Molecular cloning and characterization of farnesyl diphosphate synthase from *Rosa rugosa* Thunb

Guo Wei*, Yudie Chen*, Jianwen Wang, and Liguo Feng
College of Horticulture and Landscape Architecture, Yangzhou University, Yangzhou, China.
*These authors contributed equally to this work.

Corresponding Author:
Liguo Feng
Yangzhou University, Yangzhou, Jiangsu, 225009, China
Email address: lgfeng@yzu.edu.cn

I suggest that you annotate this information on the cladogram image and provide additional details regarding it in your results section?

Abstract

Rosa rugosa, a renowned ornamental plant, is cultivated for its essential oil containing valuable monoterpenes, sesquiterpenes, and other compounds widely used in the floriculture industry. Farnesyl diphosphate synthase (FPPS) is a key enzyme involved in the biosynthesis of sesquiterpenes and triterpenes. In this study, we successfully cloned and characterized a full-length cDNA encoding FPPS from *R. rugosa*, identified as RrFPPS1. Phylogenetic analysis showed that RrFPPS1 belong to the Rosaceae-FPPS clade. Moreover, Transcriptome and RT-qPCR analysis revealed that RrFPPS1 had tissue-specific expression patterns. Subcellular localization studies revealed that RrFPPS1 were cytoplasmic. *In vitro* enzymatic assays showed that RrFPPS1 produced farnesyl diphosphate (FPP) using isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) as substrates, which would provide precursor for sesquiterpenes and triterpenes biosynthesis in plant. Additionally, our research uncovered that RrFPPS1 expression was enhanced under salt treatment conditions. These significant findings contribute to an improved understanding of terpene biosynthesis in *R. rugosa* and open new avenues for advancements in horticultural practices and fragrance industries. The knowledge gained from this study can potentially lead to the development of enhanced varieties of *R. rugosa* with improved aroma, medicinal properties, and better resilience to environmental stressors.

acting on.... (cited one function of this biosynthesis)
belongs
transcriptome
was
provide a precursor
in the plant
These two sentences look similar. I suggest making a change to avoid visual similarity as it might cause confusion.

Introduction

Terpenoids are the largest and most diverse groups of natural products, with over 80,000 structures identified to date (Christianson 2017). They ~~played~~ play various physiological and ecological roles that essential for plant growth and development (Dudareva 2007; Zhou & Pichersky 2020). Notably, they are involved in various physiological processes such as hormone biosynthesis, photosynthesis, electron transport and membrane components. Moreover, terpenoids also act as important signaling and defense mediators between plants and their environments (Vandermoten

et al. 2009). Owing to their remarkable properties, they have significant applications in the food, pharmaceutical, and agricultural industries, making them highly valuable resources (Hohl 2004).

In plants, terpenoids are synthesized from two independent pathways: the methylerythritol phosphate pathway (MEP) pathway in plastids (Lichtenthaler et al. 1997; Rodriguez-Concepcion & Boronat 2002) and the mevalonate (MVA) pathway in the cytosol or peroxisomes (Chen et al. 2011; Sapir-Mir et al. 2008). The MVA-pathway mainly provides the precursor farnesyl diphosphate (FPP; C_{15}) for the formation of sesquiterpenes, triterpenes and sterols (J N Valitova 2016; Natalia Dudareva 2005), while the plastidic MEP-pathway provides the precursors geranyl diphosphate (GPP; C_{10}), geranylgeranyl diphosphate (GGPP;C_{20}) and geranylfarnesyl diphosphate (GFPP;C_{25}) for the biosynthesis of monoterpenes, diterpenes and sesterterpens(Chen et al. 2019; Duilioarigoni 1997; Huang et al. 2017; Schwender et al. 1997) . These reactions are all catalyzed by enzymes commonly referred to as prenyltransferases. FPPS belongs to the family of short-chain prenyltransferases that also includes GPP synthase (GPPS), GGPP synthase (GGPPS) and GFPP synthase (GFPPS) (Chen et al. 2021; Cui et al. 2022; Song et al. 2023). These enzymes function at the branch points of isoprenoid metabolism and can thus play a regulatory role in controlling IPP flux into different terpenoid families (Koyama 1998).

Farnesyl diphosphate synthase (FPPS) is a vital enzyme responsible for catalyzing the head-to-tail condensation of IPP and DMAPP to produce FPP (Pichersky & Raguso 2018). Numerous studies have investigated the biochemical and molecular biology characteristics of FPPS in a variety of plant species, such as Arabidopsis, Rice, Maize, *Ginkgo biloba*, *Chimonanthus praecox*, and *Hedychium coronarium* (Cunillera et al. 1997; Kazutsuka Sanmiya a & Yamamoto 1997; Lan et al. 2013; Larkins 1996; Peng Wang1 2004; Xiang et al. 2010). In Arabidopsis, *AtFPPS1* is primarily expressed in roots and flowers, while *AtFPPS2* is expressed mainly in flowers (Cunillera et al. 1996). In *Maize*, FPPS1 plays a key role in the biosynthesis of farnesyl diphosphate for ubiquinone production, while FPPS2 is primarily responsible for the biosynthesis of isofucosterol. FPPS3, on the other hand, is closely linked to the production of the sesquiterpene volatile (*E*)-*β*-caryophyllene in response to the root-chewing herbivore *Diabrotica virgifera virgifera* (Tang et al. 2022).

R. rugosa, native to Eastern Asia, has been cultivated in the floriculture industry worldwide and is renowned for its essential oil extracted from petals, which is used in perfumes, cosmetics, and pharmaceuticals (Bendahmane et al. 2013; Caser & Scariot 2022). *R. rugosa* 'Zi Zhi' is a continuously flowering variety that offers significant possibilities for yielding more essential oil (Bai et al. 2021). The essential oil of *R. rugosa* mainly consists of monoterpenes, sesquiterpenes and other valuable compounds (Ueyama et al. 1990). Despite significant knowledge about monoterpene biosynthesis, little is known about the molecular mechanisms of sesquiterpene biosynthesis in *R. rugosa*. Plant growth and development are often influenced by environmental stresses, with salinity stress being one of the most severe challenges (Zhu 2001). FPPS is closely associated with the production of terpenoids, however, studies examining the impact of abiotic stress on FPPS, particularly in *R. rugosa*, are limited (Tian et al. 2018).

In this study, we aimed to identify potential FPPS from *R. rugosa* as the first objective and

What does IPP mean? Write down what the abbreviation stands for so that any reader of your article can recognize what you're writing about.

give a space

give a space

give a space

I suggest rewriting this paragraph as it was a bit confusing.

It was not mentioned before? Put what it means the first time the abbreviation appears. Then just use the abbreviation.

Gene or protein? A person who doesn't know it may have the idea that it's a gene or a protein. I always suggest writing down what it is. Example: AtFPPS1 gene.

has

about the

to characterize *FPPS* genes underlying the molecular mechanism of FPP-derived terpene formation in *R. rugosa* as the second objective. Additionally, we also aimed to investigate whether FPPS from *R. rugosa* exhibits any response to salt stress. Using a combination of bioinformatics, transcriptomic, molecular and biochemical approaches, we were able to successfully identify *RrFPPS1* as the key gene responsible for FPP formation in *R. rugosa*.

Has this been previously documented, or is it the designated nomenclature for the FPPS genes identified in R. rugosa? If there is existing literature on this topic, please discuss it within the introduction section.

Materials & Methods

Plant Materials

The experiment utilized plant materials ~~were~~ collected from the resource nursery of Yangzhou University (32.391°N, 119.419°E), Yangzhou, China. Three-year-old cutting seedlings of *R. rugosa* 'ZiZhi' were used to collect plant tissues from leaves and flowers. The samples were immediately frozen in ~~the~~ liquid nitrogen and then stored at -80℃ for RNA extraction.

RNA Extraction and Gene Expression Analysis

Total RNA was extracted from the collected plant tissues using the TaKaRa MiniBEST Plant RNA Extraction Kit (TaKaRa) as per the manufacturer's instructions. The concentration and quality of the total RNA were evaluated through NanoDrop 1000 analysis and gel electrophoresis. For reverse transcription, first-strand cDNA was synthesized from total RNA using a PrimeScript™ II 1st Strand cDNA Synthesis Kit (Takara) following the manufacturer's protocol. The HiScript III RT SuperMix for qPCR (Vazyme) was used for RT-qPCR analysis as per the manufacturer's instructions.

The expression levels of RrFPPS in different tissues (leaves and flowers) were analyzed by Previous transcriptome data (Wang et al. 2022) of *R. rugosa* 'Zi Zhi' providing the per kilobase of exon model per million mapped fragments (FPKM) of *RrFPPSs* and reverse transcription quantitative real-time PCR (RT-qPCR). ChamQ SYBR qPCR Master Mix (Vazyme) was used for RT-qPCR on a CFX96 Real-time PCR platform (Bio-Rad). The internal control genes used were *5.8S rRNA*. The primers for *RrFPPS* and the internal control genes were designed using the Genscript online website (https://www.genscript.com) and synthesized by Tsingke Biotech (Beijing). The primer sequences are provided in the supplementary file (Supplementary Table 1).

A 20 μL reaction mixture containing 1 μL of cDNA template, 10 μM of the primers, and 10 μL of 2×ChamQ SYBR Color qPCR Master Mix was used according to the manufacturer's instructions. Reactions were performed by an initial incubation at 95 °C for 30 s and then cycled at 95 °C for 10 s and 60 °C for 30 s for 39 cycles. Three independent biological replicates with three technical replicates were prepared for each sample. The relative transcript levels of *RrFPPSs* in different tissues were calculated using the $2^{-\triangle\triangle Ct}$ method (Livak & Schmittgen 2001) based on the threshold cycle (Ct) values generated from the CFX Manager software ~~(~~Bio-Rad~~)~~. All the experiments were carried out at least in triplicate. Values were analyzed by Anova ~~(~~Graphpad Prism~~)~~. P-values less than 0.05 were considered significant.

from BioRad (REF)

using Graphpad Prism (REF)

In Supplementary Table 1, there are designations for 'F' and 'R.' I believe it would be beneficial to include in the methodology section that these correspond to 'forward' and 'reverse' (or is that not the case?). Furthermore, an explanation of the meaning of these letters should also be provided within the table.

I suggest that authors pay attention to how this abbreviation is written, as it is written in different ways throughout the text. Standardization is necessary.

In Figure 2, there are various organisms. It is necessary to provide a description of who they are here in the methodology. For instance, 'X sequences of ... were utilized to create...'

In-silico Analysis of RrFPPS and Phylogenetic Analysis

To predict RrFPPS, sequence homology and deduced amino acid sequence comparisons were carried out using ~~BLASP~~ BLASTP at NCBI from the proteome of *R. rugosa* 'Zi Zhi' (unpublished). Candidate genes were further confirmed by Pfam (http://pfam.xfam.org/) and the Conserved Domain Database. The physicochemical properties of RrFPPS were analyzed using the online software ~~Extasy~~ Expasy (https://web.expasy.org/protparam/). Multiple alignments of amino acid sequences were conducted with CLUSTALX and GENEDOC. Multiple sequence alignment of *RrFPPS* sequences, as well as the phylogenetic tree, was constructed by the neighbor-joining method of MEGA-X software (Kumar et al. 2018). The GenBank accessions for protein sequences of FPPS are shown in Supplementary Table 2.

using ProtParam tool from the online...

What substitution method was taken into consideration? It is necessary to mention.

This section appears somewhat unclear as it specifies the sources of the sequences and to which organisms these sequences belong. There is some confusion regarding the inputs under consideration.

Isolation and Cloning of the RrFPPS Coding Sequence

The open reading frame (ORF) of RrFPPS was amplified using PrimeSTAR Max DNA Polymerase (Takara) and PCR primers designed with NEB Tm Calculator (https://tmcalculator.neb.com/). Reverse-transcribed cDNA served as the PCR template. PCR products were purified using the FastPure Gel DNA Extraction Mini Kit (Vazyme), cloned into pEASY-Blunt cloning vectors (~~Transgen~~ TransGen Biotech), and transformed into E. coli Trans1-T1 (Transgen Biotech) competent cells. The PCR-positive clones were selected and sequenced by Sangon Biotech (Shanghai, China). The sequence information of RrFPPS is shown in Supplementary Table 3.

Scientific names are always italicized

When reviewing your supplementary materials 2 and 3, I have noticed a high degree of similarity, and I believe they are positioned inversely within the text. Supplementary material 2 is referenced as containing GenBank accessions but lacks the corresponding codes, only providing sequences. Conversely, supplementary material 3 includes this information. I suggest consolidating these details into a single file to avoid redundancy.

Subcellular Localization of the RrFPPS Proteins

The subcellular localization of RrFPPS was predicted using the bioinformatics analysis website WoLF PSORT (https://wolfpsort.hgc.jp/). The ORF of RrFPPS was amplified with specific primers, and inserted into the pCAMBIA 1300-35S-sGFP vector using the ClonExpress II One Step Cloning Kit (Vazyme) with Sac1/Xba1 restriction enzymes. The constructs and the empty vector (control) were transformed into Agrobacterium tumefaciens strain GV3101+P19 by the freeze-thaw method. Single positive Agrobacterium clones were grown in LB medium until the OD600 value reached 0.5-0.6 and then infiltrated into 5-6 week-old Nicotiana benthamiana leaves. The fluorescence of the tobacco plant leaves was examined 2 days after infiltration at 488 nm using an LSM 880 confocal microscope (Zeiss) to obtain images of the GFP fluorescence signal.

What does 'LB' stand for? I am using this format to suggest that for any abbreviation, its meaning should be explicitly written.

Scientific names are always italicized.

Purification of Recombinant RrFPPS Proteins and Enzymatic Assay

To perform enzymatic assays, RrFPPS was cloned downstream of the (His)6-tag sequences in pEasy Blunt E1 plasmids to express RrFPPS-His recombinant proteins. The constructed vectors were verified through DNA sequencing and transformed into chemically competent E. coli BL21(DE3) pLysS cells (Transgen). Single positive colonies were selected and grown in LB medium containing ampicillin (100 μg/mL) at 37 °C in a shaking incubator until the OD600 of the culture reached 0.4-0.6. Next, 0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG) was

Scientific names are always italicized.

added, and the culture was incubated at 16 °C and 200 rpm for 12 h to induce the recombinant proteins. The induced cells were collected by centrifugation at 5,000 g for 10 min and stored at -80°C.

The cells were resuspended in binding buffer (20 mM Tris-HCl, 500 mM NaCl, 10% glycerol, 10 mM imidazole, and 0.5 mM PMSF, pH 7.5) and sonicated to disrupt them. The lysate was centrifuged at 12,000 rpm for 30 min at 4°C, and Ni-NTA agarose (Sangon Biotech) was added to the supernatant for affinity purification following the manufacturer's instructions. To obtain desalted protein, the protein was added to the desalting buffer (250 mM MOPS, 100 mM MgCl2, 50% glycerol, pH 7.5) and then applied to a Sephadex Desalting Gravity Column (Sangon Biotech) according to the manufacturer's instructions.

To determine the in vitro activity of RrFPPS using GC-MS, the enzymatic activity assays were performed in a final volume of 100 µL reaction volume containing 66 µM IPP, 44 µM DMAPP, 10 µL desalted enzyme solution, and ddH2O up to 100 µL. After incubation at 30 °C for 2 hours, the assay mixture was hydrolyzed overnight at 30 °C using 1 unit of calf intestinal alkaline phosphatase (CIP) and 1 unit of Apyrase from potatoes. The volatile products were adsorbed by Headspace solid-phase microextraction using 30 µm (CAR/PDMS layer) and 50 µm (DVB layer) (Supelco Inc., Bellefonte, PA) overnight. The volatile compounds were collected from headspace and were subjected to GC–MS (Clarus SQ8T, PerkinElmer, USA) analysis. The experimental procedure involved heating the initial temperature to 50 ℃ and holding it for 1 min. It was then heated up to 120 ℃ at 5 ℃ min−1, increased to 200 ℃ at 8 ℃ min−1, and then increased to 250℃ at 12 ℃ min−1 for 7 min. The MS conditions included the emission current of 200 µA, the ionization energy of 70 eV, and the mass scan range was 29-600 amu.

The meaning of this abbreviation only appears in the results section, specifically in 'Functional characterization of the recombinant RrFPPS1.' It is necessary to have it explained beforehand

Another abbreviation without explanation.

Quantification of transcript levels of *RrFPPS* genes in salt treatment

In brief, three plants of *R. rugosa* were dug out from the sand, and their roots were treated with 0.1 M NaCl solution immediately for 1 h. Another three plants that were soaked in deionized water for 1 h were set as the control group (CK). The leaves of salt treatment and CK plants (three replications for each) were picked and frozen by liquid nitrogen immediately and stored at −80°C. RNA isolation, cDNA synthesis, and quantification were performed as previous.

How?

It was only described how it was prepared, but what about quantification? I believe it is not as clear as mentioned. I suggest making modifications to this paragraph.

Results

Identification of *RrFPPS* genes in Rosa genome

To investigate terpene biosynthesis in *R. rugosa*, particularly FPPS-derived terpene biosynthesis, we conducted a search for possible *FPPS* genes using the Arabidopsis thaliana FPPS (AAB07248.1) protein sequence as a query through BLASTP (Cunillera et al. 1996). The analysis revealed two potential FPPS (Rru05G067530 and Rru05G000330) that appeared to be full length, then they were designated as RrFPPS1 and RrFPPS2. The full-length cDNA sequence of RrFPPS1 was 1153 bp, including an 82 bp 5'-untranslated region and a 42 bp 3'-untranslated region. On the other hand, the full-length cDNA sequence of RrFPPS2 was 1449 bp, comprising a 104 bp 5'-untranslated region and a 316 bp 3'-untranslated region. The ~~open reading frames (ORFs)~~ of

I believe this statement is better suited for the methodology section.

As the methodology section has already addressed this, I believe there is no need to reiterate the meaning of the abbreviation.

RrFPPS1 and RrFPPS2 were each 1029 bp in length, encoding a protein of 342 amino acids. The predicted molecular mass of the RrFPPS1 protein was 39.57 kDa with a theoretical pI of 5.39. Similarly, the predicted molecular mass of the RrFPPS2 protein was 39.42 kDa with a theoretical pI of 5.52.

Sequence alignment analysis of RrFPPSs protein sequences with other plant FPPS proteins revealed the presence of five highly conserved regions (I-V) that are essential for substrate binding and catalytic activity, as typically seen in prenyltransferases. (Fig 1). These regions comprise the substrate binding pocket, substrate-Mg^{2+} binding site, catalytic site, and two aspartate-rich regions, namely Asp-Rich Motif 1 and 2. The first is FARM (first Asp-rich motif) in region II, with the sequence $DD_{XX(2-4)}D$ (D = aspartic acid and X = any amino acid), while the second is SARM (second Asp-rich motif) in region V, with the sequence $DD_{XX}D$. The FARM and SARM are suggested to function as binding sites for the diphosphate moieties of isopentenyl diphosphate (IPP) and allylic substrates (Kush 1996) (Fig 1).

The way it is currently written implies that this sentence should be placed in the discussion section. I suggest either rewriting it or relocating it to the 'Discussion' section.

To investigate the evolutionary relationships between RrFPPSs and FPPS proteins from other species, we selected typical FPPS proteins from *Rosa chinensis*, *Arabidopsis thaliana*, *Fragaria vesca*, *Malus domestica*, *Centella asiatica*, *Panax ginseng*, *Prunus persica* and *Potentilla anserine*, and constructed a phylogenetic tree ~~using the neighbor-joining method in MEGA X software(Kumar et al. 2018)~~. The phylogenetic analysis revealed that despite high sequence homology, all plant FPPSs were separated into two major clades: Clade of angiosperm, including *Rosa chinensis* G/FPPS1 (A0A2P6Q231), *Rosa chinensis* FPPS2 (A0A2P6QLH7), *Fragaria vesca* FPPS (XP_004294906), *Malus domestica* FPPS (AAM08927), *Centella asiatica* FPPS (AAV58896), *Lupinus albus* FPPS (P49351), *Humulus lupulus* FPPS (BAB40665), *Panax ginseng* FPPS (AAY87903), *Prunus persica* FPPS (XP_007211529), *Vitis vinifera* FPPS (AAX76910), *Potentilla anserina* FPPS (XP_050367607), *Hevea brasiliensis* FPPS (AAM98379), *Artemisia annua* FPPS (AAD17204), *Oryza sativa* FPPS (O04882), *Zea mays* FPPS (AAQ14871), *Salvia miltiorrhiza* FPPS (AEZ55677), *Paeonia lactiflora* FPPS (AKJ26301), *Jasminum sambac* FPPS (AIY24422); the other is gymnosperm, including *Ginkgo biloba* FPPS (AAR27053) and *Picea abies* FPPS (ACA21460). Our analysis showed that RrFPPs have high homologous to FPPSs from other species. As RrFPPS2 showed 100% identity to RcG/FPPS1(*Rosa chinensis*), hence, we focus on the biochemical function of RrFPPS1 hereafter (Fig 2).

I believe it is unnecessary to write this, as it is already described in the methodology.

I believe that this entire section should be incorporated into the methodology, particularly the codes and their sources, as I have mentioned previously. The authors may also describe the species found in the cladogram, but as it is currently written, it conveys the impression of methodology.

Tissue expression profiling of the *RrFPPS* genes

To investigate the tissue-specific expression patterns of RrFPPS in *R. rugosa*, both transcriptome analysis and RT-qPCR were used to evaluate and quantify transcript levels. Transcriptome analysis showed that a high transcript of RrFPPS2 was detected in all tissues while RrFFPPS1 only detected a high transcript in S7 of the flower. In flowers, the expression level of RrFPPS2 was higher than RrFPPS1, and both genes had increased expression from S1 to S7, reaching the maximum level at S7 stage. Moreover, RrFPPS2 had the highest transcript in branches (Fig 3A). The RT-qPCR results of leaves and flowers from 3-year-old cutting seedlings also revealed that both RrFPPS1 and RrFPPS2 were expressed in two tissues, but with distinct expression patterns. Interestingly,

What would S1 - S7 represent? These terms have only appeared here in the results section. I believe this should be included in the 'Materials and Methods' section.

Please be mindful that from this point onward, all figure numbers in the text are out of sequence. For instance, in the upcoming sections, you reference Fig 3A and Fig 3B; however, in Figure 3, there are no designations of A and B. I suggest reviewing all the figures and their corresponding references throughout the text.

both genes showed higher relative expression levels in flowers compared to leaves (Fig 3B).

Subcellular localization of the RrFPPS proteins

FPP is commonly to be synthesized through the cytosolic mevalonic acid (MVA) pathway (32). In this study, the WoLFPSORT program was used to predict the subcellular localization of RrFPPSs, which was found to be localized in the cytosol. To validate this prediction, RrFPPSs were inserted into the pCAMBIA 1300-35S-sGFP vector without the stop codon, which resulted in the fusion of these genes with the GFP protein, driven by the 35S promoter, when expressed in *N.benthamiana* leaves. The subcellular localization of RrFPPS2 was consistent with the result in a previous report (Conart et al. 2023). The subcellular localization results showed that RrFPPS1 was found to localize in the cytosol (Fig 4). These findings support the notion that RrFPPS1 is responsible for producing FPP in the cytosol, which is then used in the biosynthesis of FPP-derived compounds.

Functional characterization of the recombinant RrFPPS1

Recent Research has revealed that RcG/FPPS1 from *Rosa chinensis* exhibits a remarkable bifunctionality as a geranyl/farnesyl diphosphate synthase producing both GPP and farnesyl diphosphate (FPP) in the cytosolic mevalonate (MVA) pathway. Building upon this knowledge, in our study, we aimed to investigate the function of RrFPPS1. To achieve this, the recombinant proteins were expressed in *E. coli* and purified using nickel-nitrilotriacetic acid (Ni-NTA). The purified proteins were then subjected to *in vitro* enzyme activity assays. The prenyl diphosphates were then converted to their corresponding alcohols using alkaline phosphatase and analyzed using Gas chromatography mass spectrometry (GC/MS). Our analyses unequivocally demonstrated that when the purified RrFPPS1 protein was incubated with DMAPP and IPP, only farnesol was detected as the product. These findings conclusively indicate that RrFPPS1 encodes a fully functional FPP synthase, elucidating its crucial role in FPP biosynthesis (Fig 5).

Expression analysis of *RrFPPS* genes in salt treatment

Terpenoids play crucial ecological roles in the interactions between plants and their environments. To gain insights into the response of *R. rugosa* plants to salt stress, we subjected *R. rugosa* plants to salt treatment and analyzed the expression levels of *RrFPPS1* and *RrFPPS2* using quantitative real-time polymerase chain reaction (RT-qPCR). Remarkably, the results revealed ~~that~~ a significant increase in the mRNA levels of *RrFPPS1* and *RrFPPS2* in *R. rugosa* leaves under salt treatments, indicating they may play defensive roles under salt stress conditions.

Discussion

The primary objective of this study was to comprehensively investigate the terpene biosynthesis in *R. rugosa*, with a particular focus on FPP biosynthesis. These results led to the identification of RrFPPSs in the *R. rugosa* 'Zi Zhi' genome (unpublished data), representing a significant achievement in understanding the terpene biosynthetic pathway in this plant species. The tissue

The references are not following a standardized style. Please select one style and consistently apply it throughout.

This section should have been included in the methodology (the initial part is already written there).

It would be more appropriate to incorporate this excerpt into the 'Discussion' section.

I recommend a rewrite, as this entire section appears to have been extracted from an already written abstract, with sentences indicating an introduction, methodology, results, and conclusion.

I suggest revising this paragraph.
I believe that Figure 7 is relevant to this topic. If indeed, the first and second sentences of the Figure 7 caption could be incorporated here.

expression profiling revealed that RrFPPS1 was expressed in all tissues with distinct expression patterns. The subcellular localization of RrFPPS1 revealed that it is a cytosolic protein, and the phylogenetic analysis showed that RrFPPS1 is more similar to RcFPPS2, suggesting potential functional similarities between these enzymes (Conart et al. 2023).

What is the correct way to be written? As described in the results?

Despite a number of FPPSs that have been characterized from both angiosperms and gymnosperms, the identification of RrFPPS1 is a significant achievement as FPPS is a key enzyme involved in terpene biosynthesis in *R. rugosa*. The presence of conserved regions (FARM and SARM), typical of prenyltransferases, in RrFPPS1, indicates that these proteins are functionally active and may play a crucial role in FPPS-derived terpene biosynthesis in *R. rugosa*. FARM in domain II has been shown to play a role in the determination of chain length for the resulting prenyl pyrophosphates through the presence of a conserved Phe residue located five amino acids upstream of the DDxxD motif. SARM in domain V is the binding site for the homoallylic substrate IPP (Ohnuma et al. 1996). Therefore, the conservation of these regions suggests that RrFPPS shares a similar mechanism of substrate binding and catalysis with other plant FPPS enzymes. The phylogenetic analysis revealed that RrFPPS1 is more similar to RcFPPS2. These findings suggest that RrFPPS1 ~~share~~ shares a close genetic relationship with other FPPSs within the Rosaceae family, further supporting its crucial role in terpene biosynthesis.

Tissue expression profiling of RrFPPS in *R. rugosa* provided valuable insights into the tissue-specific expression patterns of RrFPPS1. The analysis revealed a preferential accumulation of RrFPPS1 mRNAs in lateral branches, followed by flowers and roots. The significantly higher expression levels of RrFPPS1 in flowers compared to leaves suggest its essential role in terpene biosynthesis, known to contribute to the enchanting flower scent that is characteristic of roses. The findings indicated that RrFPPS1 is not only a key enzyme of floral scent volatiles biosynthesis, but also more likely has an important role in the biosynthesis of other terpenoids, such as terpene alcohols, sterols, ubiquinone, photosynthetic pigments, and plant hormones, all of which are essential for plant growth, development and immunity (Degenhardt et al. 2009).

Functional characterization of the recombinant RrFPPS1 protein revealed that it was able to produce FPP from IPP and DMAPP substrates, indicating that it is functionally active and ~~play~~ plays a crucial role in the synthesis of important secondary metabolites. Sesquiterpenes, in particular, are responsible for imparting pungent or aromatic flavors to specific plant tissues and are released in response to wounding and insect attack (Rosenkranz et al. 2021). In the context of roses, sesquiterpenes have been traditionally used as flavors and fragrances and possess various biological properties, such as anticancer and antimalarial activities (Zhou et al. 2020). Our observations suggest that two *RrFPPSs* gene ~~expression~~ expressions were significantly responsive to salt treatment. RrFPPSs may provide a substrate for sesquiterpenoid and/or triterpenoids biosynthesis to cope with the challenging environment. These observations highlight the potential ecological significance of RrFPPSs in the adaptation of *R. rugosa* to its habitat.

In my opinion, this sentence appears to be disconnected from the current context of the text. I recommend either rephrasing it or relocating it within a different paragraph of the discussion.

Despite the progress made in this study, it is essential to acknowledge existing limitations. In order to fully elucidate the specific roles and functions of *RrFPPS*1 gene, Virus-induced gene silencing (VIGS) or CRISPR-Cas9 techniques may be necessary to manipulate the expression of the

RrFPPS1 precisely. These approaches would provide a more comprehensive insight in *R. rugosa*. With this additional information at our disposal, we can enhance our comprehension of the underlying mechanisms governing terpene biosynthesis and potentially explore novel targets for genetic modification or breeding programs in the future.

I believe there could be a more in-depth discussion regarding why RrFPPS clustered with the isoform 2 of RcFPPs.

Conclusions

The cloning and characterization of the farnesyl pyrophosphate synthases from *R. rugosa* reported here are involved in terpene biosynthesis. The relative expression of RrFPPS1 was the highest in the flower and the relative expression of RrFPPS2 was the highest in the branch. Subcellular localization analysis revealed that RrFPPS1 and RrFPPS2 were cytoplasmic. These findings open up new avenues for the overexpression of the native *FPPS* gene into *R. rugosa*. Such advancements hold great promise for the floriculture industry and may lead to the development of rose varieties with enhanced fragrance, medicinal properties, and resilience to environmental stresses.

Is there any existing literature demonstrating that this has been proposed for another species? If so, please provide a discussion based on this to illustrate its potential validity for R. rugosa.

I suggest, in a certain manner, bringing your objectives here in order to describe that they have been addressed.

Acknowledgements

All authors thank the anonymous reviewers and editors for their valuable comments and suggestions.

thanks :)

Funding

Please pay attention to the sections starting from 'Funding,' as I believe the font sizes may differ.

The study was funded by the National Natural Science Foundation of China (Grant No. 32171861). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the National Natural Science Foundation of China (Grant No. 32171861).

Competing Interests

The authors declare that they have no competing interests.

AUTHOR CONTRIBUTION:

- Guo Wei conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Yudie Chen performed the experiments, analyzed the data, prepared figures and/or tables, authored drafts of the article, and approved the final draft.
- Jianwen Wang analyzed the data and approved the final draft.
- Liguo Feng conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability: The raw data are available in the Supplemental File.

References

Please pay close attention to the final formatting of the references, as some of them have line breaks.

Bai M, Liu J, Fan C, Chen Y, Chen H, Lu J, Sun J, Ning G, and Wang C. 2021. KSN heterozygosity is associated with continuous flowering of Rosa rugosa Purple branch. *Horticulture Research* 8. 10.1038/s41438-021-00464-8

Bendahmane M, Dubois A, Raymond O, and Bris ML. 2013. Genetics and genomics of flower initiation and development in roses. *J Exp Bot* 64:847-857. 10.1093/jxb/ers387

Caser M, and Scariot V. 2022. The Contribution of Volatile Organic Compounds (VOCs) Emitted by Petals and Pollen to the Scent of Garden Roses. *Horticulturae* 8. 10.3390/horticulturae8111049

Chen F, Tholl D, Bohlmann J, and Pichersky E. 2011. The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *Plant J* 66:212-229. 10.1111/j.1365-313X.2011.04520.x

Chen Q, Jiang T, Liu Y-X, Liu H, Zhao T, Liu Z, Gan X, Hallab A, Wang X, He J, Ma Y, Zhang F, Jin T, Schranz ME, Wang Y, Bai Y, and Wang G. 2019. Recently duplicated sesterterpene (C25) gene clusters in Arabidopsis thaliana modulate root microbiota. *Science China Life Sciences* 62:947-958. 10.1007/s11427-019-9521-2

Chen Q, Li J, Ma Y, Yuan W, Zhang P, and Wang G. 2021. Occurrence and biosynthesis of plant sesterterpenes (C25), a new addition to terpene diversity. *Plant Communications* 2. 10.1016/j.xplc.2021.100184

Christianson DW. 2017. Structural and Chemical Biology of Terpenoid Cyclases. *Chem Rev* 117:11570-11648. 10.1021/acs.chemrev.7b00287

Conart C, Bomzan DP, Huang XQ, Bassard JE, Paramita SN, Saint-Marcoux D, Rius-Bony A, Hivert G, Anchisi A, Schaller H, Hamama L, Magnard JL, Lipko A, Swiezewska E, Jame P, Riveill G, Hibrand-Saint Oyant L, Rohmer M, Lewinsohn E, Dudareva N, Baudino S, Caissard JC, and Boachon B. 2023. A cytosolic bifunctional geranyl/farnesyl diphosphate synthase provides MVA-derived GPP for geraniol biosynthesis in rose flowers. *Proc Natl Acad Sci U S A* 120:e2221440120. 10.1073/pnas.2221440120

Cui B, Zhao X, Chen G, Qin S, Liu Y, Han Z, Yan Q, Han L, Liu S, Liu S, and Liu H. 2022. Functional Analysis of Two Terpene Synthase Genes Isolated from the Flowers of Hosta 'So Sweet'. *Horticulturae* 8. 10.3390/horticulturae8050447

Cunillera N, Arró M, Delourme D, Karst F, Boronat A, and Ferrer A. 1996. Arabidopsis thaliana Contains Two Differentially Expressed Farnesyl-Diphosphate Synthase Genes. *Journal of Biological Chemistry* 271:7774-7780. 10.1074/jbc.271.13.7774

Cunillera N, Boronat A, and Ferrer A. 1997. The Arabidopsis thaliana FPS1 gene generates a novel mRNA that encodes a mitochondrial farnesyl-diphosphate synthase isoform. *J Biol Chem* 272:15381-15388. 10.1074/jbc.272.24.15381

Degenhardt J, Kollner TG, and Gershenzon J. 2009. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry* 70:1621-1637. 10.1016/j.phytochem.2009.07.030

Dudareva JGN. 2007. The function of terpene natural products in the natural world. *Nature Chemical Biology* 3:408-411.

Duilioarigonis, Christophlatzel, Wolfgangeisenreich, Adelbertbacher, Andmeinharth. Zenk. 1997. Terpenoid biosynthesis from 1-deoxy-D-xylulose in higher plantsby intramolecular skeletal rearrangement. PNAS 94:10600 – 10605.

Hohl SAHaRJ. 2004. Isoprenoids: Remarkable Diversity of Form and Functio. *Lipids* 39:293–309.

Huang AC, Kautsar SA, Hong YJ, Medema MH, Bond AD, Tantillo DJ, and Osbourn A. 2017. Unearthing a sesterterpene biosynthetic repertoire in the Brassicaceae through genome mining reveals convergent evolution. *Proceedings of the National Academy of Sciences* 114. 10.1073/pnas.1705567114

J N Valitova AGS, F V Minibayeva. 2016. Plant sterols: Diversity, biosynthesis, and physiological functions. *Biochemistry* 81(8):819-834.

Kazutsuka Sanmiya a TIb, Makoto Matsuoka a, Mitsue Miyao , and Yamamoto N. 1997. Cloning of a cDNA that

encodes farnesyl diphosphate synthase and the
blue-light-induced expression of the corresponding gene in the leaves of
rice plants. *Biochimica et Biophysica Acta* 1350:240-246.
Koyama KOaT. 1998. Enzymatic Aspects of Isoprenoid Chain ElongationKyozo. *Chemical Reviews* 98:1263-1276.
Kumar S, Stecher G, Li M, Knyaz C, and Tamura K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis
across Computing Platforms. *Mol Biol Evol* 35:1547-1549. 10.1093/molbev/msy096
Kush KAaA. 1996. Cloning and characterization of cDNA encoding farnesyl diphosphate synthase from rubber tree
(Hevea brasiliensis). *Plant Molecular Biology* 30:935-946.
Lan JB, Yu RC, Yu YY, and Fan YP. 2013. Molecular cloning and expression of Hedychium coronarium farnesyl
pyrophosphate synthase gene and its possible involvement in the biosynthesis of floral and wounding/herbivory
induced leaf volatile sesquiterpenoids. *Gene* 518:360-367. 10.1016/j.gene.2013.01.007
Larkins CPLaBA. 1996. Identification of a maize endosperm-specific cDNA encoding farnesyl
pyrophosphate synthetase. *Gene* 171:193-196.
Lichtenthaler HK, Schwender J, Disch A, and Rohmer M. 1997. Biosynthesis of isoprenoids in higher plant
chloroplasts proceeds via a mevalonate-independent pathway. *FEBS Lett* 400:271-274. 10.1016/s0014-
5793(96)01404-4
Livak KJ, and Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and
the 2(-Delta Delta C(T)) Method. *Methods* 25:402-408. 10.1006/meth.2001.1262
Natalia Dudareva SA, Irina Orlova, Nathalie Gatto, Michael Reichelt, David Rhodes,Wilhelm Boland, and Jonathan
Gershenzon. 2005. The nonmevalonate pathway supports bothmonoterpene and sesquiterpene formationin
snapdragon
flowers. *PNAS* 102:933-938.
Ohnuma S, Hirooka K, Hemmi H, Ishida C, Ohto C, and Nishino T. 1996. Conversion of product specificity of
archaebacterial geranylgeranyl-diphosphate synthase. Identification of essential amino acid residues for chain length
determination of prenyltransferase reaction. *J Biol Chem* 271:18831-18837. 10.1074/jbc.271.31.18831
Peng Wang1 ZL, Liang Guo1, Wenchao Li1, Min Chen2, Yan Pi1, Yifu Gong3, Xiaofen Sun1,and Kexuan Tang1.
2004. Cloning and Functional Analysis of a cDNA Encoding
Ginkgo biloba Farnesyl Diphosphate Synthase. *Molecules and Cells* 18:150-156.
Pichersky E, and Raguso RA. 2018. Why do plants produce so many terpenoid compounds? *New Phytol* 220:692-702.
10.1111/nph.14178
Rodriguez-Concepcion M, and Boronat A. 2002. Elucidation of the methylerythritol phosphate pathway for isoprenoid
biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiol* 130:1079-
1089.
10.1104/pp.007138
Rosenkranz M, Chen Y, Zhu P, and Vlot AC. 2021. Volatile terpenes - mediators of plant-to-plant communication.
Plant J 108:617-631. 10.1111/tpj.15453
Sapir-Mir M, Mett A, Belausov E, Tal-Meshulam S, Frydman A, Gidoni D, and Eyal Y. 2008. Peroxisomal
localization
of Arabidopsis isopentenyl diphosphate isomerases suggests that part of the plant isoprenoid mevalonic acid
pathway is compartmentalized to peroxisomes. *Plant Physiol* 148:1219-1228. 10.1104/pp.108.127951
Schwender J, Zeidler J, Groner R, Muller C, Focke M, Braun S, Lichtenthaler FW, and Lichtenthaler HK. 1997.
Incorporation of 1-deoxy-D-xylulose into isoprene and phytol by higher plants and algae. *FEBS Lett* 414:129-134.
10.1016/s0014-5793(97)01002-8
Song S, Jin R, Chen Y, He S, Li K, Tang Q, Wang Q, Wang L, Kong M, Dudareva N, Smith BJ, Zhou F, and Lu S.
2023. The functional evolution of architecturally different plant geranyl diphosphate synthases from geranylgeranyl
diphosphate synthase. *The Plant Cell* 35:2293-2315. 10.1093/plcell/koad083
Tang HV, Berryman DL, Mendoza J, Yactayo-Chang JP, Li QB, Christensen SA, Hunter CT, Best N, Soubeyrand E,
Akhtar TA, Basset GJ, and Block AK. 2022. Dedicated farnesyl diphosphate synthases circumvent isoprenoid-derived
growth-defense tradeoffs in Zea mays. *Plant J* 112:207-220. 10.1111/tpj.15941
Tian X, Wang Z, Zhang Q, Ci H, Wang P, Yu L, and Jia G. 2018. Genome-wide transcriptome analysis of the salt
stress
tolerance mechanism in Rosa chinensis. *PLoS One* 13:e0200938. 10.1371/journal.pone.0200938
Ueyama Y, Hashimoto S, Nii H, and Furukawa K. 1990. The essential oil from the flowers ofRosa rugosa Thunb.
var.plena Regel. *Flavour and Fragrance Journal* 5:219-222. 10.1002/ffj.2730050406
Vandermoten S, Haubruge E, and Cusson M. 2009. New insights into short-chain prenyltransferases: structural
features, evolutionary history and potential for selective inhibition. *Cell Mol Life Sci* 66:3685-3695. 10.1007/s00018-

009-0100-9
Wang J, Wang P, Xu M, Chen Y, and Feng L. 2022. Systematic Identification and Analysis of OSC Gene Family of Rosa rugosa Thunb. *Int J Mol Sci* 23. 10.3390/ijms232213884
Xiang L, Zhao K, and Chen L. 2010. Molecular cloning and expression of Chimonanthus praecox farnesyl pyrophosphate synthase gene and its possible involvement in the biosynthesis of floral volatile sesquiterpenoids. *Plant Physiol Biochem* 48:845-850. 10.1016/j.plaphy.2010.08.015
Zhou F, and Pichersky E. 2020. More is better: the diversity of terpene metabolism in plants. *Current Opinion in Plant Biology* 55:1-10. 10.1016/j.pbi.2020.01.005
Zhou L, Yu C, Cheng B, Wan H, Luo L, Pan H, and Zhang Q. 2020. Volatile compound analysis and aroma evaluation of tea-scented roses in China. *Industrial Crops and Products* 155. 10.1016/j.indcrop.2020.112735
Zhu J-K. 2001. Plant salt tolerance. *TRENDS in Plant Science* 6:66-71.

Figure 1

Sequence alignment of the deduced amino acid of *R.rugosa* FPPS and FPPSs from other related proteins.

This alignment was shaded using Genedoc software to show conserved amino acid residues in black and similar residues in red and blue with the similarity of 75% and more than 50% respectively. The five conserved domains of prenyltransferases are boxed and marked by Roman numerals(I-V), which are the substrate binding pocket, aspartate-rich regions 1, catalytic site, substrate-Mg 2+ binding site and aspartate-rich regions 2. The highly conserved first and second aspartate rich regions is presented in domains II and V.

For me is 'pink'. Is red or pink?

For me is 'dark blue'. Is black ou another color?

respectively (?)

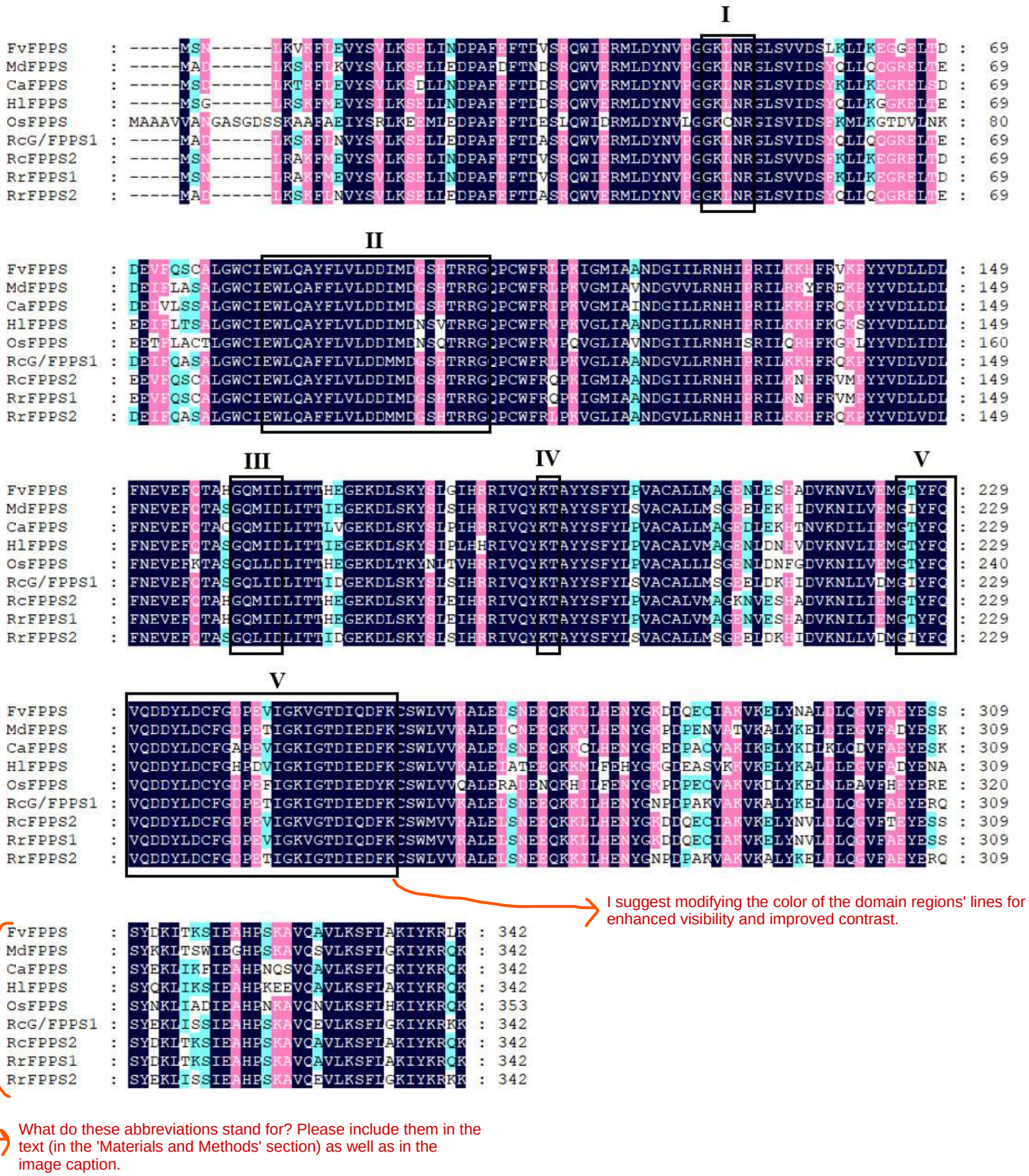

I
II
III
IV
V
V
FvFPPS
MdFPPS
CaFPPS
HlFPPS
OsFPPS
RcG/FPPS1
RcFPPS2
RrFPPS1
RrFPPS2
I suggest modifying the color of the domain regions' lines for enhanced visibility and improved contrast.
What do these abbreviations stand for? Please include them in the text (in the 'Materials and Methods' section) as well as in the image caption.

Figure 2

Phylogenetic tree of the amino acid sequences of FPPS of different organisms constructed by the neighbor-joining method on MEGAX.

Pay attention to the consistent spelling of program names, as some have been spelled differently throughout the text.

The numbers in the figure represent the genetic distance. Bootstrap values are shown as a percentage of 1000 replicates. RrFPPS1 and RrFPPS2 are marked by red dots. The scale bar on the bottom-left is representative of the degree of difference among sequences wherein a distance of 0.05 infers a 5% difference among sequences.

For this figure, I have some questions and suggestions to be addressed:

1) Where is A. thaliana mentioned in the text? It is not present in the image.

2) Why are only isoforms 1 and 2 of R. chinensis and R. rugosa displayed? What about the other species?

3) The text mentions the names of eight species, but why are other names appearing? If the mentioned names are significant, please explain why they were chosen. It may be more effective to begin with a broader discussion before addressing the specific species.

4) Building on the previous point, if it's important, I suggest adding a star (or another symbol) next to those names in the image.

5) I recommend discussing the significance of FPPS being present in R. rugosa as well to emphasize the relevance of the research.

6) Consider marking the angiosperm and gymnosperm groups on the image.

7) In the "Discussion" section, it would be more interesting to delve further into the issue of angiosperms and gymnosperms because, as suggested, the findings could contribute to the floriculture industry. What would be the impact of the contribution from these groups, given that gymnosperms do not have flowers, yet the results indicate an increased expression of FPPS in flowers?

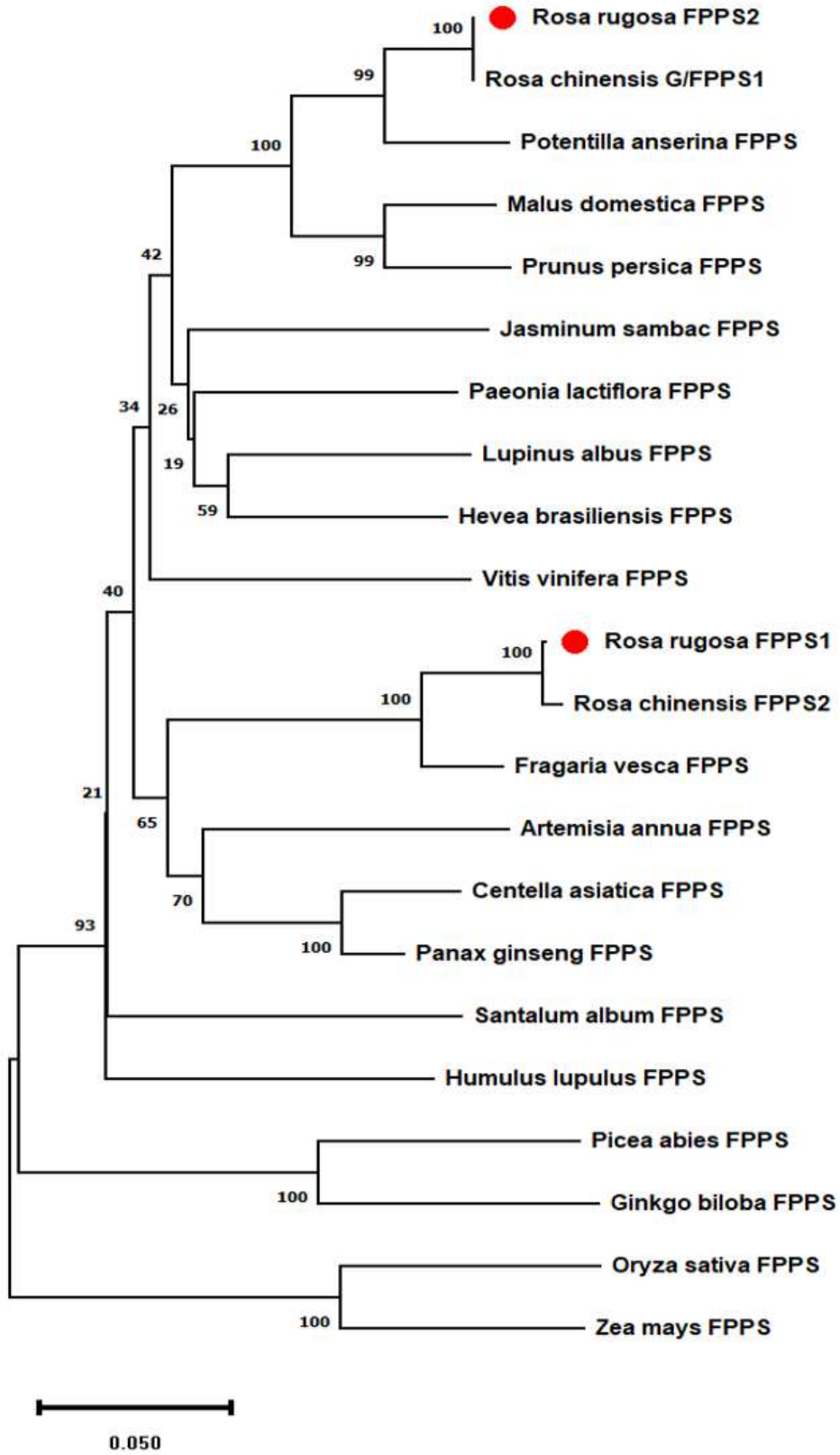
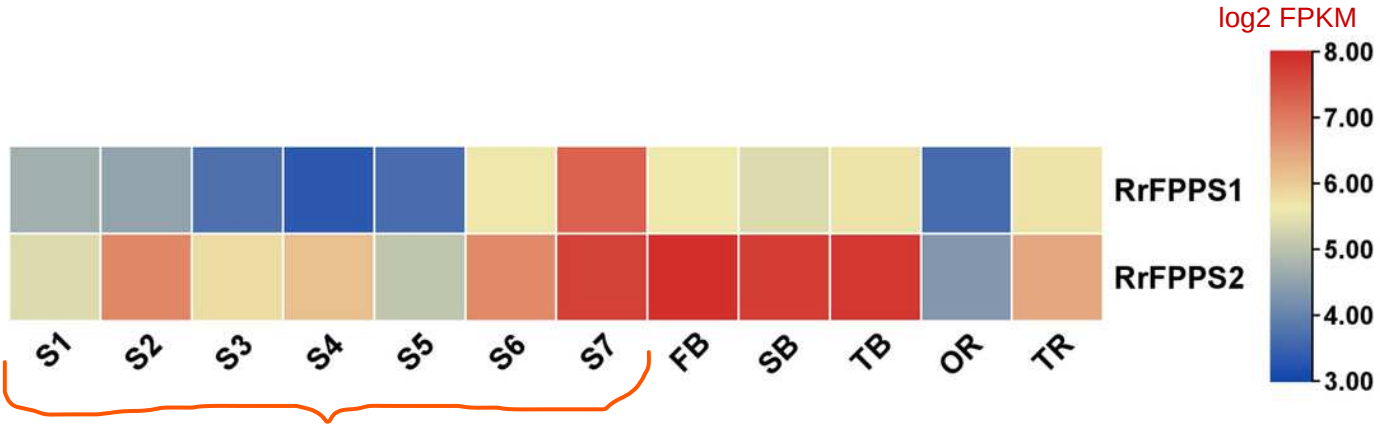

Rosa rugosa FPPS2
Rosa chinensis G/FPPS1
Potentilla anserina FPPS
Malus domestica FPPS
Prunus persica FPPS
Jasminum sambac FPPS
Paeonia lactiflora FPPS
Lupinus albus FPPS
Hevea brasiliensis FPPS
Vitis vinifera FPPS
Rosa rugosa FPPS1
Rosa chinensis FPPS2
Fragaria vesca FPPS
Artemisia annua FPPS
Centella asiatica FPPS
Panax ginseng FPPS
Santalum album FPPS
Humulus lupulus FPPS
Picea abies FPPS
Ginkgo biloba FPPS
Oryza sativa FPPS
Zea mays FPPS
100
99
100
42
99
34
26
19
59
40
100
100
21
65
70
93
100
100
100
0.050

Figure 3

What would the term 'digital' mean?

Digital gene expression profiles in the roots, lateral branches, and flowers of *R. rugosa* ‘Zi Zhi’.

A heat map was generated based on the normalized Log2 FPKM represented by the blue-yellow gradation. The numbers in the heat map were FPKM from RNA-seq data. The OR and TR columns represent the roots of open-air and tissue culture seedlings, respectively The columns of FB, SB, and TB represent the primary, secondary, and tertiary lateral branches, respectively. The columns of S1–S7 represent the seven flower stages, respectively.



The floral stages are not clearly explained in the text. I suggest addressing this in the text (in the 'Materials and Methods' section).

In the text, you use 'Fig 3A' and 'Fig 3B,' but there is no such differentiation here. Please, be mindful of this.

Figure 4

Expression of RrFPPS1 and RrFPPS2 in the flower and leaf of *R. rugosa* 'Zi Zhi'. obtain with RT-qPCR.

RrFPPS1 and *RrFPPS2* expression values were normalized to the levels of *5.8S RNA* expression in respective stage. Data are presented as mean ± SE, n = 3.

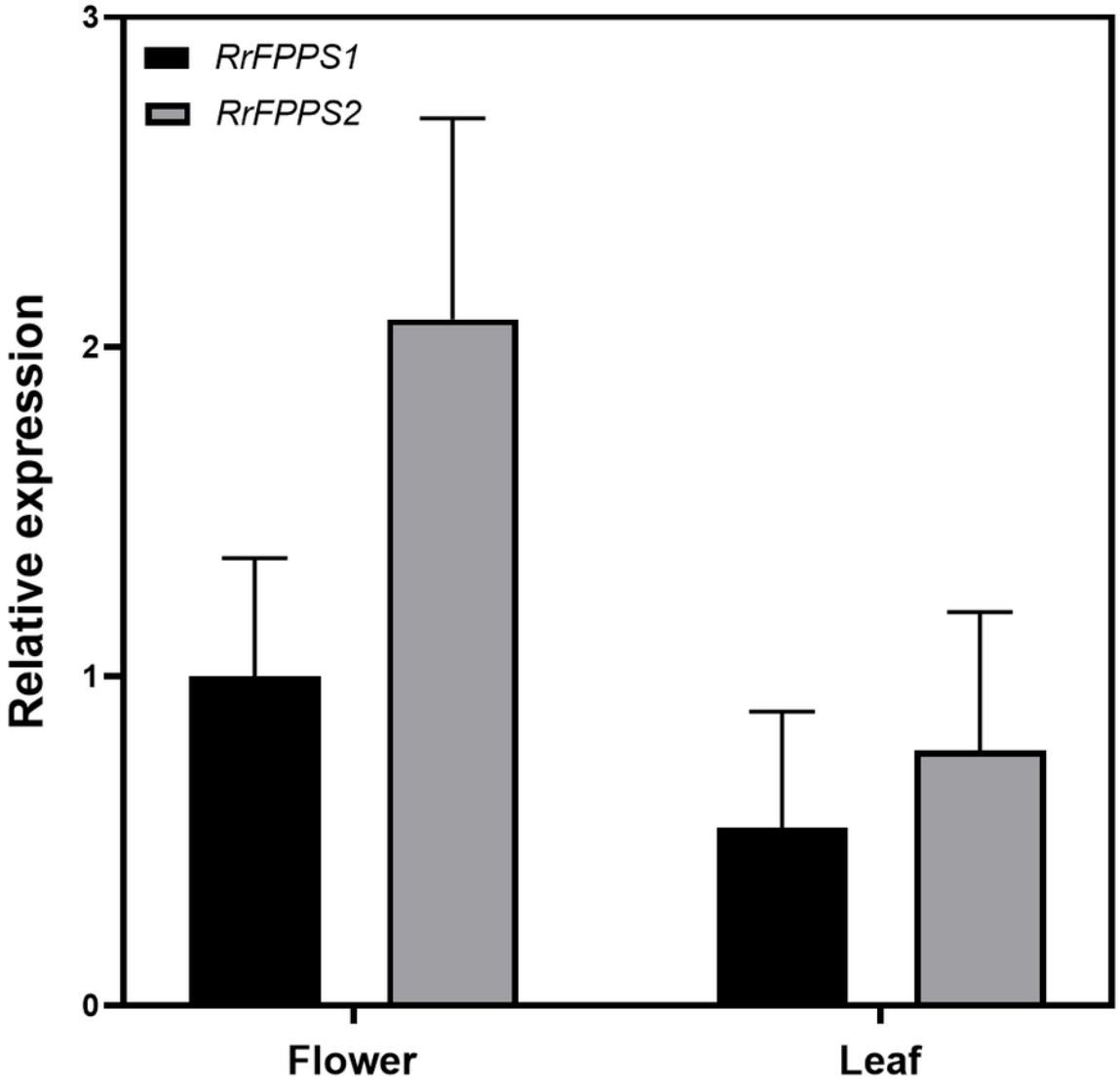


Figure 5

Scientific names are always italicized.

Subcellular localization of RrFPPS1 and RrFPPS2 in the N. benthamiana leaves.

GFP signals from the adaxial leaf surface were observed by using a confocal laser scanning microscope.

GFP , green fluorescence image;

Autofluorescence, chlorophyll autofluorescence image;

Merge, overlay of GFP and autofluorescence.

Would this serve as your control? If so, please make that explicit.

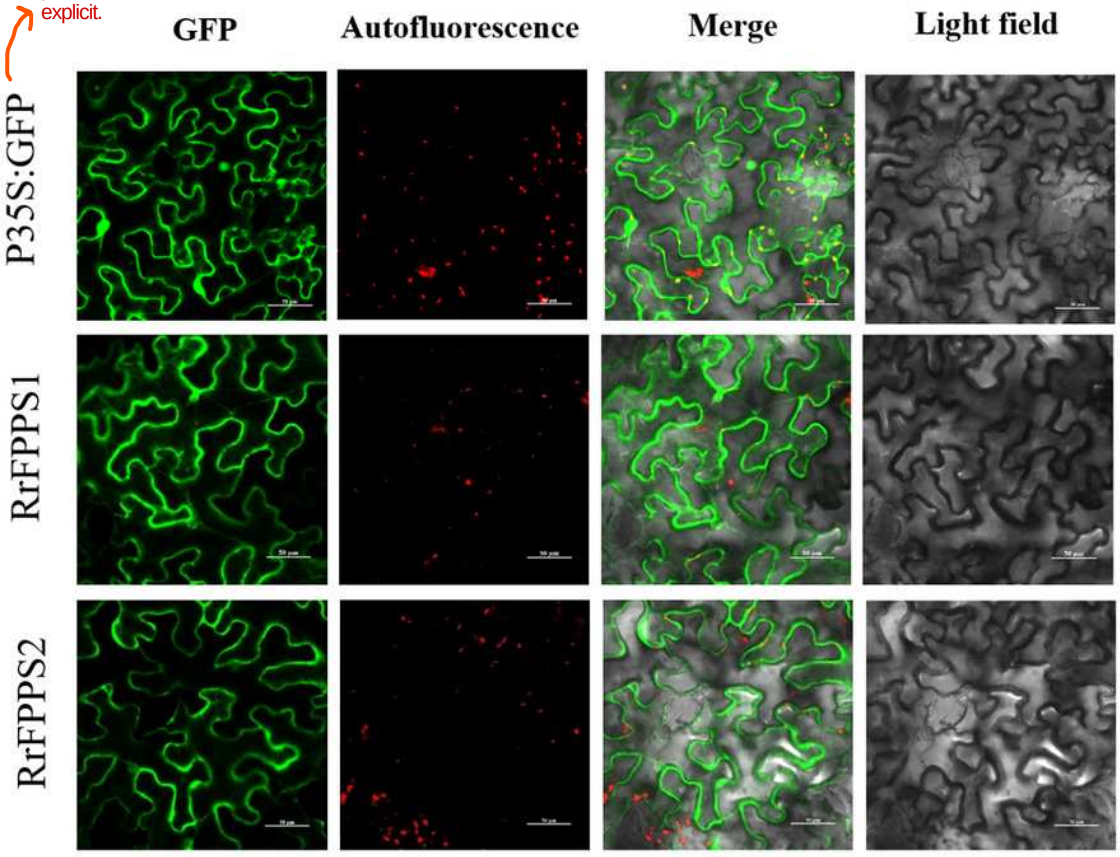


Figure 6

GC-MS analysis of the reaction products catalyzed by purified recombinant RrFPPS1 incubated with IPP and DMAPP.

a.The reaction products catalyzed by purified recombinant RrFPPS1 (IPP and DMAPP were added to the reaction mixture).

b.The mass spectrogram of the reaction products catalyzed by purified recombinant RrFPPS1.

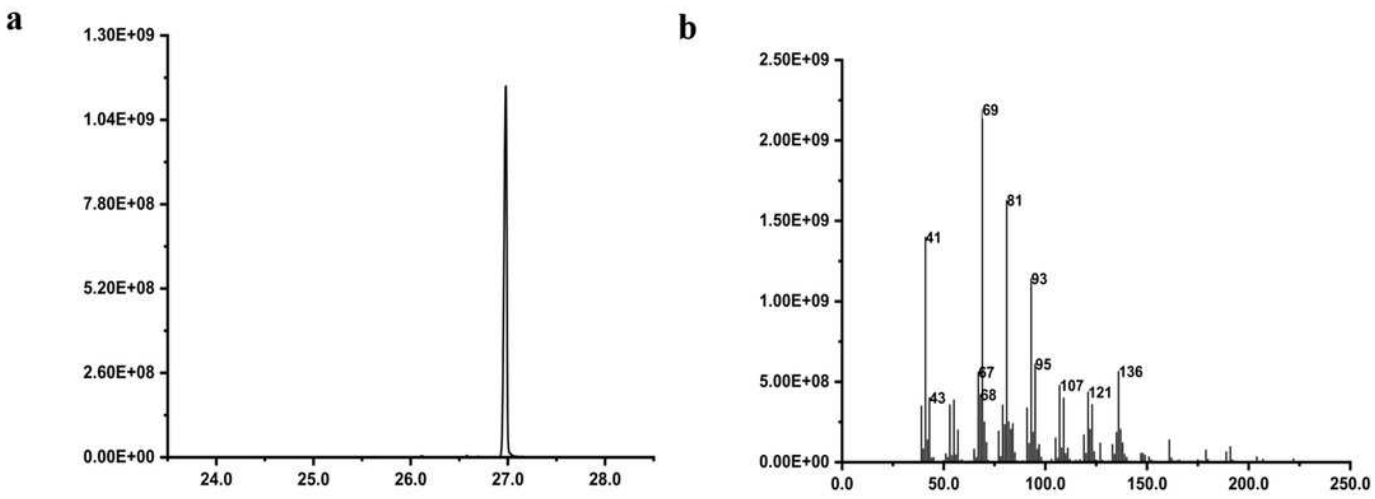


Figure 7

Change in RrFPPSs RNA levels in response to salt treatment.

RrFPPSs expression increases in response to salt treatment. Expression changes of the RrFPPSs gene were determined in *R. rugosa* treated with salt in comparison with that in control cultures after the initiation of treatment: 1 h. Data presented are $2^{-\Delta\Delta Ct}$ levels calculated relative to the special tissue. Data are presented as mean ± SE, n = 3. Asterisks indicate statistically significant differences relative control and salt treatment samples (t-test : *, P < 0.05; **, P < 0.01).