



Article

Systematic Identification and Analysis of OSC Gene Family of *Rosa rugosa* Thunb

Jianwen Wang [†], Pengqing Wang [†], Mengmeng Xu, Yudie Chen and Liguang Feng ^{*}

College of Horticulture and Landscape Architecture, Yangzhou University, Yangzhou 225009, China

^{*} Correspondence: lgfeng@yzu.edu.cn; Tel.: +86-514-8797-1026

[†] These authors contributed equally to this work.

Abstract: The oxidosqualene cyclase family of *Rosa rugosa* (*RrOSC*) provides a starting point for the triterpenoid pathway, which contributes to the medicinal value of the extraction of tissues of *Rosa rugosa*. However, the structure and function of key *RrOSC*s of active triterpenoids remain ambiguous. In this study, a total of 18 *RrOSC* members with conservative gene structures and motifs were identified based on the genome of *Rosa rugosa*. The *RrOSC*s were located on three chromosomes including two gene clusters that derived from gene replication. The phylogenetic relationship divided *RrOSC*s into six groups, and the *RrOSC*s of GI and GIV that were represented by lupeol or α -amyrin were identified as likely to include candidate genes for producing active triterpenoids. Considering the high expression or specific-tissue expression of the candidates, *RrOSC1*, *RrOSC10*, *RrOSC12*, and *RrOSC18* were considered the key genes. *RrOSC12* was identified in vitro as lupeol synthase. The results provided fundamental information and candidate genes for further illustration of the triterpenoid pathway involved in the pharmacological activities of *Rosa rugosa*.

Keywords: *Rosa rugosa* Thunb; oxidosqualene cyclase; biosynthesis; triterpenes



Citation: Wang, J.; Wang, P.; Xu, M.; Chen, Y.; Feng, L. Systematic Identification and Analysis of OSC Gene Family of *Rosa rugosa* Thunb. *Int. J. Mol. Sci.* **2022**, *23*, 13884. <https://doi.org/10.3390/ijms232213884>

Academic Editor: Frank M. You

Received: 17 September 2022

Accepted: 9 November 2022

Published: 11 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Sterols and triterpenes are important components of plant metabolites, and nearly 200 different skeletons of these compounds have been identified [1]. In addition to components of cell membranes and roles in plant growth and development, triterpenes perform key roles in commercial applications by biotic or abiotic function, such as preservatives, flavor modifiers, and cholesterol-lowering agents. For example, the sweetness of licorice roots is attributable to the presence of the triterpenoid saponin glycyrrhizin. Evidence of the health benefits of saponins is increasing, and their expanding application in food, cosmetics, and pharmaceutical industries is attracting commercial attention [2]. In the aspect of biochemical reaction, 2,3-Oxidosqualene is the common precursor of all the sterols and triterpenoids. The cyclization processes of 2,3-Oxidosqualene including protonation, cyclization, rearrangement, and deprotonation produce the diversiform skeletons of sterols and triterpenes [3–5]. During the cyclization processes, “chair-chair-chair (C-C-C)” and “chair-boat-chair (C-B-C)” configurations are present prior to cyclization into diverse triterpenes. Sterols are mainly formed through the “C-B-C” configuration, and triterpenoids are mainly formed through the “C-C-C” configuration [1].

Oxidosqualene cyclases (OSCs) are the pivotal enzyme in promoting the diversification of triterpenes, which can catalyze cyclization processes. Phillips et al., (2006) [6] divided plant OSCs into two groups based on the nature of their presumed catalytic intermediates, that is, protosteryl and dammarenyl cations, which possess distinct stereochemistry and ring configurations [6,7]. The protosteryl cation adopts the C-B-C configuration and is an intermediate leading to the formation of cycloartenol, lanosterol, parkeol, and cucurbitadienol tetracyclic triterpene skeletons (6-6-6-5). Isoarborinol is an unusual pentacyclic triterpene (6-6-6-6-5) derived from the additional D-ring expansion of the protosteryl cation

based on the C-B-C configuration. Most pentacyclic triterpene skeletons, however, are derived from the dammarenyl cation by D-ring expansion to form lupeol or further E-ring expansion to form β -amyrin via the C-C-C conformation [1]. More than 150 OSC genes have been identified in plants and their catalytic function were mostly identified by heterologous expression in yeast. The *Arabidopsis thaliana* genome contains 13 OSC genes, including cycloartenol synthase (*AtCAS1*), lanosterol synthase (*AtLSS1*), and β -amyrin synthase (*LUP4*), and several OSCs with mixed functions [8–12]. The amyirin synthase of *Eriobotrya japonica* (*EjAS*) can catalyze the conversion of 2,3-oxidosqualene into α - and β -amyrin at a ratio of 17:3 [13]. In OSCs of *Malus domestica* (*MdOSC*s), *MdOSC1* and *MdOSC3* can produce α -amyrin as the main product, with β -amyrin and lupeol as the minor products. *MdOSC4* is characterized as germanicol synthase, and *MdOSC5* produces lupeol as a major product [13,14]. Further, the potential of plant OSCs has been exploited in engineered yeast. *MdOSC1* introduced *Saccharomyces cerevisiae* achieve 5.8 times the maximum α -amyrin production [14].

Rosa rugosa is a shrub of Rosaceae and native to North China, Northeast China, Japan, Korea, and Russia, and has a long history of application as a medicinal plant. Its petals and roots have been widely used to staunch bleeding [15], reduce blood lipids, treat diabetes, and as a cure for cancer or cardiovascular diseases [16,17]. In the *R. rugosa* root extract, the active triterpenoids Kaji ichigoside F1, rosamultin, euscaphic acid, and tormentic acid showed anti-injury and anti-inflammatory effects on animal models; among these compounds, rosamultin had a strong inhibitory effect on human immunodeficiency virus-1 protease [18,19]. These reports point to the key role that terpenoids play in the pharmacological activities of *R. rugosa*. However, the OSC genes of *R. rugosa* (*RrOSC*s) involved in the synthesis of pharmacological terpenoids (e.g., rosamultin) haven't been reported. This study was aimed at discovering the potential key *RrOSC*s in the active terpenoids synthesis pathway of *R. rugosa* by systematic analysis of the *RrOSC* family.

2. Results

2.1. Identification and Phylogenetic Analysis of *RrOSC*s

A total of 18 *RrOSC* genes were identified, and the predicted proteins ranged from 540 to 770 amino acids in length and from 62.283 to 85.712 kDa in molecular weight (Table S1). All OSCs (except *RrOSC4*, which is extremely short and could belong to a pseudogene or unknown coding gene) were divided into 6 Groups (GI–GVI), which were named by common catalytic function based on the phylogenetic branches (Figure 1). OSCs in GI would produce mixed products and the main components of the mixed products were different; OSCs of GIII including *RrOSC9* and *RrOSC13* would be cycloartenol synthases. OSCs of GIV including *RrOSC12* would be lupeol synthases. OSCs of GV including *RrOSC14*, *RrOSC 15*, and *RrOSC 16* would be lanosterol synthases. No *RrOSC*s belonged to GII (β -amyrin synthases) or GVI (specific to monocotyledons).

2.2. The Conserved Motifs and Gene Structures of *RrOSC*s

The exon number of *RrOSC*s was in the range of 14–21. The SQHop_cyclase_C and/or SQHop_cyclase_N domains located by constant exon number (18–20 exons) indicated gene structures of *RrOSC*s were conserved (Figure 2b). Analysis of motifs (Figure 2a) showed that the domains of most *RrOSC*s composed of motifs 3-8-5-10-2 (corresponding to SQHop_cyclase_C) and motifs 6-7-4-1 (corresponding to SQHop_cyclase_N) were highly aligned. Interestingly, Group I members lack one motif5 in the N-terminal compared while other *RrOSC*s comprised of two motif5. And motif6 and/or motif7 were absent in SQHop_cyclase_N of Group I members *RrOSC1* and *RrOSC3*.

2.3. Clusters and Gene Duplications of *RrOSC*s

The *RrOSC*s are distributed on three (Chr2, Chr4, and Chr7) of all seven chromosomes (Figure 3a). Most members are located on the gene clusters of Chr2 and Chr4 while only *RrOSC18* is distributed on the Chr7. According to the gene duplication type (Table S3),

Figure 1. Phylogenetic analysis of 95 *oxidosqualene cyclases* (OSCs) of different plant species. The phylogenetic dendrogram was generated using the neighbor-joining method with 1000 bootstrap replicates (numbers in branches). Groups were distinguished with different colors and OSCs of *Rosa rugosa* (RrOSCs) were highlighted. The information on plant species and corresponding OSCs are listed in Table S2.

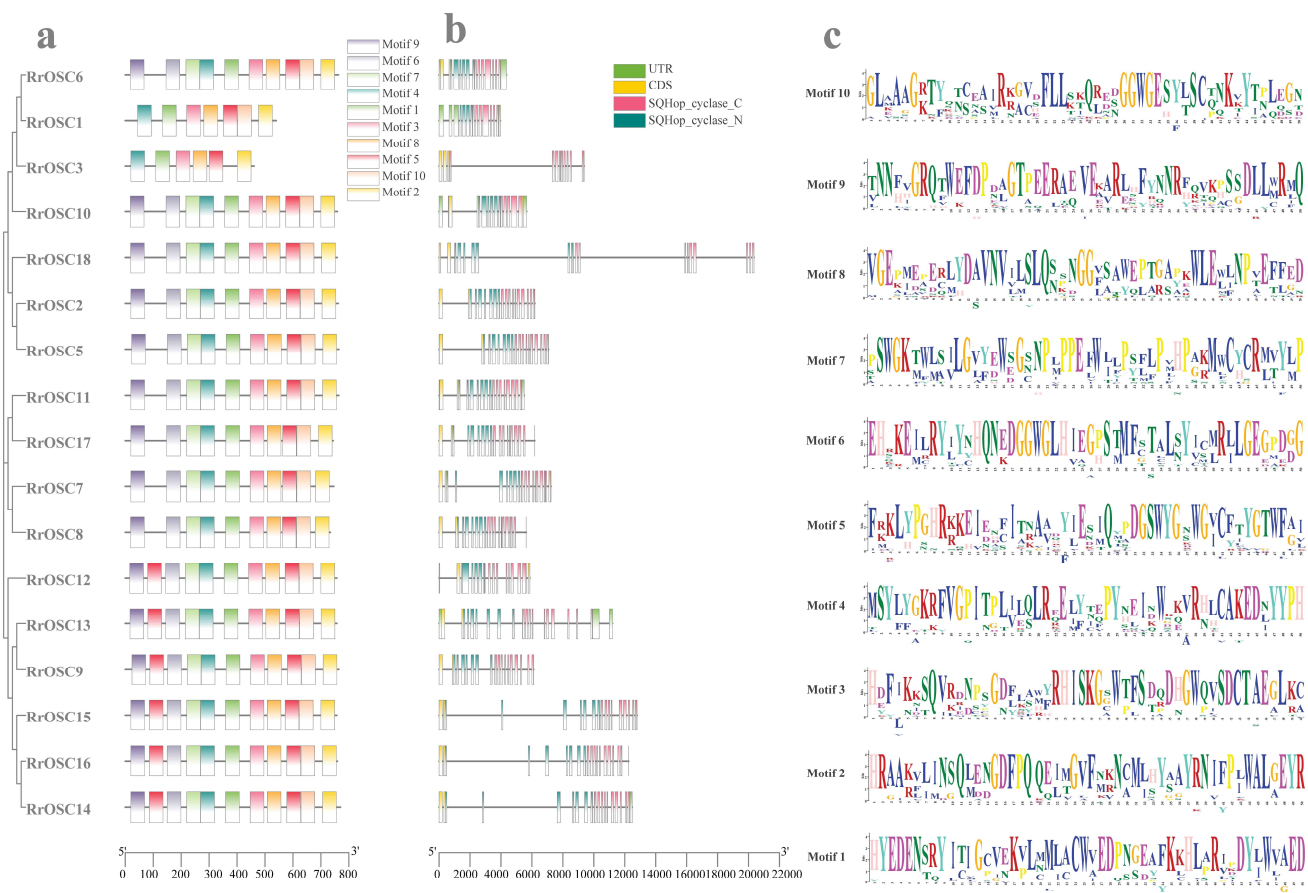


Figure 2. Motifs (a), gene structures (b), and motif sequences (c) of *RrOSCs*. The coloring boxes represent motifs whose locations were labeled by the scale plate of amino acid residue (a). Exons (rectangles) separated by introns (lines) were colored yellow (coding sequence) and green (untranslated region), and the locations were labeled by the nucleotide scale plate. Domains were colored purple (SQHop_cyclase_C) and blue (SQHop_cyclase_N) (b). The top 10 significantly enriched motifs are listed as sequence logos (c).

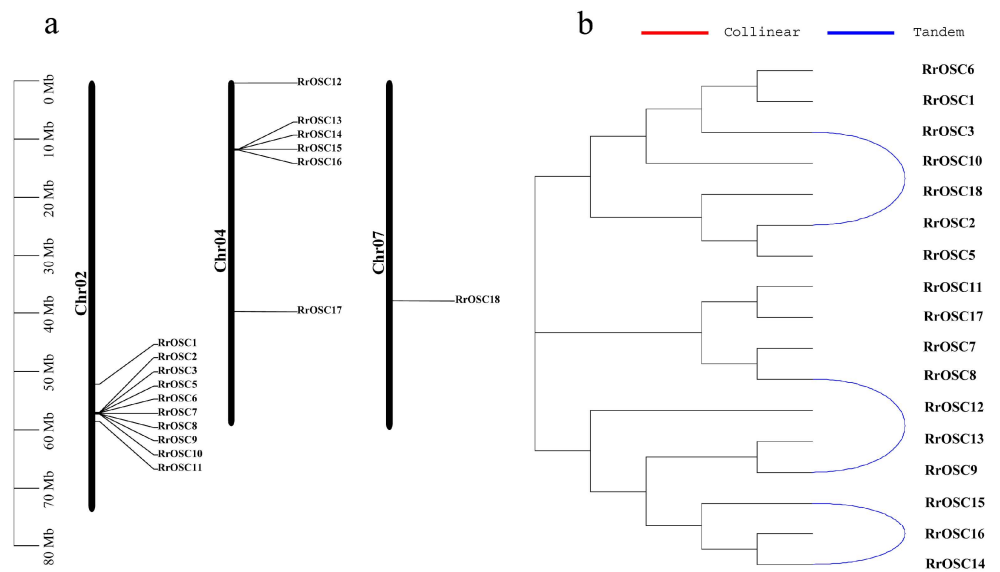


Figure 3. Chromosomal localization (a) and tandem duplication (b) analysis of *RrOSCs*. Blue lines indicated the duplication gene pairs according to the synteny analysis.

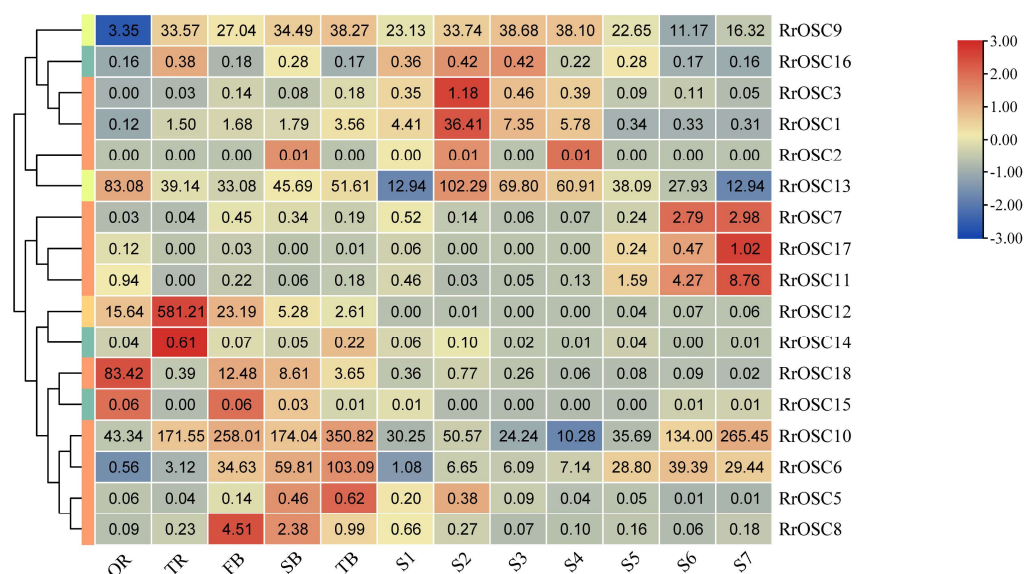


Figure 4. Digital gene expression profiles in the roots, lateral branches, and flowers of *R. rugosa* “Zi Zhi”. A heat map was generated based on the normalized Log2 FPKM represented by the blue-yellow gradation. The numbers in the heat map were FPKM from RNA-seq data. The OR and TR columns represent 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 highly-expressed *RrOSC6*, *RrOSC10*, and *RrOSC13* in the lateral branches of *R. rugosa* showed two expression patterns. The expression of *RrOSC6* and *RrOSC13* increased from the first lateral branch to the third lateral branch. The expression of *RrOSC10* was the highest, decreased from the primary lateral branch to the secondary lateral branch, and increased from the secondary lateral branch to the tertiary lateral branch. The three genes all showed the highest abundance in the third lateral branch, and only *RrOSC6* specific expressed in branches.

In roots (Figure 4, TIR or OR), the expression levels of *RrOSC10*, *RrOSC12*, *RrOSC13*, and *RrOSC18* were high (FPKM > 50), and *RrOSC12* was strongly expressed in TIR but almost not at all in OR. In flowers, the expression levels of *RrOSC9*, *RrOSC10*, and *RrOSC13* were higher. The expression levels of *RrOSC9* and *RrOSC13* were high in the early stage, decreased at the end of the flowering stage, and reached a maximum at S2 and S3. The expression of *RrOSC10* was the highest, and it increased significantly at the end of the flowering stage and reached the maximum at S7. In the lateral branches and petals, more than half of the genes were not expressed or had low expression.

2.5. Functional Characterization of *RrOSC12*

To elucidate the function of *RrOSC12*, we detected its products by expression in yeast. Authentic lupeol showed a dominant peak at 17.8 min (Figure 5b). A peak with the corresponding retention time was observed in the total ion chromatogram of *RrOSC12* expressed yeast, and mass spectra confirmed that it was that of lupeol (Figure 5c). These results demonstrate that *RrOSC12* catalyzes 2,3-oxidosqualene to yield lupeol.

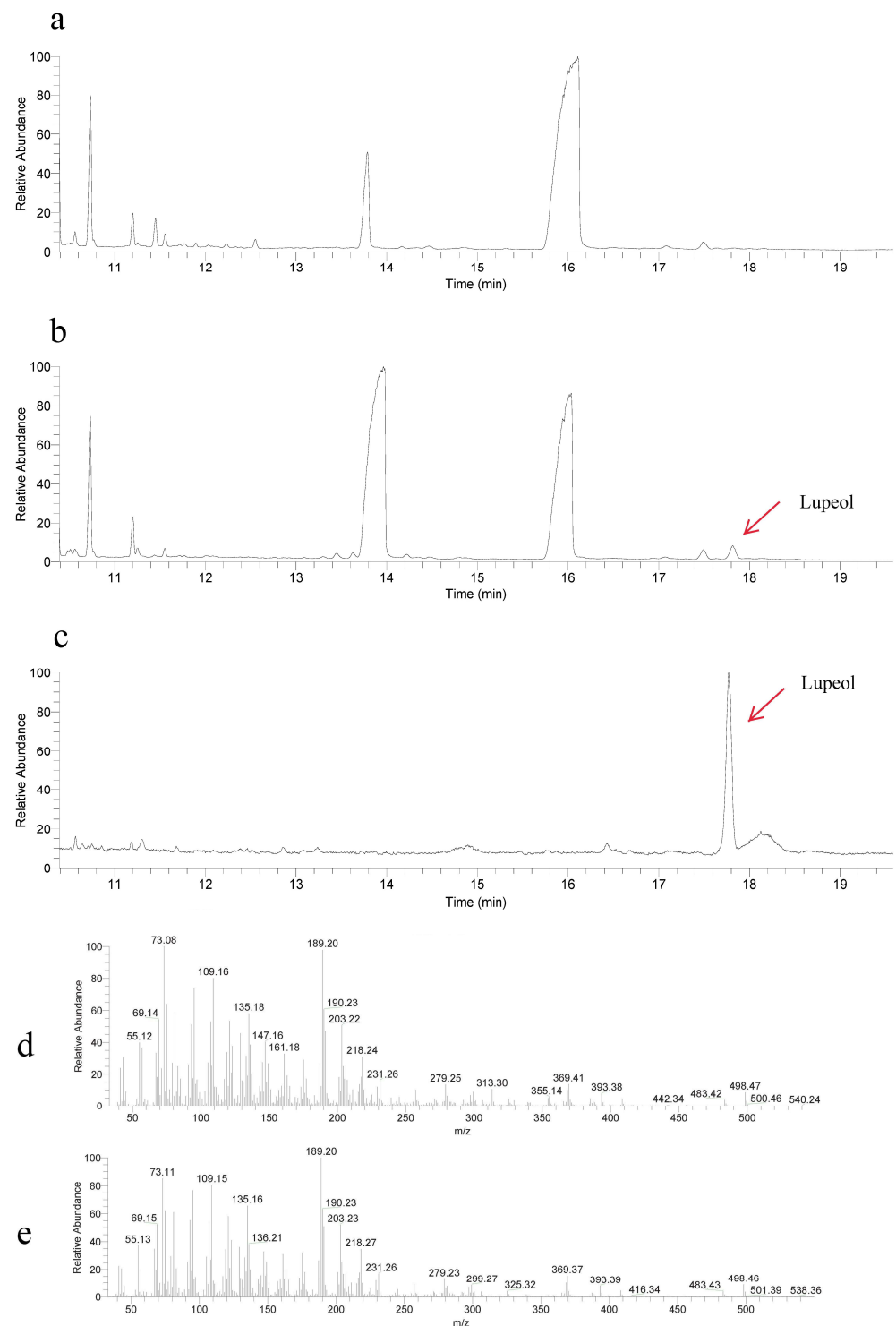


Figure 5. Gas chromatography–mass spectrometry analysis of yeast extracts overexpressing the *R. rugosa* lupeol synthase gene. Chromatogram for the yeast with empty vector (a), transformed yeast with the *RrOSC12* gene (b), and the authentic lupeol standard (c). Mass spectra in a retention time of 17.8 min for lupeol in the transformed yeast (d) and the authentic lupeol standard (e).

3. Discussion

The OSC gene number (18 members) of *R. rugosa* was more than the 14 members found in *Arabidopsis* and 11 members in rice. A comparison of gene structure and motifs showed that the gene structure difference of *RrOSCs* in the same subgroup didn't affect their highly-conserved domains including consistent motif structure. Additionally, many *RrOSCs* in the

same subgroup shared more similar expression patterns, such as two different subgroups of GI: *RrOSC7*, *RrOSC11*, and *RrOSC17* all highly expressed in flowers or *RrOSC6*, *RrOSC10*, and *RrOSC18* all highly expressed in roots (or branches).

The OSCs in the same subgroup in GI could produce mix-products with accordant main components. *RrOSC1*-*RrOSC6*, *RrOSC10*, and *RrOSC18* were clustered with *MdOSC3*, *EjAS*, and *MdOSC1* of Rosaceae plants as a subgroup. *MdOSC1* yields α -amyrin, β -amyrin, and lupeol at a ratio of 85:13:2 [13], *EjAS* produces α -amyrin and β -amyrin at a ratio of 17:3, and *MdOSC3* manufactures α -amyrin, β -amyrin, and lupeol at a ratio of 85:14:1 [20]. This subgroup was most likely a cluster of α -amyrin synthases. In addition, the tandem replication pair *RrOSC2*-*RrOSC3* and proximal duplication genes *RrOSC6*, and *RrOSC10* indicated that gene duplications contributed to the functional redundancy of the seven α -amyrin synthase candidates. *RrOSC7*, *RrOSC8*, *RrOSC11*, and *RrOSC17* combined with *MdOSC4* and *MdOSC5* were clustered into another subgroup. The main product of *MdOSC4* is germanicol, that of *MdOSC5* is lupeol, and both contain the by-product β -amyrin. Thus, the four candidates are unlikely to be α -amyrin synthases. According to the structural comparison, the rosamultin of *R. rugosa* should be formed from α -amyrin through cytochrome P450 oxidation to introduce functional groups and finally glycosylated by glycosyltransferase. Except for the four low abundance genes, above, *RrOSC*s, *RrOSC1*, *RrOSC10*, and *RrOSC18* were considered as the candidate genes of the rosamultin synthesis pathway.

Lupeol is an important pentacyclic triterpenoid. The applications of lupeol derivatives in the field of medicine are extensively studied because of their pharmacological activities, including antinephrotoxicity, antihepatotoxicity, and anticancer, anti-inflammatory, anti-heart disease, antiarthritis, and antidiabetes properties [21]. Lupeol was also found to be involved in the formation of epicuticular wax crystals on the stem and hypocotyl surfaces [22]. Only OSC of *R. rugosa* (*RrOSC12*) belongs to GIV, whose members would catalytic substrate to produce lupeol. The yeast expression proved its catalytic activity of lupeol. Though the lupeol synthase activity in vivo of *RrOSC12* needs to be studied further, its remarkably high mRNA abundance in TR and the lupenone and betulinic acid (downstream products of lupeol) accumulation in TR (unpublished metabolome data) indicated that lupeol contributes partly to the medicinal value of *R. rugosa* roots. In addition, the chromosome terminal locations (near telomere) of *RrOSC12* provide insight into DNA methylation, which could be the entry point for further exploring the pathway gene cluster to jointly modify the substrate (i.e., the gene cluster of *RrOSC12* and corresponding downstream P450 genes).

The enormous metabolic engineering potential of OSCs is increasingly apparent, e.g., *MdOSC1* [14,23], enzymes cycloartenol synthase gene of *Panax ginseng* and β -amyrin synthase genes of *Glycyrrhiza glabra*, and *P. ginseng* [24,25]. Based on the genetic transformation method in hairy roots of *R. rugosa* that has been established by our own lab, we would like to overexpress *RrOSC12* in *R. rugosa* hairy roots to increase lupeol production. In addition, the two α -amyrin synthase candidates *RrOSC10/18*, which were highly expressed in roots, should be tested further for increasing the potential of α -amyrin and that of the downstream product rosamultin.

4. Materials and Methods

4.1. Plant Materials

The three-year-old cutting seedlings of *R. rugosa* “Zi Zhi” were planted in germplasm in the resource nursery of Yangzhou University (32.391° N, 119.419° E), Yangzhou, China. Samples including roots (OR), primary lateral branch (PB), secondary lateral branch (SB), tertiary lateral branch (TB), flowers of seven stages (S1–S7), and roots (TR) of tissue culture seedlings (two-month-old micro-cutting seedlings) were collected and frozen using liquid nitrogen, and then stored at -80°C for RNA extraction (Figure 6).



Figure 6. The lateral branches (a), primary lateral branch (1), secondary lateral branch (2), tertiary lateral branch (3). Roots (b) and roots of tissue culture seedlings (c) (4). Petals at different development stages (d), S1: Large bud stage, S2: Reddish stage, S3: Flowering initiation stage, S4: Flower bud half-opening stage, S5: Initial opening stage, S6: Semi-opening stage, S7: Full opening stage.

4.2. Identification of RrOSCs

RrOSCs were predicted by HMMER 3.0 from the proteome of *R. rugosa* “Zi Zhi” (unpublished) using the hidden Markov model of the squalene-hopene cyclase N-terminal domain (PF13249) and squalene-hopene cyclase C-terminal domain (PF13243) [26]. The candidate genes with both domains were confirmed by Pfam (<http://pfam.xfam.org/>, accessed on 1 September 2022) and the Conserved Domain Database [27,28]. The dubious OSC genes were checked by manual inspection. The isoelectric point and molecular weight of RrOSC proteins were obtained by ProtParam tool of Extasy [29].

4.3. Phylogenetic Analyses

RrOSC proteins and 78 OSC proteins of different plant species were aligned using the MAFFT algorithm (<https://mafft.cbrc.jp/alignment/server/index.html>, accessed on 1 September 2022) [30]. The phylogenetic dendrogram was constructed by the neighbor-joining method of MEGA-X [31], with 1000 bootstrap replications. In addition, the p-distance model of the substitution type, pairwise deletion of Gaps/Missing data treatment, and uniform rates of rates among sites were selected for phylogenetic analysis. The 78 OSC genes of *Arabidopsis thaliana*, *Lotus japonicus*, *Panax ginseng*, *Abies magnifica*, *Adiantum capillus-veneris*, *Avena strigosa*, *Betula platyphylla*, *Costus speciosus*, *Cucurbita pepo*, *Dioscorea zingiberensis*, *Glycyrrhiza glabra*, *Kalanchoe daigremontiana*, *Kandelia candel*, *Luffa cylindrica*, *Oryza sativa*, *Pisum sativum*, *Polypodiodes niponica*, *Rhizophora stylosa*, *Ricinus communis*, *Artemisia annua*, *Aster sedifolius*, *Bruguiera gymnorrhiza*, *Euphorbia tirucalli*, *Medicago truncatula*, *Nigella sativa*, *Polygala tenuifolia*, *Solanum lycopersicum*, *Vaccaria hispanica*, *Olea europaea*,

Taraxacum officinale, *Aster tataricus*, *Stevia rebaudiana*, *Malus pumila*, and *Eriobotrya japonica* are shown in Table S2.

4.4. Gene Structures and Conserved Motifs Analysis

Conserved motifs of RrOSC proteins were predicted by the MEME suite (<http://meme-suite.org/>, accessed on 1 September 2022) with the default parameters and a maximum motif number of 8 [32]. The conserved motifs and gene structures with domain location of RrOSCs were illustrated by the gene structure view tool of TBtools toolkit 1.086 [33] based on the genomic DNA and mRNA sequences of the RrOSCs (Tables S5 and S6).

4.5. Chromosome Location and Synteny Analysis

According to the GFF3 file of the reference genome, RrOSC gene location was illustrated by the Gene location visualize function of the GTF/GFF tool of TBtools toolkit 1.086 [33].

Inter-species synteny analysis was conducted by reciprocal BLASTP search for potential homologous gene pairs ($E < 10^{-5}$, top five matches) of the whole genome, and the syntenic regions were predicted by the MCScanX tool of TBtools toolkit 1.086 [34]. The gene duplication events were predicted according to syntenic regions by the 'duplicate_gene_classifier' program of MCScanX [34].

4.6. Expression Analysis and Functional Characterization

The RNA of the above 12 tissues (R, TR, PB, SB, TB, and S1–S7) were extracted by the MiniBEST Universal RNA Extraction Kit (TaKaRa, Kusatsu, Japan) for transcriptome libraries building and sequencing on the Illumina NovaSeq 6000 platform. The raw reads could be accessed in the SRA of NCBI (PRJNA725330) and BIG Sub of China National Center for Bioinformation (PRJCA012932). After low-quality reads filtering, clean reads were used for analysis. Based on clean reads, read counts of each gene were calculated by mapping to the genome of *R. rugosa* "Zi Zhi" using HISAT 2.2.1 (<http://daehwankimlab.github.io/hisat2/>, accessed on 1 September 2022). The sequencing quality was examined by Q30 value, percentage of clean reads, and alignment coverage (Table S7). Fragments per kilobase of exon model per million reads mapped (FPKM) of RrOSCs were calculated by normalization of read counts using a self-built R script.

RrOSC12 was cloned by the polymerase chain reaction (PCR). The coding sequence of RrOSC12 was sub-cloned to pESC-leu and transformed to yeast strain AM94. Culture extracts of the RrOSC12 expressed yeast were derivatized with trimethylsilylating agents and submitted to gas chromatography–mass spectrometry analysis.

Supplementary Materials: The supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms232213884/s1>.

Author Contributions: Conceptualization, L.F. and J.W.; data curation, P.W., M.X. and Y.C.; writing—Original draft, P.W. and J.W.; writing—Review & editing, J.W. All authors have read and agreed to the published version of the manuscript.

Funding: The study was funded by the National Natural Science Foundation of China (grant numbers 31972454, 32171861, 32002076), the National Key R&D Program of China (grant number 2018YFD1000400), the Jiangsu Provincial Natural Science Foundation (grant number SBK2020043530), the Jiangsu Provincial Natural Science Research Project of Universities (grant number 20KJB210004), and the Jiangsu Provincial Agricultural Science and Technology Independent Innovation Fund (grant numbers CX(20)3023, CX(20)3026), and supported by the high-end talent support program of Yangzhou University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Xu, R.; Fazio, G.C.; Matsuda, S.P. On the origins of triterpenoid skeletal diversity. *Phytochemistry* **2004**, *65*, 261–291. [\[CrossRef\]](#) [\[PubMed\]](#)
- Osbourne, A.; Goss, R.J.; Field, R.A. The saponins: Polar isoprenoids with important and diverse biological activities. *Nat. Prod. Rep.* **2011**, *28*, 1261–1268. [\[CrossRef\]](#) [\[PubMed\]](#)
- Haralampidis, K.; Trojanowska, M.; Osbourne, A.E. Biosynthesis of triterpenoid saponins in plants. *Adv. Biochem. Eng. Biotechnol.* **2002**, *75*, 31–49. [\[CrossRef\]](#) [\[PubMed\]](#)
- Shibuya, M.; Zhang, H.; Endo, A.; Shishikura, K.; Kushiro, T.; Ebizuka, Y. Two branches of the lupeol synthase gene in the molecular evolution of plant oxidosqualene cyclases. *Eur. J. Biochem.* **1999**, *266*, 302–307. [\[CrossRef\]](#)
- Basyuni, M.; Oku, H.; Tsujimoto, E.; Kinjo, K.; Baba, S.; Takara, K. Triterpene synthases from the Okinawan mangrove tribe, Rhizophoraceae. *FEBS J.* **2007**, *274*, 5028–5042. [\[CrossRef\]](#) [\[PubMed\]](#)
- Phillips, D.R.; Rasbery, J.M.; Bartel, B.; Matsuda, S.P.T. Biosynthetic diversity in plant triterpene cyclization. *Curr. Opin. Plant Biol.* **2006**, *9*, 305–314. [\[CrossRef\]](#) [\[PubMed\]](#)
- Xue, Z.; Duan, L.; Liu, D.; Guo, J.; Ge, S.; Dicks, J.; ÓMáille, P.; Osbourne, A.; Qi, X. Divergent evolution of oxidosqualene cyclases in plants. *New Phytol.* **2012**, *193*, 1022–1038. [\[CrossRef\]](#) [\[PubMed\]](#)
- Xiang, T.; Shibuya, M.; Katsube, Y.; Tsutsumi, T.; Otsuka, M.; Zhang, H.; Masuda, K.; Ebizuka, Y. A new triterpene synthase from *Arabidopsis thaliana* produces a tricyclic triterpene with two hydroxyl groups. *Org. Lett.* **2006**, *8*, 2835–2838. [\[CrossRef\]](#)
- Kushiro, T.; Shibuya, M.; Masuda, K.; Ebizuka, Y. A novel multifunctional triterpene synthase from *Arabidopsis thaliana*. *Tetrahedron Lett.* **2000**, *41*, 7705–7710. [\[CrossRef\]](#)
- Lodeiro, S.; Xiong, Q.; Wilson, W.K.; Kolesnikova, M.D.; Onak, C.S.; Matsuda, S.P.T. An oxidosqualene cyclase makes numerous products by diverse mechanisms: A challenge to prevailing concepts of triterpene biosynthesis. *J. Am. Chem. Soc.* **2007**, *129*, 11213–11222.
- Xiong, Q.; Wilson, W.K.; Matsuda, S.P. An *Arabidopsis* oxidosqualene cyclase catalyzes iridal skeleton formation by Grob fragmentation. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 1285–1288. [\[CrossRef\]](#) [\[PubMed\]](#)
- Segura, M.J.R.; Meyer, M.M.; Matsuda, S.P.T. *Arabidopsis thaliana* LUP1 converts oxidosqualene to multiple triterpene alcohols and a triterpene diol. *Org. Lett.* **2000**, *2*, 2257–2259. [\[CrossRef\]](#) [\[PubMed\]](#)
- Andre, C.M.; Legay, S.; Deleruelle, A.; Nieuwenhuizen, N.; Punter, M.; Brendolise, C.; Cooney, J.M.; Lateur, M.; Hausman, J.F.; Larondelle, Y.; et al. Multifunctional oxidosqualene cyclases and cytochrome P450 involved in the biosynthesis of apple fruit triterpenic acids. *New Phytol.* **2016**, *211*, 1279–1294. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yu, Y.; Chang, P.; Yu, H.; Ren, H.; Hong, D.; Li, Z.; Wang, Y.; Song, H.; Huo, Y.; Li, C. Productive Amyrin Synthases for Efficient α -Amyrin Synthesis in Engineered *Saccharomyces cerevisiae*. *ACS Synth. Biol.* **2018**, *7*, 2391–2402. [\[CrossRef\]](#) [\[PubMed\]](#)
- Hashidoko, Y. The phytochemistry of *Rosa rugosa*. *Phytochemistry* **1996**, *43*, 535–549. [\[CrossRef\]](#)
- Lee, S.Y.; Kim, J.D.; Lee, Y.H.; Rhee, H.; Choi, Y.S. Influence of extract of *Rosa rugosa* roots on lipid levels in serum and liver of rats. *Life Sci.* **1991**, *49*, 947–951. [\[CrossRef\]](#)
- Na, J.R.; Oh, D.R.; Han, S.; Kim, Y.J.; Choi, E.; Bae, D.; Oh, D.H.; Lee, Y.H.; Kim, S.; Jun, W. Antistress Effects of *Rosa rugosa* Thunb. on Total Sleep Deprivation-Induced Anxiety-Like Behavior and Cognitive Dysfunction in Rat: Possible Mechanism of Action of 5-HT₆ Receptor Antagonist. *J. Med. Food* **2016**, *19*, 870–881. [\[CrossRef\]](#)
- Jung, H.-J.; Nam, J.-H.; Choi, J.; Lee, K.-T.; Park, H.-J. 19 α -hydroxyursane-type triterpenoids: Antinociceptive anti-inflammatory principles of the roots of *Rosa rugosa*. *Biol. Pharm. Bull.* **2005**, *28*, 101–104. [\[CrossRef\]](#)
- Park, J.C.; Kim, S.C.; Choi, M.R.; Song, S.H.; Yoo, E.J.; Kim, S.H.; Miyashiro, H.; Hattori, M. Anti-HIV protease activity from *rosa* family plant extracts and rosamultin from *Rosa rugosa*. *J. Med. Food* **2005**, *8*, 107–109. [\[CrossRef\]](#)
- Liu, X.; Su, M.; Wu, S.; Li, H. Clone of Amyrin Synthase Gene Conservative District from *Eriobotrya japonica* ‘Jie Fang Zhong’. *Subtrop. Plant Sci.* **2013**, *42*, 1–4. (In Chinese)
- Liu, K.; Zhang, X.; Xie, L.; Deng, M.; Chen, H.; Song, J.; Long, J.; Li, X.; Luo, J. Lupeol and its derivatives as anticancer and anti-inflammatory agents: Molecular mechanisms and therapeutic efficacy. *Pharm. Res.* **2021**, *164*, 105373. [\[CrossRef\]](#) [\[PubMed\]](#)
- Guhling, O.; Hobl, B.; Yeats, T.; Jetter, R. Cloning and characterization of a lupeol synthase involved in the synthesis of epicuticular wax crystals on stem and hypocotyl surfaces of *Ricinus communis*. *Arch. Biochem. Biophys.* **2006**, *448*, 60–72. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yu, Y.; Rasool, A.; Liu, H.; Lv, B.; Chang, P.; Song, H.; Wang, Y.; Li, C. Engineering *Saccharomyces Cerevisiae* for High Yield Production of A-Amyrin Via Synergistic Remodeling of A-Amyrin Synthase and Expanding the Storage Pool. *Metab. Eng.* **2020**, *62*, 72–83. [\[CrossRef\]](#) [\[PubMed\]](#)
- Liang, Y.; Zhao, S.; Zhang, X. Antisense Suppression of Cycloartenol Synthase Results in Elevated Ginsenoside Levels in Panax Ginseng Hairy Roots. *Plant Mol. Biol. Rep.* **2009**, *27*, 298–304. [\[CrossRef\]](#)
- Dai, Z.; Wang, B.; Liu, Y.; Shi, M.; Wang, D.; Zhang, X.; Liu, T.; Huang, L.; Zhang, X. Producing Aglycons of Ginsenosides in Bakers’ Yeast. *Sci. Rep.* **2014**, *4*, 3698. [\[CrossRef\]](#)

26. Mistry, J.; Finn, R.D.; Eddy, S.R.; Bateman, A.; Punta, M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* **2013**, *41*, e121. [[CrossRef](#)]
27. Finn, R.D.; Bateman, A.; Clements, J.; Coghill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; et al. Pfam: The protein families database. *Nucleic Acids Res.* **2014**, *42*, D222–D230. [[CrossRef](#)]
28. Finn, R.D.; Coghill, P.; Eberhardt, R.Y.; Eddy, S.R.; Mistry, J.; Mitchell, A.L.; Potter, S.C.; Punta, M.; Qureshi, M.; Sangrador-Vegas, A.; et al. The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* **2016**, *44*, D279–D285. [[CrossRef](#)]
29. Gasteiger, E.; Gattiker, A.; C. Hoogland, C.; Ivanyi, I.; Appel, R.D.; Bairoch, A. ExPASy: The Proteomics Server for in-Depth Protein Knowledge and Analysis. *Nucleic Acids Res.* **2003**, *31*, 3784–3788. [[CrossRef](#)]
30. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
31. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [[CrossRef](#)] [[PubMed](#)]
32. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME SUITE: Tools for motif discovery and searching. *Nucleic Acids Res.* **2009**, *37*, W202–W208. [[CrossRef](#)] [[PubMed](#)]
33. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202. [[CrossRef](#)]
34. Wang, Y.; Tang, H.; Debarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.H.; Jin, H.; Marler, B.; Guo, H.; et al. MCScanX: A toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **2012**, *40*, e49. [[CrossRef](#)] [[PubMed](#)]