPS provide a theoretical foundation for further investigation into the molecular mechanism of aroma formation in Zanthoxylum armatum and for breeding superior varieties through molecular biology approaches.

Full Text

Preamble

Cloning and Expression Analysis of the ZaGGPPS Gene from Zanthoxylum armatum

Shuwen Guan[1],[2], Yi Wang[2]*, Jiabo Hao[2], Xiaolong Yuan[2], Bin Lu[2], Xianzhong Li[1]

[1]College of Forestry, Southwest Forestry University, Kunming 650224, China
[2]Key Laboratory of Forestry Plant Cultivation and Utilization, Key Laboratory of Rare and Endangered Forest Plants of State Forestry Administration, Yunnan Academy of Forestry, Kunming 650201, China

Abstract

To elucidate the molecular mechanism of terpenoid metabolism in Zanthoxylum armatum and the influence of grafting on its flavor profile, this study designed specific primers based on transcriptome data and cloned a novel full-length cDNA sequence of geranylgeranyl pyrophosphate synthase (GGPPS) from Z. armatum using RT-PCR, designated as ZaGGPPS. Bioinformatics analysis of ZaGGPPS was performed using NCBI, ProParam, SignalP 4.1 server, DNAMAN, and MEGA 7.0 software, and its expression levels were compared between grafted and seedling trees. The results revealed that ZaGGPPS contains a complete open reading frame (ORF) of 1,086 bp, encoding 361 amino acids. The predicted protein has a relative molecular mass of 39,079.14 Da and a theoretical isoelectric point (pI) of 6.38. BLAST analysis indicated that the protein belongs to the GGPPS family, containing two characteristic aspartate-rich motifs ("DDXXXXD" and "DDXXD") and five conserved functional domains. Phylogenetic analysis showed that Z. armatum is closely related to Rutaceae species including sweet orange (Citrus sinensis), clementine mandarin (C. clementina), and pomelo (C. maxima). Quantitative real-time PCR demonstrated that ZaGGPPS expression levels in Z. armatum tissues, from highest to lowest, were: seedling leaves, grafted leaves, seedling stems, and grafted stems. Geranylgeranyl pyrophosphate synthase is a key enzyme in the terpenoid biosynthetic pathway of Z. armatum, and grafting can significantly affect ZaGGPPS expression in both leaves and stems. The cloning and characterization of ZaGGPPS provide a theoretical foundation for further investigation into the molecular mechanisms of aroma formation in Z. armatum and for breeding improved varieties through molecular biological approaches.

Keywords: Zanthoxylum armatum, geranylgeranyl pyrophosphate synthase, cloning, bioinformatics analysis

---

Zanthoxylum armatum belongs to the Rutaceae family and the genus Zanthoxylum, which comprises approximately 250 species worldwide, with 39 species and

14 varieties found in China (Du et al., 2013). Commonly cultivated species include Z. armatum, Z. schinifolium, and Z. bungeanum (Chen et al., 2018). Z. armatum is widely distributed in Guizhou, Guangxi, and Yunnan provinces, offering significant economic and ecological benefits. Its fruits are used as culinary spices and traditional medicine, while the species features robust root systems and tolerance to drought and poor soils (Zhang et al., 2010; Wang et al., 2016). Consequently, cultivation of Z. armatum has become an important strategy for employment and income generation in economically disadvantaged and arid regions of western China (Shu et al., 2018).

The essential oil of Z. armatum represents the key component of its aroma, with terpenoid compounds accounting for 67.122% of the oil, including -pinene, -pinene, -caryophyllene, and germacrene. Therefore, terpenoids constitute crucial aromatic components in Z. armatum (Zhang et al., 2010). Terpenoids, built from isoprene structural units, represent the largest class of plant metabolites and participate in numerous physiological and biochemical processes, including respiration, photosynthesis, growth and development, defense, signal transduction, and reproduction. Many terpenoid compounds possess significant economic and medicinal value (Qi et al., 2016). Geranylgeranyl pyrophosphate synthase (GGPPS) plays a pivotal role in terpenoid biosynthesis (Wei et al., 2016). Plant terpenoid synthesis occurs in two stages: first, the methylerythritol phosphate (MEP) and mevalonic acid (MVA) pathways generate dimethylallyl diphosphate (DMAPP) and isopentenyl pyrophosphate (IPP). Subsequently, GGPPS catalyzes three distinct reactions: (1) one molecule of farnesyl pyrophosphate (FPP) with one molecule of IPP; (2) one molecule of geranyl pyrophosphate (GPP) with two molecules of IPP; or (3) one molecule of DMAPP with three molecules of IPP. All three reactions ultimately produce geranylgeranyl pyrophosphate (GGPP) (Yao et al., 2017). GGPP not only regulates carbon flux in plant isoprenoid metabolism but also serves as a universal precursor for numerous compounds, including polyprenols, gibberellins, chlorophyll, abscisic acid, carotenoids, coenzyme Q, strigolactones, polyphenols, vitamin E, and vitamin K, and participates in protein prenylation (Yamamura et al., 2014; Li et al., 2015; Zhang et al., 2015; Han, 2015). Thus, GGPPS is essential for terpenoid synthesis in plants. Li et al. (2019) reported that upregulation of NtGGPPS1 in tobacco increased diterpenoid and plastid pigment content, while Wei et al. (2016) found that specific silencing of NtGGPPS1 in tobacco K326 resulted in dwarfism, slow growth, reduced plastid pigment content, and delayed flowering. To date, GGPPS genes have been cloned from various plants including Brassica oleracea (Xue et al., 2018), Nicotiana tabacum (Wei et al., 2016), Camellia sinensis (Yao et al., 2017), Cephalotaxus oliveri (Qi et al., 2016), and Blumea balsamifera (Xia et al., 2016), but no studies have reported the ZaGGPPS gene from Z. armatum. This study clones and characterizes the ZaGGPPS gene to provide a theoretical basis for further investigation into the molecular mechanisms of aroma formation in Z. armatum and for molecular breeding of superior varieties.

1.1 Materials and Instruments

Materials: Young shoots, leaves, and stems were collected from grafted and seedling Z. armatum trees in the arboretum of Yunnan Academy of Forestry (102°44 42.57 , 25°08 50.67), immediately preserved in liquid nitrogen, transported to the laboratory, and stored at -80°C.

Instruments: HiFi DNA polymerase (TransGen, Beijing); pUMT vector, T4 ligase, and DH5 competent cells (Sangon Biotech, Shanghai); RNA extraction kit (Qiagen); plasmid extraction kit (Cwbiotech); Super RT kit (Takara). PCR thermal cycler (Bio-Rad), real-time fluorescence quantitative PCR system (ABI), and centrifuge (Beckman).

1.2 RNA Extraction and cDNA Synthesis

Leaves and stems from seedling and grafted Z. armatum trees were ground into fine powder in liquid nitrogen. Total RNA was extracted using a plant RNA extraction kit according to the manufacturer' s protocol. RNA quality was assessed using 1% agarose gel electrophoresis, and concentration was measured using a NanoDrop™ 2000 spectrophotometer. Intact RNA with appropriate concentration was selected for cDNA synthesis using a reverse transcription kit with 1 g of total RNA, following the manufacturer' s instructions. The synthesized cDNA was stored at -20°C, while remaining RNA was kept at -80°C.

1.3 Cloning of ZaGGPPS Gene

Using the GGPPS gene from sweet orange (Citrus sinensis, NCBI accession: XP_006466719.1) as a template, local BLAST searches were performed against Z. armatum transcriptome data to obtain the Z. armatum geranylgeranyl pyrophosphate synthase gene (ZaGGPPS) sequence. Specific primers were designed based on this sequence: ZaGGPPS1F (5' -ATGACTTGTGTGAATATCGG-3') and ZaGGPPS1R (5' -TCAATTCTGCCTATAAGCAA-3'). Using cDNA from young shoots of grafted and seedling Z. armatum as templates and ZaGGPPS1F/ZaGGPPS1R as primers, the full-length ZaGGPPS gene fragment was amplified with HiFi DNA polymerase. After electrophoresis verification, the target fragment was ligated into the pUMT vector, confirmed by PCR, and sent to Sangon Biotech for sequencing.

1.4 Protein Sequence Analysis of ZaGGPPS

The cDNA sequence of ZaGGPPS was analyzed using NCBI ORF Finder to obtain the amino acid sequence. Physicochemical properties were predicted using ProParam. SignalP 4.1 server was used to predict signal peptides. The amino acid sequence was searched against NCBI to retrieve homologous plant

GGPPS protein sequences. Sequence similarity was analyzed using DNAMAN, and a phylogenetic tree was constructed using MEGA 7.0.

1.5 Transcription Pattern Analysis

After successful synthesis of the first cDNA strand, specific primers TZaGGPPSF (5' -ACGCCAAAACTTAAACGCCG-3') and TZaGGPPSR (5' -GAGAGTTTCTTCTTTGGTAA-3') were used to detect ZaGGPPS expression in stems and leaves of grafted and seedling trees, with ubiquitins (NCBI accession: MK953729) as the reference gene. SYBR Green (Invitrogen) was used to detect PCR products. The 25 L reaction mixture contained: 12.5 L 2×SYBR Green master mix, 0.5 L each of forward and reverse primers (10 M), 1 L template cDNA, and 10.5 L ddH O. PCR was performed on an ABI 7300 thermal cycler (Applied Biosystems, Foster City, CA, USA) with the following program: denaturation (95°C, 10 min), amplification quantification (45 cycles of 95°C for 15 s, 57°C for 10 s, and 72°C for 15 s with single fluorescence measurement), melting curve analysis (60°C to 95°C at 0.1°C increments with continuous fluorescence measurement), and final cooling to 40°C. The elongation factor 1-alpha (EF1a) gene served as an internal control for normal expression. Each sample was analyzed with at least two independent biological replicates and three technical replicates per biological replicate to ensure reproducibility and reliability. Relative gene expression was calculated using the comparative Ct method, where f = 2^(-ΔΔCt), ΔΔCt = (Ct_target_gene_test - Ct_reference_gene_test) - (Ct_target_gene_control - Ct_reference_gene_control).

2.1 RNA Extraction and Cloning of ZaGGPPS

Total RNA was successfully extracted from young shoots of grafted and seedling Z. armatum using a plant total RNA extraction kit and reverse-transcribed into cDNA. Using ZaGGPPS1F and ZaGGPPS1R as specific primers, the full-length ZaGGPPS gene (NCBI accession: MK953731) was successfully cloned. Analysis with NCBI ORF Finder revealed that ZaGGPPS contains a complete full-length cDNA open reading frame (ORF) of 1,086 bp, encoding 361 amino acids (Fig. 1 [Figure 1: see original paper]).

2.2 Protein Sequence Analysis of ZaGGPPS

The ZaGGPPS ORF was predicted using NCBI ORF Finder, and protein physicochemical properties were analyzed using ProParam. The results showed that the protein comprises 361 amino acid residues with a molecular weight of 39,079.14 Da and a theoretical pI of 6.38. BLAST analysis against NCBI revealed high similarity between the deduced 361-amino-acid sequence and functionally characterized GGPPS proteins from sweet orange (Citrus sinensis), clementine mandarin (C. clementina), pomelo (C. maxima), and white pear (Pyrus bretschneideri). Specifically, ZaGGPPS shares 82.32% similarity

with CsGGPPS from sweet orange, 82.04% with CcGGPPS from clementine mandarin, 82.04% with CmGGPPS from pomelo, and 73.53% with PbGGPPS from white pear. Sequence analysis identified two conserved aspartate-rich motifs characteristic of GGPPS proteins, "DDXXXXD" and "DDXXD" (where X represents any amino acid), as well as five characteristic functional domains (I-V) of polyprenyl synthases (Fig. 2 [Figure 2: see original paper]).

Note: Za. Zanthoxylum armatum; Cc. Citrus clementina (Accession No: XP_006425738.1); Cs. C. sinensis (Accession No: XP_006466719.1); Cm. C. maxima (Accession No: AJT59420.1); Pb. Pyrus bretschneideri (Accession No: XP_009379182.1).

The deduced ZaGGPPS protein sequence was aligned against NCBI to retrieve GGPPS sequences from other plants. Multiple sequence alignment was performed using Cluster W in MEGA 7.0, and a phylogenetic tree was constructed using the Neighbor-Joining algorithm with 1,000 bootstrap replicates (Fig. 3 [Figure 3: see original paper]). The results demonstrated that ZaGGPPS clusters with GGPPS proteins from sweet orange, clementine mandarin, and pomelo.

2.3 Transcription Pattern Analysis of ZaGGPPS

Young shoots from grafted and seedling Z. armatum were collected and preserved in liquid nitrogen. Total RNA was extracted from stems and leaves of grafted trees and seedling trees, and reverse-transcribed into cDNA. Specific primers TZaGGPPSF and TZaGGPPSR were used to quantify ZaGGPPS expression in these tissues (Fig. 4 [Figure 4: see original paper]). The results showed that ZaGGPPS expression was highest in seedling leaves, followed by grafted leaves, then seedling stems, and lowest in grafted stems.

3 Conclusion

Previous studies have demonstrated that although GGPPS constitutes a large gene family with significant divergence among homologs—such as differences in subcellular localization and expression patterns of encoded proteins—certain homologs play crucial regulatory roles in terpenoid metabolism and significantly affect terpenoid synthesis in plants (Wang et al., 2018). Multiple sequence alignment revealed high similarity between Z. armatum GGPPS and GGPPS proteins from other plants. The amino acid sequence contains five conserved domains common to GGPPS proteins, as well as two specific functional motifs: the first aspartate-rich motif (FARM) and the second aspartate-rich motif (SARM), with sequences "DDLPCMD" and "DDILD" , respectively. FARM and SARM represent critical active sites where GGPPS catalyzes the binding of Mg^{2} -bridged diphosphate groups to substrate molecules (Song & Poulter, 1994; Hemmi et al., 2003). Studies by Kai et al. (2010) and Wang et al. (2010) reported that SmGGPPS from Salvia miltiorrhiza and CgGGPPS from hazel (Corylus avellana) both exhibited higher expression in leaves than in stems.

Our transcription pattern analysis revealed consistent results, with ZaGGPPS showing higher expression in leaves than in stems for both grafted and seedling trees. However, grafting downregulated ZaGGPPS expression in both leaves and stems compared to seedling trees, suggesting that grafting can influence ZaGGPPS expression in Z. armatum.

The primary flavor characteristics of Zanthoxylum species are pungency and aroma. Aroma components are represented by volatile oils containing alcohols, alkenes, aldehydes, esters, ketones, and other volatile substances, which serve as key indicators for quality evaluation. In Z. armatum, terpenoid aroma components are dominated by monoterpenes, followed by diterpenes, including limonene, eucalyptol, terpineol, phellandrene, pinene, myrcene, sabinene, germacrene, and caryophyllene (Chen et al., 2018). Current research on GGPPS in plant aroma has primarily focused on tobacco and tea plants. For instance, Lin et al. (2014) found that increased carotenoid and diterpenoid content in tobacco leaves enhanced aroma intensity. Yao et al. (2017) reported that CsGGPPS expression in tea buds and leaves increased with maturity, explaining why leaves of certain maturity are preferred for processing. Wang et al. (2019) suggested that the CsGGDPS7 gene is closely associated with terpenoid aroma compound synthesis in tea. However, no studies have investigated the impact of GGPPS on Z. armatum aroma. This study provides a theoretical foundation for future research on the molecular mechanisms of aroma formation in Z. armatum and for manipulating terpenoid content and peel aroma quality through genetic engineering.

Aroma components from Zanthoxylum species possess substantial economic and medicinal value. In pharmaceuticals, these components promote transdermal drug absorption and can be incorporated into topical formulations. In food applications, they serve as non-toxic preservatives to replace synthetic compounds such as sodium benzoate and potassium sorbate. In agriculture, they effectively inhibit pest reproduction, reducing the need for chemical pesticides (Chen et al., 2018). Therefore, while strengthening basic research on Zanthoxylum aroma components, we must also enhance product development and utilization to promote ecological and economic sustainability.

References

Chen Q, Tao XB, Huang YL, et al., 2018. Research progress on the aroma of Chinese prickly ash [J]. *Chinese Condiment*, 43(1):189-194.

Du LJ, Zheng GH, Niu XQ, 2013. Research progress in biological activities of *Zanthoxylum* L. [J]. *Chinese Journal of Tropical Crops*, 34(5): 995-999.

Hemmi H, Noike M, Nakayama T, et al., 2003. An alternative mechanism of product chain-length determination in type III geranylgeranyl diphosphate synthase [J]. *European Journal of Biochemistry*, 270(10): 2186-2194.

Han LM, 2015. Clone and expression characterization of a new gene encod-

ing geranylgeranyl pyrophosphate synthase in *Isatis indigotica* Fortune [J]. *Genomics and Applied Biology*, 34(6):1172-1178.

Kai G, Liao P, Zang T, et al., 2010. Characterization, expression profiling, and functional identification of a gene encoding geranylgeranyl diphosphate synthase from *Salvia miltiorrhiza* [J]. *Biotechnology and Bioprocess Engineering*, 15(2): 236-245.

Lin SF, Wang RG, Zhang XL, et al., 2014. Cloning and tissue expression analysis of geranylgeranyl pyrophosphate synthase small subunit gene from *Nicotiana tabacum* [J]. *Tobacco Science and Technology*, 58(2):70-75.

Li ZF, Wei P, Xia YZ, et al., 2015. Whole genome identification and analysis of tobacco GGPPS synthase gene family [J]. *Tobacco Science and Technology*, 48(6):1-8.

Li F, Wang R, Wu MZ, et al., 2019. Study on NtGGPPS1 function in *Nicotiana tabacum* based on overexpression technology [J]. *Journal of Henan University (Natural Science Edition)*, 49(2):180-185.

Qi XQ, Hu CX, Li DW, 2016. Cloning and sequence analysis of geranylgeranyl pyrophosphate synthase gene from *Cephalotaxus oliveri* Mast. [J]. *Jiangsu Agricultural Sciences*, 44(4):79-82.

Song L, Poulter CD, 1994. Yeast farnesyl-diphosphate synthase: site-directed mutagenesis of residues in highly conserved prenyltransferase domains I and II [J]. *Proceedings of the National Academy of Sciences USA*, 91(8): 3044-3048.

Shu ZY, Wang JY, Gong W, et al., 2018. Water and fertilizer effects on antioxidant enzyme activities of *Zanthoxylum armatum* seedlings under waterlogging stress [J]. *Journal of Sichuan Agricultural University*, 36(2):203-209, 216.

Wang Y, Miao Z, Tang K, 2010. Molecular cloning and functional expression analysis of a new gene encoding geranylgeranyl diphosphate synthase from hazel (*Corylus avellana* L. Gasaway) [J]. *Molecular Biology Reports*, 37(7): 3439-3444.

Wei P, Meng LJ, Chen QS, et al., 2016. Cloning and functional analysis of geranylgeranyl pyrophosphate synthase gene NtGGPPS1 from *Nicotiana tabacum* [J]. *Tobacco Science and Technology*, 49(4):8-15.

Wang JY, Gong W, Bao XL, et al., 2016. Coupling effects of water and fertilizer on diurnal variation of photosynthesis of *Zanthoxylum bungeanum* Maxim 'hanyuan' seedling leaf [J]. *Acta Ecologica Sinica*, 36(5):1321-1330.

Wang Z, Li F, Jin LF, et al., 2018. Research progress of geranylgeranyl pyrophosphate synthase gene in higher plants [J]. *Plant Physiology Journal*, 54(4):557-564.

Wang Z, Huang XJ, Cao HL, et al., 2019. Cloning and expression analysis of CsGGDPS7 gene in tea plant (*Camellia sinensis*) [J]. *Acta Botanica Boreali-Occidentalia Sinica*, 39(1):32-41.

Xia QF, Zhao Z, Liu HC, et al., 2016. Cloning and analysis of geranyl pyrophosphate synthase (GPPS) sequence of *Blumea balsamifera* L. DC based on transcriptome information [J]. *Journal of Mountain Agriculture and Biology*, 35(4):23-29.

Xue SL, Jiang M, Chang JQ, et al., 2018. Cloning and expression analysis of geranylgeranyl pyrophosphate synthase gene (*BoaGGPPS1*) in *Brassica oleracea* var. *alboglabra* [J]. *Jiangsu Journal of Agricultural Sciences*, 34(2):259-265.

Yamamura Y, Mizuguchi Y, Taura F, et al., 2014. Transcriptional activation of a geranylgeranyl diphosphate synthase gene, GGPPS2, isolated from *Scoparia dulcis* by treatment with methyl jasmonate and yeast extract [J]. *Journal of Natural Medicine*, 68(4):748-753.

Yao XQ, Yue C, Yang GY, et al., 2017. Cloning and expression analysis of geranylgeranyl diphosphate synthase gene CsGGDPS in tea plant [J]. *Journal of Tea Science*, 37(1):86-96.

Zhang Y, Peng YH, Zeng DQ, et al., 2010. Insecticidal activity of essential oil from *Zanthoxylum armatum* fructification against two mosquito species [J]. *Guihaia*, 30(2):274-279.

Zhang M, Su P, Liu YJ, et al., 2015. Cloning and bioinformatics analysis of geranylgeranyl diphosphate synthase gene of *Tripterygium wilfordii* [J]. *China Journal of Chinese Materia Medica*, 40(6):1066-1070.

Note: Figure translations are in progress. See original paper for figures.