OPEN ACCESS

GAVIN PUBLISHERS

Molecular Biology: Open Access

Tu D, et al. Mol Bio Open Acc: MBOP-103.

Research Article

DOI: 10.29011/MBOP-103. 00003

Cloning and Expression of Sgcyp749a22 from *Siraitia Grosvenorii*

DongpingTu[1,2,3], Xiaojun Ma[4*], Liuping Wang[1], Lichun Zhao[1,2], Huan Zhao[4], Changming Mo[5], Qi Tang[6*]

[1]Guangxi University of Chinese Medicine, China
[2]Guangxi Engineering Technology Research Center of Zhuang and Yao Yao, China
[3]Collaborative Innovation Center of Research on Functional Ingredients from Agricultural Residues of Guangxi University of Chinese medicine, China
[4]Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, China
[5]Guangxi Academy of Agricultural Sciences, China
[6]Horticulture & Landscape College, Hunan Agriculture University, China

[*]**Corresponding author:** Xiaojun Ma, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, China. Tel: +86-13501187416; Email: mayixuan10@163.com; Qi Tang, Horticulture & Landscape College, Hunan Agriculture University, China. Tel: +86-15873157602; Email: tangqi423@sina.com

Citation: Tu D, Ma X, Wang L, Zhao L, Zhao H, et al. (2018) Cloning and Expression of Sgcyp749a22 from *Siraitia Grosvenorii*. Mol Bio Open Acc: MBOP-103. DOI: 10.29011/MBOP-103. 00003

Received Date: 30 January, 2018; **Accepted Date:** 22 February, 2018; **Published Date:** 07 March, 2018.

Abstract

Siraitia grosvenorii have been used for thousands of years in China as a natural sweetener and as a folk medicine. Its active ingredient Mogroside V is a new sweetener with the therapeutic function, has been drawing more and more people's attention. CYP450 plays an essential role in the Mogrosides biosynthesis pathway, and SgCYP749A22 acts a candidate gene in the biosynthesis pathway from the preliminary study. However, little is known about the SgCYP749A22 gene in *S. grosvenorii*. Here, based on the transcriptome date, a full-length cDNA sequence of SgCYP749A22 was firstly cloned by RT-qPCR and rapid-amplification of cDNA ends (RACE) strategies. SgCYP749A22 is 2069 bp in length (GenBank accession No. AEM42984. 1) and contains a complete Open Reading Frame (ORF) of 1560 bp. The deduced protein was composed of 519 amino acids, molecular weight is 59.46 kDa, theoretical PI is 9.25, and was predicted to possess Cytochrome P450 domains. SgCYP749A22 gene was highly expressed in roots, diploid fruit and the fruits treated with hormones combining with pollination. On 10 days after treating with pollination, the expression of SgCYP749A22 had the higher level then decreased which was consist with the mogrosides biosynthesis of the fruits of *S. grosvenorii*. These results provided a scientific basis for the underlying mechanism of biosynthesis pathway of Mogrosides in *S. grosvenorii*.

Keywords: Biosynthesis; CYP450; Expression Pattern; Race Clone; *Siraitia grosvenorii*; SgCYP749A22

Introduction

S. grosvenorii is an important traditional Chinese medicine. Its fruits have many effects such as detoxification, preventing cough and eliminating sputum, cooling blood and soothing the stomach, clearing the lung and lubricating the intestine [1]. It is also the unique medicinal plant only distributing in the mountainous areas in the northern of Guangxi in China [2]. Its active ingredient, named Mogrosides V, is high sweetness, good effects of antitussive and expectorant, low in calories, non-toxic, good taste and broad prospects for development [3,4]. But the problems of the low content and high production cost of mogroside V restrict the development. The content of morgroside V was only 3 ‰ to 4 ‰ in the fresh fruits, but the price of the content of 50 % and 90 % morgroside V extracts were 610 $·kg^{-1}$ and 3050 $·kg^{-1}$, respectively. In recent years, the research on improving the content of morgroside V has become a hot research direction. The cultivation and breeding technology have been studied for decades, and the difficulty and potential of improving the mogroside V content are limited [5]. So we believe that the regulation of mogroside V synthesis key enzyme genes to solve the problem of low content of morgroside V, is a more direct and effective new way. This is an urgent need to carry out these genes basic research. It's reported that mogrol is the nucleus parent and its lack may be the main reason of the low content of mogroside V [6]. Because in the *S. grosvenorii* fruit growth development, the content of mogrol continued to decreasing, such as 0.028 % in 10 days, 0.025 % in 15 days, but only 0.0059 % in 30 days and almost reducing to zero in 40 days. Low glycosidase almost total conversed to high glycosides at the later stage of development, but at the same time

the content of glucose is still as high as 20 % [7,8]. Therefore, the regulation of the CYP450 gene, the biosynthesis enzyme of mogrol, may increase the supply of mogrol at the later stage of fruit development, thereby increase the content of mogroside V.

CYP450s have a wide range of biological activities, their general enzymatic function is to catalyze regiospecific and stereospecific oxidation of non-activated hydrocarbons, play the role of hydroxyl [9,10]. Such as CYP450-1 showed the activity of monooxygenase and played the hydroxylation role in the C-11 of Nanping E, while CYP450-2 played a positive role in the hydroxylation of C-7 and C-13 [11], meanwhile CYP450-2 also was shown recently to encode a GA 20-oxidase, converted GA14 to GA4 by removal of C-20 [12]. Entkaurene oxidase, as a member of CYP450, catalyzed the three oxidation steps between ent-kaurene and ent-kaurenoic acid [13], As showed in existing studies, CYP450 is the key enzyme in the synthesis of sweet glycosides and expression in different tissues of specific, especially higher expression level in the root and flower with more saponin synthesis more than other tissues significantly. Oat root saponin deficient mutant, Sad2, was a member of CYP450, can catalyze the β-amyrin to oat root saponin whose expression is limited to a root apical epidermis [14]. CYP88D6 (with the activity of β-amyrin -11- oxidase, expressed in the root and stolon), CYP93E3 (catalytic the hydroxylation of C-24 of β-amyrin) and CYP72A154 (catalytic 11-carbonyl-β-amyrin to synthesis glycyrrhizic acid, express in root, stem and stolon) were confirmed that they were the CYP450 all participate in the biosynthesis of licorice triterpenoid saponiN [15,16]. Carell et al. [17] identified the CYP716A12 taking part in the biosynthesis of hemolytic saponin by studying Medicago truncatula saponin, the defects of CYP716A12 can be generated soybean saponin but not synthetic hemolytic saponin. Dammarenediol was been hydroxylation by CYP450, and glycosylation was generated to Ginsenoside by glycosylation transfer enzyme in Panax ginseng at last, among of the CYP749A22 was proved to relate to the Ginsenoside synthesis [18].

CYP450 is a key enzyme in the biosynthesis pathway of Mogrolsides [6]. As showed in the pathway of mogrosides biosynthesis in (Figure 1), as a precursor for mogrosides synthesis, mogrol synthesis regulates the content of the end product of mogroside V, especially the enzyme regulation. CYP450 is the key enzyme catalyzing cucurbitadienol to mogrol, restrict the biosynthesis of mogrol, then mogrol was been glycosylation by glycosylation transfer enzyme to generate Mogrosides, that mogrol is the precursor of mogrosides biosynthesis and CYP450 play important role in this pathway. Meanwhile, the study of Xu [19,20] showed that the metabolites in vivo of mogrosides were mainly the products catalyzed by CYP450. In the preliminary transcriptome study, SgCYP749A22 take part in the biosynthesis pathway of mogrosides, which belong to CYP749A22. In our study full-length cDNA sequence of SgCYP749A22 from *S. grosvenorii* was isolated, and its expression patterns were investigated. The results of SgCYP749A22 function in the biosynthetic pathway may help to reveal the underlying mechanism of biosynthesis pathway of Mogrosides in *S. grosvenorii*, which lay a foundation for further exploring the biological functions finally to regulate the biosynthesis of mogrosides by regulating the gene and provide a scientific basis for transgenic breeding and an important target gene for genetic engineering for *S. grosvenorii*.

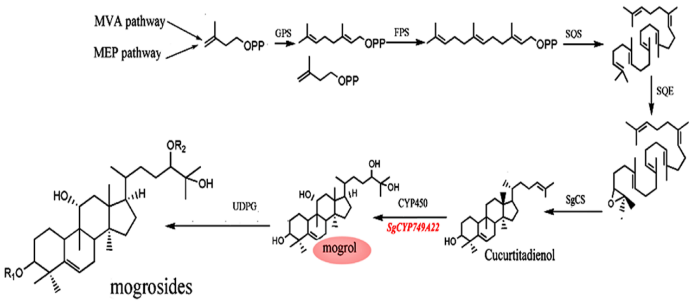

Figure 1: The biosynthesis pathway of mogrosides in *S. grosvenorii*.

Materials and Methods

Plant materials

S. grosvenorii cultivar 'Yongqing NO. 1' (diploid fruits with seeds) and the 'Yaoyuanwuzi NO. 1' (triploid fruits with seedless) were all planted in the experimental farm of Guangxi Medicinal Botanical Garden, Nanning, China. All the plants' samples used in the study were identified by Dr. Xiaojun Ma who breed the teo *S. grosvenorii* cultivars as the first person to complete. There were five sample groups, including A group (the different development period of triploid fruits: on the flowering day and 3 h, 1 day, 3 days, 5 days, 10 days and 15 days after treating with growth regulators of our lab' patent) , B group (the 10 days old seedless fruits with different treatment: pollination, growth regulators and the pollination combing with growth regulators) , C group (the different development period of diploid fruits: 10 days, 20 days, 30 days, 40 days, 50 days, 60 days and 70 days fruits), D group (the different tissues of the diploid plant: root, stem, leaf, flower, peel and pulp) , E group (the fruits of 40 days old of diploid, triploid and tripliod plants). All the samples were collected and stored in liquid nitrogen then stored at -80°C for tissue expression analysis.

RNA isolation and synthesis of SgCYP749A22 cDNA

Total RNA was extracted by the Trizol (Invitrogen, USA) method [6]. Quality and quantity of each total RNA sample were assessed in agarose gels (1%, w/v), spectrophotometricaly at 260 and 280nm (Bio-Rad, Nano Drop 2000), respectively. cDNA was synthesized by a reverse transcription kit (TaKaRa Prime Script™, China) according to the manufacturer's instructions. The cDNA was stored at -20°C for the later study.

Cloning of SgCYP749A22 by RACE Method

The primers used in the study are showed in (Table 1). The first-strand cDNA was used as the template for SgCYP749A22 core fragment amplication based on the unigenes of 454datA [6]. The PCR products were got by RACE kit (Clontech, USA) then subjected to electrophoresis on a 1% agarose gel for detection and purification. The amplified subjective fragments were cloned into the pGM-Tvector (Tiangen). Recombinant plasmids were transformed into Escherichia coli, selected by blue/white screening, and verified by PCR. Nucleotide sequencing was performed by Shanghai Sangon Biological Engineering Technology and Service Company, China.

Gene name	primers
5′RACEC YP749A22	CGGTGACCAACTGAATCGGGGAGTGAGA
3′RACEC YP749A22	TCCCCGATTCAGTTGGTCACCGTTTGCC
5′ORF	ATGGTTGGTATGAGGATATACGTTT
3′ORF	TCAAGGTTTGTGATGTGATAATGG
18sRNA	CTTCGGGATCGGAGTAATGA
	GCGGAGTCCTAGAAGCAACA
SgCYP 749A22	GATTCTACGGCGATATTCCTT
	AATGGATGAAGTATGACCTGAA

Table 1: Primers for gene cloning and RT-qPCR detection.

PCR of ORF of SgCYP749A22

The primers used in the study were showed in (Table 1). The PCR reaction contained 2 μL of cDNA, 2 μL of gene-specific primers, 25 μL of 2 × Trans Taq TM HiFi PCR Super Mix, 19 μL of double-distillated H_2O, in a final volume of 50 μL. The condition of PCR was as followed: 94˚C for 5 min, 94˚C for 40 secs, 50˚C for 40 secs, 72˚C for 1 min, 35 cycles, at last 72˚C for 10 min.

Isolation and Bioinformatics Analysis of SgCYP749A22

The Physical and chemical property of SgCYP749A22 protein were predicted by ExPASy Proteomics Server (http://www.Expasy.Ch/tools/protparam.html). The protein secondary structure and molecular modeling of SgCYP749A22 were predicted by SWISS-MODEL [21-24] (http://swissmodel.Expasy.org/). Then the signal peptide was predicted by SignalP 4.0 Server (http://www.Cbs.Dtu.dk/services/SignalP/) and Target 1.1 Server (http://www.Cbs.Dtu.dk/services/TargetP/), and the protein subcellular localization and the transmembrane region were predicted by SubLoc.v1. 0 [25] (http://www.bioinfo.tsinghua.edu.cn/SubLoc/) and TMHMM Server v.2.0 (http://www.cbs.Dtu.dk/services/TMHMM/). Conserved domains were analyzed by Blastp and ExPASy PROSITE (http://www.Expasy.ch/prosite/), finally to determine the relationship between CYP749A22 and other proteins, phylogenetic analysis was constructed for 11 proteins of different species using MEGA version 5.0 by the neighbor-joining method with 1000 bootstrap replicates.

RT-qPCR of SgCYP749A22

The RNA extracted from all the samples respectively, each sample was used for RT-qPCR. RT-qPCR was performed by ABI 7500 fluorescence quantitative PCR instrument. Each reaction contained 1 μL of cDNA, 1 μmol·L^{-1} of gene-specific primers, 0.5 μL of Rox, 10 μL of double-distillated H_2O and 12.5 μL of 2 × SYBR Green Master Mix (SYBR® Premix Ex TaqTMII (Tli RNaseH Plus), TaKaRa Prime Script™, China), in a final total volume of 25 μL. The 18S rRNA gene was used as the reference genE [26-28]. The primer sequences of genes in the study were showed in Table 1. RT-qPCR analysis condition was that: 95˚C for 10 min, followed by the 40 cycles of 95˚C for 30 secs, 95˚C for 5sec and 60˚C for 34sec. All experiments were used in triplicate for each sample and relative gene expression levels were calculated by the relative standard curve method. Each sample have three biological replicates, three techniques replicates and a negative control.

Results

Molecular cloning of full-length cDNAs and characterization of SgCYP749A22

A full-length cDNA clone was obtained using 5′-/3′-RACE extension methods. Two specific primers, one of 5′RACE CYP749A22 for 5′-RACE, and the other of 3′RACE CYP749A22 for 3′- RACE were designed (Table 1) to yield a 1630 bp 5′-cDNA ends sequence and a 461 bp 3′-cDNA ends sequence (Figure 2). The sequence analysis confirmed that the clone is a SgCYP749A22 gene. The full-length comprises 2069 bp. Its ORF is 1560 bp (Figure 3). The cloned cDNA has been submitted to GenBank under the accession number AEM42984. 1.

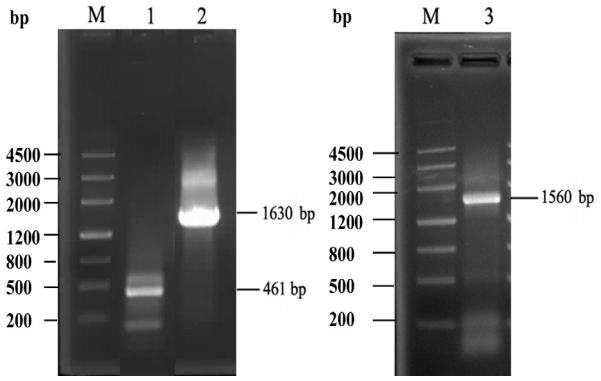

Figure 2: Amplification and detection the 5′-RACE, 3′-RACE end and ORF of SgCYP749A22. (M) D4500 DNA marker; (1) 5′-RACE product; (2) 3′-RACE product;(3) open reading frame (ORF) product.

```
130   acggttggtagaggatatacgtttcaagttttgtgttgggggggttagtggttttgggtgtgatgaagcttgtggtgagattatggatg
      M  V  G  M  R  I  Y  V  S  S  F  V  L  G  G  L  V  V  L  G  V  M  K  L  V  V  R  L  W  M
220   agccccacccgaacccagaagacgatgcggtcgcaggggattcgaggccctccttaccagttcattcaaggcaacaccaaagagatgtac
      S  P  I  R  I  Q  K  I  M  R  S  Q  G  I  R  G  P  P  Y  Q  F  I  Q  G  N  I  K  E  M  Y
310   accaaaagaacgcaagccaccgccacccccacggacctctcccataacattctccccagagttatgcctcatgttcactcctggctcaac
      T  K  R  M  Q  A  T  A  T  P  M  D  L  S  H  N  I  L  P  R  V  L  P  H  V  H  S  W  L  N
400   cattacgggaagaacttccttcagtggttaggtactgaagctcagttcgtcgttacagaccctgagatgatcaaggagatgctccatgtt
      H  Y  G  K  N  F  L  Q  W  L  G  T  E  A  Q  F  V  V  T  D  P  E  M  I  K  E  M  L  H  V
490   cgacaaaggaatttcccaaaagcaaaactccatggctacatacagaaaattttcggagatggtactgtcacagccgagggtgaaagacgg
      R  Q  R  N  F  P  K  A  K  L  H  G  Y  I  Q  K  I  F  G  D  G  T  V  T  A  E  G  E  R  W
580   gccaagctacggaaaatagccaatttttctttcaatggagatagcctcaaaagcatgatcccagcgatgactgggtgtgctgaggccatg
      A  K  L  R  K  I  A  N  F  S  F  N  G  D  S  L  K  S  M  I  P  A  M  T  G  C  A  E  A  M
670   atcgaaaggtggaagcattatgaaggcaaagagatcgacgtcttcgaacaaatgaagatatacacatccgacgtcatttcacatacggct
      I  E  R  W  K  H  Y  E  G  K  E  I  D  V  F  E  Q  M  K  I  Y  T  S  D  V  I  S  H  T  A
760   ttcggaagcagttatgaacaaggaaagcatatcttctacacgctgcagcagatcagcgagctatatatcgagggtgcatataaagttagt
      F  G  S  S  Y  E  Q  G  K  H  I  F  Y  T  L  Q  Q  I  S  E  L  Y  I  R  N  A  Y  K  V  S
850   ttcccagtaataagtaggatttttaaacacaaagaccatatcgaaggagagcggcttcaaaagagaatgaaagaatgcttcatgaaagtc
      F  P  V  I  S  R  I  F  K  H  K  D  H  I  E  G  E  R  L  Q  K  R  M  K  E  C  F  M  K  V
940   atcaagacgagggaagcaaagtttttgagtggcgaagcagatgggtatggcaatgactttctgggattgctggtgaaggcgaagaacgac
      I  K  T  R  E  A  K  F  L  S  G  E  A  D  G  Y  G  N  D  F  L  G  L  L  V  K  A  K  N  D
1030  ccagaaatatctcagaggatttctgtggatgaaattgtggacgaatgcaagacattttactttgccgggcatgagaccaccagcatgttg
      P  E  I  S  Q  R  I  S  V  D  E  I  V  D  E  C  K  T  F  Y  F  A  G  H  E  T  T  S  M  L
1120  ctcacttggactatgtttctgcttgctctgcatcaagaatggcagaggaggcaagaaaggaagcgttcaggatatttggagacagcgat
      L  T  W  T  M  F  L  L  A  L  H  Q  E  W  Q  E  E  A  R  K  E  A  F  R  I  F  G  D  S  D
1210  aacccaacagcagaaggccttgcaaagttaaaaacgatgacgatgatcatcaacgaatgcctgaggctgtatccgccggcgatgacagta
      N  P  T  A  E  G  L  A  K  L  K  T  M  T  M  I  I  N  E  C  L  R  L  Y  P  P  A  M  T  V
1300  tcgcggacagttgagaaggaagtgagactggggaggctggtgcttccgaagacgataaccgtggccgtcccgaccattgcagttcatcag
      S  R  T  V  E  K  E  V  T  L  G  R  L  V  L  P  K  T  I  T  V  A  V  P  T  I  A  V  H  Q
1390  gacacagagtattgggggggaagacgcacatgttttcaaaccagaaagattttcagaagggatcgccaaagtcatcgaaagcaactcggct
      D  T  E  Y  W  G  E  D  A  H  V  F  K  P  E  R  F  S  E  G  I  A  K  V  I  E  S  N  S  A
1480  gggtacctcccatttgggttgggtcctcgaatctgtgtggggatgaactttgctttgaacgaagccaaaattgtaatctccatgattcta
      G  Y  L  P  F  G  L  G  P  R  I  C  V  G  M  N  F  A  L  N  E  A  K  I  V  I  S  M  I  L
1570  cggcgatattccttcacgctgtcgccggcctactctcactcccgattcagttggtcaccgtttgcccccagcatggacttcaggtcata
      R  R  Y  S  F  T  L  S  P  A  Y  S  H  S  P  I  Q  L  V  T  V  C  P  Q  H  G  L  Q  V  I
1660  cttcatccattatcacatcacaaaccttga  1689
      L  H  P  L  S  H  H  K  P  *
```

Figure 3: ORF and deduced amino acid sequences of SgCYP749A22. The 130 - 1689 bp were ORF (compose of 1560 bp); deduced protein was composed of 519 amino acids.

Bioinformatics and Analysis of SgCYP749A22

Physical and chemical property analysis of SgCYP749A22: The SgCYP749A22 protein atomic formula is $C_{2689}H_{4218}N_{726}O_{738}S_{29}$, it encodes a deduced protein of 519 amino acids. Its predicted molecular weight is 59.46 kDa and the theoretical Isoelectric Point (PI) is 9.25. The total number of negatively charged residues (Asp+Glu) are 52, and there are 65 positively charged residues (Arg+Lys). The instability index is computed to be 39.62, this classifies the protein is stable. The aliphatic index is 86.76 and the grand average value of hydropathicity is -0.192.

The secondary structure prediction and three-dimensional (3D) modeling: SgCYP749A22 protein contains 51.06 % alpha helix, 42.58 % loop and 6.36 % strand. A three-dimensional structural model was constructed by SWISS-MODEL. The amino acid of 38-514 were used to establish the model with the 4d75.1. A protein as the template. The template described as Cytochrome P450 3A4. The sequence homology was 25.00 % (Figure 4).

Figure 4: Three-dimensional structure of SgCYP749A22.

The signal peptide, subcellular localization and transmembrane region prediction of SgCYP749A22: The SgCYP749A22 has no signal peptide, but contains a highly conserved leading peptide localized in the chloroplast, including the mitochondrial targeting peptide (mTP, 0.197), the secretory pathway signal peptide (SP, 0.931) and the other peptide (0.006), the predict credibility of which was 2. The SgCYP749A22 protein was localized in the cytoplasmic most likely, the Reliability Index (RI) is 1 and the expected accuracy is 56%. The SgCYP749A22 protein localized in the endoplasmic reticulum transmembrane.

Domains prediction of SgCYP749A22: The results of the NCBI Conserved Domain Database showed that SgCYP749A22 contains the p450 (92-489) domain, CypX (106-511) domain, P450-cycloAA-1 (352-499) domain, belong to the Cytochrome P450, the Cytochrome P450 (Secondary metabolites biosynthesis, transport and catabolism, defense metabolism) and the Cytochrome P450 (cyclodipeptide synthase-associated, members of this subfamily are cytochrome P450 enzymes that occur next to tRNA-dependent cyclodipeptide synthases), respectively. It contains PLN02290 (29-514, a member of the superfamily cl2078 and PLNO2738 (92-500, carotene beta-ring hydroxylase) multiple domains. While the prediction of InterProScan results showed that there were 4 types in total 12 conversed domains (Figure 4), including Cytochrome P450 domain (IPR001128) which consists of PR00385 (P450), SSF48264 (Cytochrome P450), PF00067 (P450) and G3DSA:1.10.630.10; Cytochrome P450, E-class, group1 (IPR002401) which consists of PR00463 (EP4501); Cytochrome P450, conserved site (IPR017972) which consists of PS00086 (Cytochrome P450), and one type has no unintegrated signatures in IPR which consists of the Cytoplasmic-Domain, Non-cytoplasmic domain, PTHR24286 (Cytochrome P450 family member), PTHR24286 (SF20, Subfamily not named) , TMhelix and Transmembrane (Transmembrane region) (Figure 5).

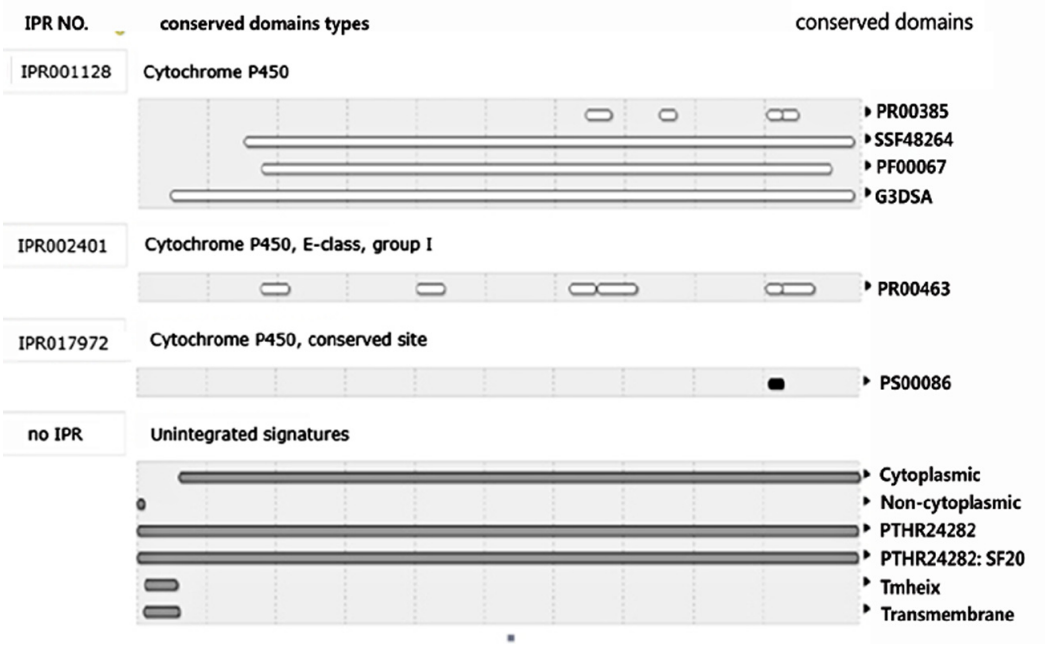

Figure 5: The conserved domains of SgCYP749A22 protein predicted by InterProScan. The conserved domains contained 4 types (Cytochrome P450 domain; Cytochrome P450, E-class, group1 domain; Cytochrome P450, conserved site domain and unintegrated signatures) in total 12 conversed domains (PR00385, SSF48264, PF00067, G3DSA, PR00463, PS00086, Cytoplasmic-Domain, Non-ytoplasmic domain, PTHR24286, PTHR24286 (SF20), TMhelix and Transmembrane).

Amino acid sequence alignment and phylogenetic tree analysis: The Blastp alignment and neighbor-joining tree were constructed for further identifying the relationships between the SgCYP749A22 and other relative protein sequences of other 10 plants already abtained in order to determine the evolutionary relationship among SgCYP749A2*2* and other species. As shown in (Figure 6,7).

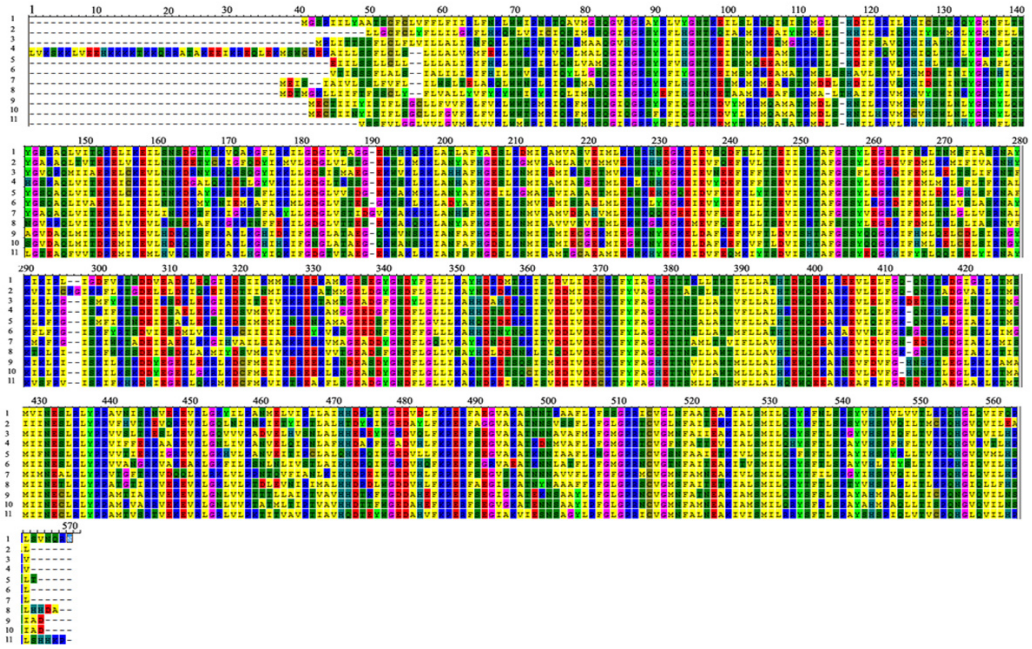

Figure 6: Homology comparison of amino acid sequence of *S. grosvenorii* CYP749A22 with other plants. The different colors represent the different amino acids and the same color represent the same amino acid or the amino acid with the similar functions. 1. *Citrus sinensis* CYP749A22 (XP_ 006484252.1); 2. J*atropha curcas* CYP749A22 (XP_ 012067825. 1); 3. *Malus domestica* CYP749A22 (XP_ 008360080. 1); 4. *Prunus mune* CYP749A22 (XP_ 008243892. 1); 5. *Pyrus X bretschneideri* CYP749A22 (XP_ 009356202. 1); 6. *Ziziphus jujuba* CYP749A22 (XP_ 015894829. 1); 7. *Ziziphus jujuba* CYP749A22 (XP_ 015894829. 1); 8. *Gossypium raimondii* CYP749A22 (XP_ 012461170. 1); 9. *Cucumis melo* CYP749A22 (XP_ 008446466. 1); 10. *Cucumis sativus* CYP749A22 (XP_ 004135430. 2); 11.*S. grosvenorii* CYP749A22 (AEM42984. 1).

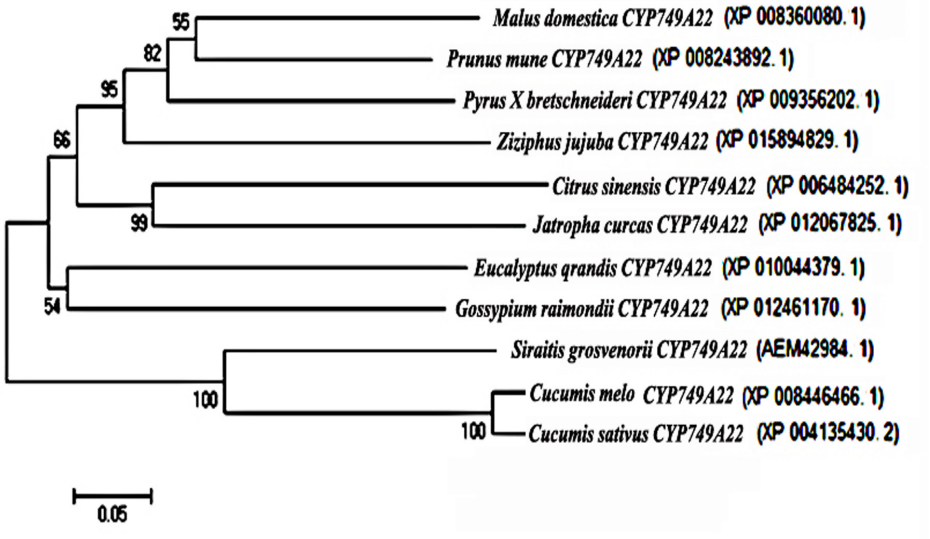

Figure 7: Phylogenetic analysis of SgCYP749A22 and other CYP749A22. SgCYP749A22 lined up with *Cucumis melo* CYP749A22.

SgCYP749A22 lined up with *Cucumis melo* (CmCYP749A22, XP_008446466.1, identify was 68%), followed by *Cucumis sativus* CYP749A22 (identify was 67%, XP_004135430.1), *Malus domestical* CYP749A22 (identify was 53 %, XP_008360080.1). which indicated that the proteins had similar structures and likely enjoyed some gene functions.

The Expression of SgCYP749A22 Gene in Different Samples

The SgCYP749A22 gene expression had a highest level in the fruits of 10 days old after pollinating in the diploid fruits, then had a sharp decreased after 20 days and 30 days, and maintained a lower level of expression during 50 days to 70 days (Figure 8A). While SgCYP749A22 gene expression had a higher level in the fruits of 3 days old after pollinating and to the highest level after 10 days, but sharply decreased after 15 days in the tripliod fruits (Figure 8C). It was constitutively expressed in all tested tissues, but at very different levels. The transcription was the highest in roots, then in stems and leaf, while the weakest in peels (Figure 8B). SgCYP749A22 was positively and significantly induced by hormone treatment. Generally, the expression level in the fruit treated with hormones combining with pollination or induced only by hormone were much higher 4 times than that in the pollination fruits (Figure 8D). SgCYP749A22 gene was highest expressed in triploid fruit and low level in the diploid fruit and tetraploid fruit (Figure 8E).

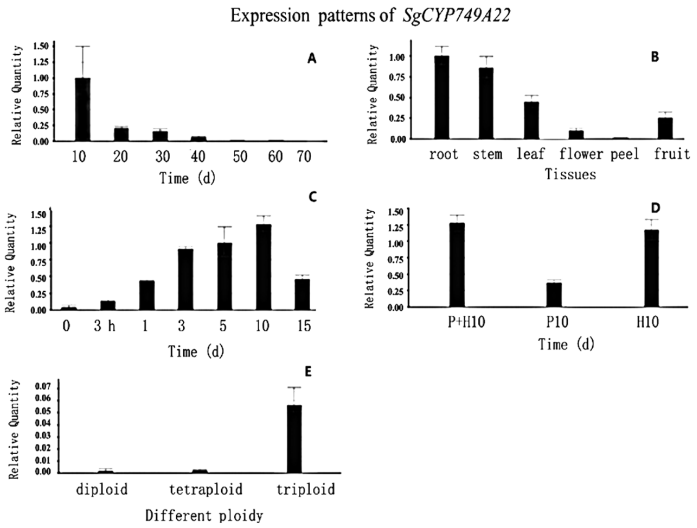

Figure 8: Expression patterns of SgCYP749A22 in the different samples of *S. grosvenorii.* (A) Expression of SgCYP749A22 gene in different development period in diploid fruit. (B) Tissue-specific expression of SgCYP749A22 gene. (C) Expression of SgCYP749A22 gene in different development period in triploid fruit. (D) Expression of SgCYP749A22 gene in different treatment (P: pollination; H: hormones; P+H: pollination combining with hormones). (E) Expression of SgCYP450-4 gene in different ploidy fruits.

Discussion

Nowadays, CYP749A22 gene has been cloned and characterized from *Panax ginseng* [21] but not from *S. grosvenorii*. Here, we firstly reported on the SgCYP749A22 gene cloning and expression. SgCYP749A22 has the highest homology with the CYP749A22 (sequence similarity is 68 %), it may play an important role in the biosynthesis of Mogrosides. The deduced amino acid sequence of SgCYP749A22 showed extensive similarity to its counter parts in other species. In *S. grosvenorii*, SgCYP749A22 is a gene required for Mogrosides biosynthesis. In this study, based on the unigenes sequence of CYP749A22, we designed specific primers and firstly cloned the full-length cDNA sequence from *S. grosvenorii*, named SgCYP749A22. The deduced SgCYP749A22 protein was observed to contain CYP450 domains, as the same with CYP749A22 protein family. Multiple alignments analysis showed that SgCYP749A22 had more than 68% sequence identity with the CYP749A22 proteins of several other species, which suggested that SgCYP749A22 proteins were highly conserved. CYP749A22 proteins were observed to be highly conserved, confirming the high degree of CYP749A22 conservation during the evolution, which reflects the selective pressure imposed by the essential functions of CYP749A22 in plants.

In the mogrosides synthesis pathway, SgCYP749A22 plays the important role in catalyzing the cucurbitadienol into the mogrol, which is the precursor substance of mogrosides of *S. grosvenorii.* Research showed that the *cucurbitadienol synthase* (*CS*) gene had the highest expression level in 10 days of the fruits, the cucurbitadienol synthesized content level was highest too at the same time with the highest content of mogrol, and then decreased [29]. SgCYP749A22 expression level was highest in 10 days, which had the same expression patterns with the *CS*, this suggested that SgCYP749A22 was related to the mogrosides biosynthesis. The results provided a scientific basis for the underlying mechanism of biosynthesis pathway of Mogrosides. But there still were many transcripts of CYP450 in the transcriptome of *S. grosvenorii*, but so far little biological function was exactly examined, therefore, a large number of related research on the CYP450 need to be done.

Conclusions

We cloned a CYP450 gene from *S. grosvenorii* for the first time named SgCYP749A22. The full-length ORF of SgCYP749A22 is 1630 bp, encoding 519 amino acids with a predicted molecular weight of 59.46kDa and an PI of 9.25. SgCYP749A22 gene is mainly expressed in roots and stems, and in fruits. This work may lay a theoretical and experimental foundation for the future research on gene functions, and the transgenic *S. grosvenorii* with varied SgCYP749A22 expression will give deeper insight into the SgCYP749A22 role in *S. grosvenorii.*

Acknowledgments

This study was supported by grants from the National Natural Science Foundation of China (Nos.81560619 and 31400275), Natural Science Foundation of Guangxi province of China (2017GXNSFBA198094) Collaborative Innovation Center of Research on Functional Ingredients from Agricultural Residues

of Guangxi University of Chinese medicine (CICAR 2016- P4) Education Department of Guangxi Zhuang Autonomous Region Education (J17095) and Program for Innovation Research Team in IMPLAD (IT1306).

References

1. State Pharmacopoeia Commission (2010) Pharmacopoeia of the People's Republic of China. Peking: Chinese Medical Science and Technology Press, Beijing. 197.
2. Bai Longhua, Ma Xiaojun, Mo Changming, et al. (2007) Study on the quantitative evaluation of germplasm resources of *Siraitia grosvenorii*. Chinese Journal of Chinese. Materia Medica 32: 2482-2484.
3. Liu Ting, Wangxuhua, Li Chun, Zhang Yi, Wu Guilan, et al. (2007) Study on the antitussive, expectorant and antispasmodic effects of saponin V from *Siraitia grosvenorii*. Chinese Pharm J 42: 1534-1536.
4. Jiang Shiqiong (1999) Research and development of sweet ingredients in *Siraitia grosvenorii.* Journal of Chinese food additives 1: 12-15.
5. Fu W, Ma XJ, Tang Q (2012) Karyotype analysis and genetic variation of a mutant in *Siraitia grosvenorii*. Molecular Biology Report 39: 1247-1252.
6. Tang Qi (2010) High-throughput sequencing of transcriptome and expression profile and clone of key enzyme in the biosynthesis of Mogrosides V in *Siraitia grosvenorii*. Peking: Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College.
7. Chen Quanbin, Yi Xianghui, Yu Lijuan, Yang Ruiyun, Yang Jianxiang (2005) Study on the variation of mogroside V and flavone glycoside in *Siraitia grosvenorii* fresh-fruits in different growth periods. Guihaia 25: 274-277.
8. Li Dianpeng, Chen Yueyuan, Pan Zhenghong, Zhang Hourui (2004) Study on variation of mogrol glycosides from fruits of *Siraitia grosvenori* in different growing ages. Guihaia 24: 546-549.
9. CHAPPLE C (1998) Molecular genetics analysis of plant cytorehrome P450 dependent monooxygenases. Annual review of plant physiology and molecular biology 49: 311-343.
10. Schuler MA, Werck-Reichhart D (2003) Functional genomics of P450. Annu Rev Plant Biol 54: 629-667.
11. Jie Hu, Hiroto Okawa, Kentaro Yamamoto, Kazuhiko Oyama, Masaaki Mitomi, et al. (2011) Characterization of two cytochrome P450 monooxygenase genes of the pyripyropene biosynthetic gene cluster from Penicillium coprobium. The Journal of Antibiotics 64: 221-227.
12. Tudzynski B, Rojas MC, Gaskin P, Hedden P (2002) The gibberllin 20-oxidase of Gibberella fujkuroi is a multifunctional monooxygenase. J Biol Chem 277: 21246-21253.
13. Tudzynski B, Hedden P, Carrera E, Gaskin P (2001) The P450-4 gene of Gibberella fujikuroi encodes ent-kaurene oxidase in the gibberellin biosynthesis pathway. Appl Environ Microbial 67: 3514-3522.
14. Qi X, Bakht S, Qin H (2006) A different function for a member of an ancient and highly conserved cytochrome P450 family: from essential sterols to plant defense. PNAS 103: 18848-18853.
15. Seki H, Ohyama K, Sawai S (2008) Licoruce β-amyrin-11-oxidase cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. PNAS 105: 14204-14209.
16. Seki H, Sawai S, Ohyama K (2011) Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glyeyrrhizin. Plant Cell 23: 4112-4123.
17. Carell M, Biazzi E, Panara F (2011) Medicago truncatula CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytie saponins. Plant Cell 23: 3070-3081.
18. Jung-Yeon Han, Hyun-Jung Kim, Yong-Soo Kwon, Yong-Eui Choi (2011) The cyt P450 enzyme CYP716A47 catalyzes the formation of protopanaxadiol from dammarenediol-II during Ginsenoside biosynthesis in Panax ginseng. Plant Cell Physiol 52: 2062-2073.
19. Huang Zhencong. Studies on the metabolism of Mogroside V in vivo and vitro. Guangxi University of Chinese Medicine: 2013
20. Lu Fenglai, Huang Zhencong, Yan Xiaojie, Chen Yueyuan, Xu Feng, et al. (2014) Stability of mogroside V in artificial gastric juice and its metabolism in vitro. Guihaia 35: 792-795.
21. Marco Biasini, Stefan Bienert, Andrew Waterhouse, Konstantin Arnold, Gabriel Studier, et al. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research. W252-W258.
22. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics 22: 195-201.
23. Kiefer F, Arnold K, Künzli M, Bordoli L, Schwede T (2009) The SWISS-MODEL Repository and associated resources. Nucleic Acids Research 37: 387-392.
24. Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. Electrophoresis. 30: 162-173.
25. Hua SuJun, Sun Zhirong. support vector machine approach for protein subcellular localization prediction. Bioinformatics in press.
26. Tu Dongping, Mo Changming, Ma Xiaojun (2015) Selection of Reference genes of *Siraitia grosvenorii* by real time PCR. Chinese Journal of Chinese Materia Medica. 40: 19-25.
27. Wei rongchang, Zhao Huan, Ma Xiaojun, Mi Ke, Mo Changming, et al. (2014) Extraction of total RNA and cloning of *SgDHAR* gene from *Siraitia grosvenorii.* Acta Pharmaceutica Sinica 49: 115-123.
28. Dongping Tu, Xiaojun Ma, Huan Zhao, Changming Mo, Qi Tang, et al. (2016) Cloning and expression ofSgCYP450-4from *Siraitia grosvenorii*. Acta Pharmaceutica Sinica B 6: 614-622.
29. Zhao huan (2014) Cloning and functional analysis of related enzymes of Mogroside V biosynthesis in *Siraitia grosvenorii*. Peking: Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College.