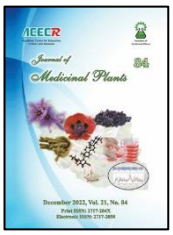
J. Med. Plants 2022; 21(84): 26-38

Institute of Medicinal Plants

Journal of Medicinal Plants

Journal homepage: www.jmp.ir



Research Article

Isolation of CYP72A154, a gene involved in glycyrrhizin biosynthesis pathway, in *Glycyrrhiza glabra* L. (Iranian licorice)

Maryam Allahdou[1,*], Mansour Omidi[2], Mohammad Reza Bihamta[2], Ali Reza Abbasi[3], Barat Ali Fakheri[4]

[1] *Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Zabol, Iran*
[2] *Department of Agronomy and Plant Breeding, Faculty of Agriculture, University of Tehran, Iran*
[3] *Department of Agronomy and Plant Breeding, Faculty of Agriculture, University of Tehran, Iran*
[4] *Department of Plant Breeding and Biotechnology, Faculty of Agriculture, University of Zabol, Iran*

ARTICLE INFO

Keywords:
Glycyrrhiza glabra
Cytochrome P450s
Homology
Bioinformatic analysis
RT-PCR

ABSTRACT

Background: Cytochrome P450s have essential roles in oxidative reactions during the biosynthesis of secondary metabolites, such as terpenoids. **Objective:** This research was aimed to identify the gene CYP72A154 as a gene involved in the glycyrrhizin biosynthesis pathway in Iranian licorice. **Methods:** CYP72A154 gene was isolated from Iranian licorice and cloned it into the *PTG19-T* vector. After confirmation of fragment length, the recombination plasmid was sent for sequencing. NCBI BLAST was used to analyze the nucleotide/ protein sequence homology between *Glycyrrhiza glabra* and other plants. The characterization of predicted amino acid sequences such as sequence homology, protein domains and functional sites, was performed using InterProscan. RT-PCR was performed to improve the relative expression this gene in the licorice root. **Results:** The query length was 314 aa, which after blasting in NCBI had about 78 to 80 % identity to the cytochrome P450 72A154 and 11-oxo-beta-amyrin 30-oxidase in species of *G. glabra*, *G. uralensis* and *G. pallidiflora*, as well as about 64 to 67 % with other species of the *Fabaceae* family. TMHMM analysis indicated the Exp number of AAs in TMHs was 22.11931, Exp number and first 60 AAs was 20.013. The results of RT-PCR revealed that the expression of this gene was comparable to the *β*-AS gene and CYP88D6, both being involved in the glycyrrhizin biosynthesis pathway. **Conclusion**: According to bioinformatics analysis and RT-PCR, it can be stated that the desired fragment belongs to the CYP72A154 gene and is also involved in the biosynthesis of glycyrrhizin.

1. Introduction

Licorice (*Glycyrrhiza glabra*) is one of the most popular medicinal plants in the world [1]. The roots and stolons of Glycyrrhiza have a large amount (8 % of the dry weight) of glycyrrhizin, an oleanane-type triterpene

Abbreviations: MVA, Mevalonate; SQS, Squalene Synthase; *β*-AS, *β*-Amyrin Synthase; EST, Expressed Sequence Tag

*Corresponding author: Maryam.allahdou@uoz.ac.ir

doi: 10.52547/jmp.21.84.26
Received 4 August 2022; Received in revised form 15 October 2022; Accepted 17 October 2022

saponin. Glycyrrhizin has high economic value, therefore the biosynthesis of glycyrrhizin has received extensive attention [2]. The major pharmacological activities of glycyrrhizin are anti-inflammatory [3], immunomodulatory ([4], antiulcer [5], antiallergy activities [6], antiviral activity against various DNA and RNA viruses including HIV, [7] and severe acute respiratory syndrome (SARS)-associated coronavirus [8] and corona virus (Covid 19) [9]. It has been reported that glycyrrhizin strongly inhibits the replication of SARS-CoV-2 in vitro. It blocks viral replication by inhibiting the major viral protease Mpro, which is essential for viral replication [10]. It has been reported that some licorice compounds have inhibitory activity for PD-1/PD-L1, (Programmed death-ligand 1) which is a ligand of PD-1, which is expressed on tumor cells surface [11].

In licorice, glycyrrhizin is mainly synthesized via the MVA pathway. SQS and *β*-AS are two genes that encode enzymes involved in the early stages of triterpene skleten formation. these genes have been successfully isolated from *Glycyrrhiza glabra* [2]. Cytochrome P450s have essential roles in oxidative reactions during the biosynthesis of secondary metabolites, such as terpenoids. As a resource for gene discovery in glycyrrhizin biosynthesis, it was used collection of licorice ESTs to search for glycyrrhizin-biosynthetic genes. As a result, research for putative P450 genes and subsequent transcript profiling-based selection of the candidate P450s led to the identification of CYP88D6. CYP88D6 was shown in vitro and in vivo (in *β*-AS-expressing yeast) to involve in glycyrrhizin biosynthesis pathway [2]. Furthermore, it was reported that CYP72A154, a second relevant P450, is responsible for C-30 oxidation in the glycyrrhizin pathway. It was expressed in an engineered yeast strain which would endogenously produce 11-oxo-b-amyrin (a possible biosynthetic intermediate between *β*-amyrin and glycyrrhizin) [12]. It was isolated an oxidosqualene cyclase cDNA, termed *GgbAS1*, from cultured cells of licorice (*Glycyrrhiza glabra*) [13]. It was cloned a gene encoding *β*-amyrin synthase (*β*-AS) involved in GA biosynthesis in *G. uralensis* and expressed in Saccharomyces cerevisiae [14].

Different genes involved in the biosynthesis of saponins have been isolated from different medicinal plants, such as *GmbAS3* from soybean [15], *β*-amyrin 11-oxidase from Iranian licorice [16], and FPS, SS, SE, and DS genes in the medicinal plant *Panax notoginseng* [17]. genes *β*-amyrin synthase, *β*-amyrin-11-oxidase, 11-oxo-beta-amyrin 30-oxidase and UDP-dependent glucosyltransferase, complete biosynthetic gene cluster committed to glycyrrhizin biosynthesis along with their corresponding promoter regions were isolated from Glycyrrhiza glabra [18].

We aimed to identify and clone the gene CYP72A154 in Iranian licorice and prove the involvement of this gene in the glycyrrhizin biosynthesis pathway through the homology of this gene with other genes involved in this pathway and RT-PCR. This is the first report of the existence of this gene in Iranian licorice of Semiram region. The identification of this gene in Iranian cultivars confirms the existence of the secondary metabolite of glycirrhizin in this species and it will be valuable for future research on metabolite engineering.

2. Materials and Methods

2.1. Plant Materials

Licorice seeds were collected from the Semirom area of Isfahan province. These seeds were disinfected and after scarification with

[DOR: 20.1001.1.2717204.2022.21.84.3.5]
[DOI: 10.52547/jmp.21.84.26]

concentrated sulfuric acid, it was cultured in the ½ MS medium. The cultures were placed at 25 °C and photoperiod 14 h for more than two months.

2.2. Gene Isolation

DNA extraction was performed using cTAB method [19]. For designing the primer, the gene CYP72A63 in Medicago trangatula and CYP72A154 in *Glycyrrhiza uralensis* were used, as they have had more than 95 % homology and are also involved in glycyrrhizin biosynthesis. The sequence of these two genes (CYP72A154, accession number: AB558153 and CYP72A63, accession number: AB558146) was aligned in DNaster software (version 6.13) with Clastal W procedure. Then, a primer pair based on consensus regions of these genes' sequences were designed. Next, using the Oligoanalyser software (version1.0.3), the melting temperature of the primers was investigated and the melting temperature was considered close to each other with a maximum difference of 1 °C. The forward primer sequence was AAATCCATTGGTCTCTCCAAAG and the reverse primer sequence was TCTTCARTGCTTTTTCTCGTTTT. The primers were synthesized by Sinaclone Company.

PCR was performed as follows: 200 ng DNA, 0.5 µM from each of the forward and reverse primers; 10 microliters of Master Mix PCR (Takara company) poured into a microtube and the reaction volume was diluted 20 µl. The PCR program was as follows: 95 °C, 3 min, 34 cycles: 94 °C for 30 seconds, 59 °C (after the temperature gradient) for 1 min, 72 °C for 1 min and finally a final expansion at 72 °C for 5 min. After amplification, the PCR product was electrophoresed in 1 % agarose gel and TAE 1X buffer. DNA extraction from agarose gel was done using Silica gel [20]. The DNA concentration (supernatant) was determined using a Nanodrop device and the peak of curve was 230 nm. The extracted samples were stored at 20 °C until use.

2.3. Molecular cloning

2.3.1. Transfer of DNA desired to T vector

Initially, the competent cell was constructed from Ecoli strain DH5α using the $CaCl_2$ method [21]. Then, the desired DNA segment was inserted into the PTG19-T vector (Vivantis company) using T4 ligase enzymes (Takara company) and ligation reaction according to the company protocol. Then, microtubes, vortexed for a few seconds were spun and exposed to a temperature of about 10 °C overnight. Recombination plasmid was transferred to Ecoli bacterial strain DH5α. After 14 hours of transfer of recombinant plasmid to bacteria, many clones were formed on the culture medium. The clones were transferred to new petri dishes and their lids were tightly closed with parafilm, and incubated at 37 °C for 14 hours overnight.

2.3.2. Extraction of recombinant plasmid

To extract the recombinant plasmid, the bacteria were cultured overnight in liquid LB medium with 100 ppm ampicillin antibiotic. After vortexing, the falcons were placed in a shaker incubator at 37 °C for 14 hours overnight. The plasmid extraction was performed from falcons containing bacteria using Li et al. [22] with slight modifications.

2.3.3. Polymerase chain reaction

To prove the DNA fragment inserted in the plasmid, the extracted plasmids were diluted ten times. PCR was performed using forward and

[DOR: 20.1001.1.2717204.2022.21.84.3.5]
[DOI: 10.52547/jmp.21.84.26]

reverse primer designed for CYP72A154 gene. The PCR program was as follows: 95 °C, 3 min, 34 cycles: 94 °C for 30 seconds, 59 °C for F1-R1 primer (after temperature gradient) for 1 min, 72 °C for 1 min, and finally a final expansion. It was performed at 72 °C for 5 min.

2.3.4. Enzymatic digestion of recombinant plasmid with BamHI enzyme

The vector used in this study was PTG 19-T vector (Vivantis company). Digestion with BamHI was performed using enzyme manufacture company. According to the genomic map, the enzymatic digestion of the recombinant plasmid with this enzyme should produce two fragments, containing plasmid and foreign DNA.

2.4. Bioinformatics analysis

After confirmation of fragment length by colony PCR and enzymatic digestion with BamHI enzyme, the plasmid was sent to Microsynth Switzerland company for sequencing.

Open reading frame (ORF) detection was conducted using the NCBI online tools (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). NCBI BLAST (http://blast.ncbi.nlm.nih.gov) was used to analyze the nucleotide/ protein sequence homology between Glycyrrhiza glabra and other plants. The characterization of predicted amino acid sequences such as sequence homology, protein domains, functional sites, and protein physicochemical properties, was performed using InterProscan (http://www.ebi.ac.uk/Tools/pfa/iprscan/), ProtParam (www.expasy. org/tools/protparam.html), and ScanProsite (http://prosite.expasy.org/ scanprosite/). The phylogenetic trees were drawn by MEGA X [23] and based on the neighbor-joining (NJ) plus maximum likelihood (ML) inference methods.

Amino acid sequences were aligned using CLUSTALW [24]. The evolutionary distances were computed using Poisson correction. A bootstrap analysis with 500 replicates was also conducted to obtain confidence levels for the branches, where branches with confidence levels more than 50 were shown. The protein secondary structure was determined by SOPMA tool (https://npsaprabi.ibcp.fr/). The sub-cellular localization for Plant proteins was predicted using ProtComp 9.0 (http://www.softberry.com). The online tool used for the prediction of transmembrane helices was TMHMM Server v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM-2. 0/).

2.5. RT-PCR

After disinfection and scarification with concentrated sulfuric acid, the seeds of the Iranian species were planted in half MS medium, exposed to 25 °C, and photoperiod of 14 hours of light and 10 hours dark. After about three months, RNA was extracted from the roots, stolons, and leaves of these plants using the cTAB-Litum chloride method [25-26]. To remove possible contamination of genomic DNA with extracted RNA, the samples were treated with DNaseI enzyme (Thermo Scientific's catalog number: EN0521) according to the manufacturer's instructions. For cDNA synthesis, 500 ng of RNA treated with DNase I was used based on the Takara cDNA synthesis kit (catalog number: RR037A) according to the manufacturer's instructions.

The primer design was done based on the sequence of the obtained segment and the largest ORF using Primer 3 software and Oliganalyser. Beta actin gene was used as internal control gene, while two genes β-AS and CYP88D6 were used as genes involved in

[DOR: 20.1001.1.2717204.2022.21.84.3.5]
[DOI: 10.52547/jmp.21.84.26]

glycyrrhizin biosynthesis pathway. There is no glycyrrhizin in licorice leaves [12], thus the leaves were used as a control sample. For all samples, each reaction was performed in four replications (two biological replications and two technical replications). The relative expression of genes was calculated using the formula Fold Chang = $2^{-\Delta\Delta ct}$ [27].

3. Results

3.1. Cloning and blast resulting

The primer designed based on the consensus regions reproduced approximately 1000 bp fragments. The fragment was transferred to the *PTG19-T* vector and the plasmid was transferred to *Ecoli* susceptible cells. The plasmid was extracted from the transformed bacteria, where recombinant plasmids containing the desired fragment were identified (Fig. 1). The recombinant plasmid was amplified by PCR with primers related to the gene and the 1000 bp fragment was amplified (Fig. 2). To confirm the inserted fragment, enzymatic digestion of the recombinant plasmid was performed with *BamHI* enzyme with two fragments separated on agarose gel (Fig. 3). This sequence was recorded in NCBI with accession number: OP056325.

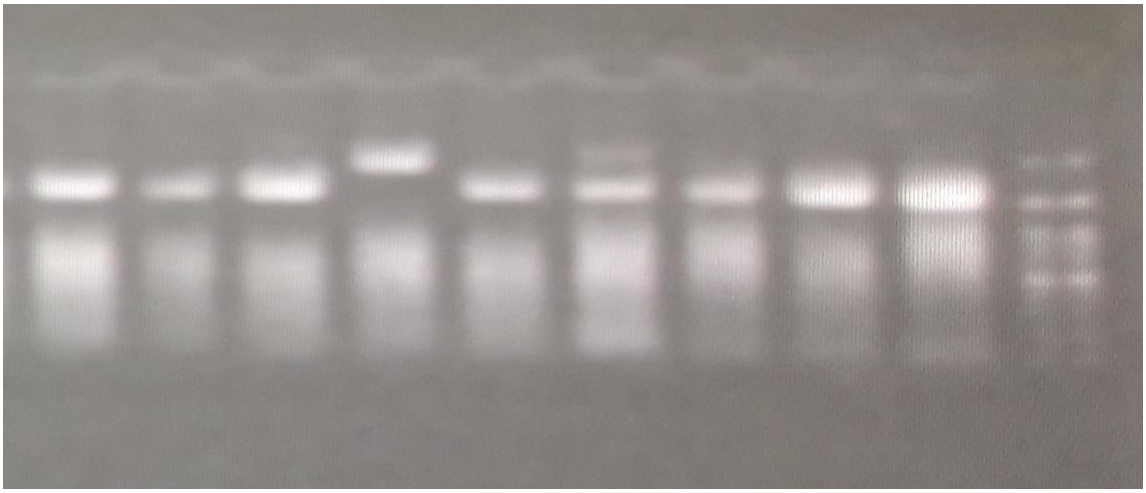


Fig. 1. The recombinant plasmid contained a 1000 bp fragment amplified with F1CYP72A154 primer

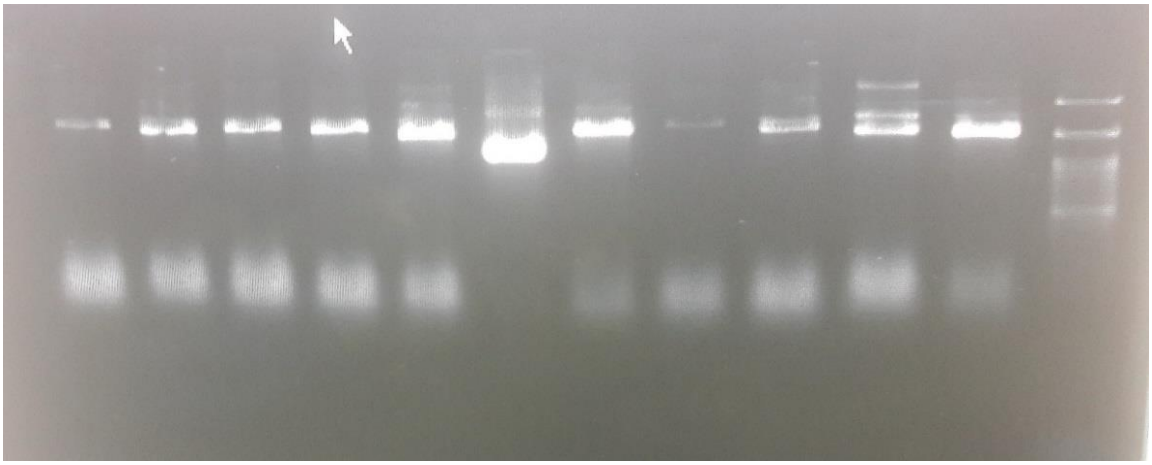


Fig. 2. Amplification of a 1000 bp fragment by an extracted plasmid containing the inserted fragment and F1CYP72A154 primer

[DOR: 20.1001.1.2717204.2022.21.84.3.5]
[DOI: 10.52547/jmp.21.84.26]

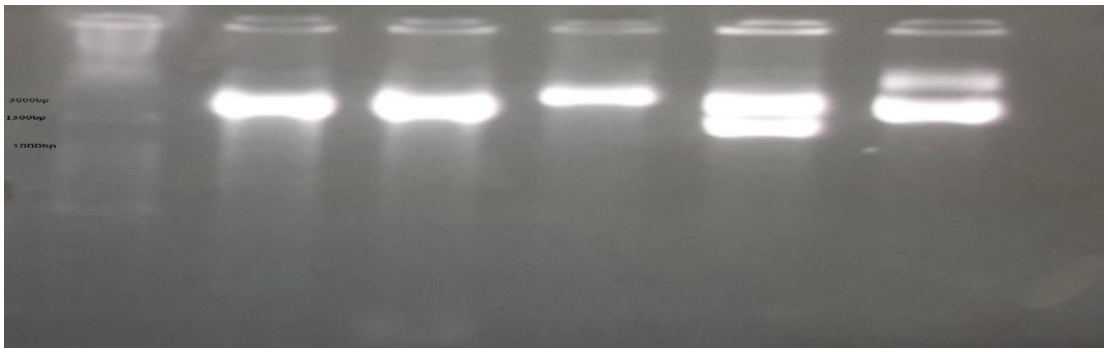


Fig. 3. Enzymatic digestion of recombinant plasmid with *BamHI* enzyme

The sequence was introduced to the www.ncbi.nlm.nih.gov and after blasting in NCBI had 78 to 80 % identity to the cytochrome P450 72A154 and 11-oxo-beta-amyrin 30-oxidase in species of *G. glabra*, *G. uralensis* and *G. pallidiflora*, and about 64 to 67 % with other species of the Fabaceae family (Table 1).

The phylogenetic tree was constructed based on a multiple sequence alignment among the resulting proteins and other proteins of cytochrome P450 from the CYP72 family. The results showed this protein had the greatest similarity with cytochrome P450 from the CYP72A subfamily and 11-oxo-beta-amyrin 30-oxidase (Fig. 4).

3.2. Physicochemical properties of proteins related to this sequence

Initial sequence characterization analysis was performed using the ProtParam tool. The results indicated that the protein related to this sequence had 523 amino acids, its molecular weight was 58997.81, and its isoelectric point was 9.26. The molecular formula was $C_{2690}H_{4251}N_{721}O_{733}S_{18}$ and its number of atoms was 8413. The half-life was estimated 30 hours (mammalian reticulocytes, *in vitro*), > 20 hours (yeast, in vivo) and > 10 hours (*E.coli*, *in vivo*). The instability index (II) was computed to be 38.25; since stable proteins have instability parameter <40, thus this protein would be classified as stable protein.

The aliphatic index, as an important factor in heat resistance, was 99.43, which indicates that this protein is resistant to high heat. The grand average of hydropathicity (GRAVY) was -0.140.

The hydropathicity pattern was calculated using ProtScale with Kyte and Doolittle method [28]. The lowest score was -2.94 in position 310, (D) in polypeptide chain while the highest score was 3.10 in position 17 (Isoleucine) which are related to the strongest and weakest locations of hydrophilicity. The hydrophobic and the hydrophilic domain were located above and below the zero line, respectively (Fig. 5).

The secondary structure of the cytochrome P450 72A154 using SOPMA tool indicated that this protein has 265 Alpha helix (50.67 %), 60 extended strand (11.47 %), 28 Beta turn (5.35 %), and 170 random coil (32.5 %) (Fig. 6).

Prediction of the sub-cellular localization for plant proteins showed that integral prediction of protein location is plasma membrane with score: 9.0. TMHMM analysis revealed Exp number of AAs in TMHs was 22.11931, Exp number, first 60 AAs was 20.013 and Total prob of N-in was 0.9868

[DOR: 20.1001.1.2717204.2022.21.84.3.5]
[DOI: 10.52547/jmp.21.84.26]

Table 1. Characterization of sequences with the desired sequence for CYP72A154 gene in *Glycyrrhiza glabra*

Gene ID	Species	Accesion number	Quary cover %	Identity %
cytochrome P450 72A154	*Glycyrrhiza glabra*	QFS19015.1	58	80.7
11-oxo-beta-amyrin 30-oxidase	*Glycyrrhiza glabra*	QBC36432.1	58	80.7
cytochrome P450 72A154	*Glycyrrhiza uralensis*	H1A988.1	58	79.82
cytochrome P450 monooxygenase	*Glycyrrhiza uralensis*	QMX78160.1	58	78.07
11-oxo-beta-amyrin 30-oxidase	*Medicago truncatula*	RHN40420.1	50	66.67
cytochrome P450 monooxygenase	*Lotus japonicus*	BAL45205.1	59	63.15
11-oxo-beta-amyrin 30-oxidase	*Glycine soja*	RZB82947.1	59	64.91
cytochrome P450 72A697	*Lotus japonicus*	QFS19021.1	59	67.26
cytochrome P450 72A154	*Glycyrrhiza pallidiflora*	QFS19016.1	58	77.19
11-oxo-beta-amyrin 30-oxidase	*Arachis hypogaea*	XP_025610734.1	58	63.15
cytochrome P450 72A698	*Pisum sativum*	QFS19018.1	58	66.67
Cytochrome P450 72A63	*Medicago truncatula*	H1A981.1	58	66.67
cytochrome P450 family 72 protein	*Medicago truncatula*	AET02541.2	58	66.67
cytochrome P450 72A699	*Trifolium pratense*	QFS19031	57	67.26
cytochrome P450 72A698	*Lens culinaris*	QFS19017	58	66.67
cytochrome P450 CYP72A613	*Trigonella foenum-graecum*	QDS03634.1	59	66.67
11-oxo-beta-amyrin 30-oxidase	*Cajanus cajan*	XP_020233595.1	58	63.16
11-oxo-beta-amyrin 30-oxidase	*Glycine soja*	RZB82945.1	59	64.91
cytochrome P450 family 72 protein	*Medicago truncatula*	AET02544.1	59	64.91
11-oxo-beta-amyrin 30-oxidase	*Glycine max*	XP_003543170	59	64.91
11-oxo-beta-amyrin 30-oxidase	*Mucuna pruriens*	RDX90634.1	58	64.04
cytochrome P450 monooxygenase	*Medicago truncatula*	BAL45201	58	64.91

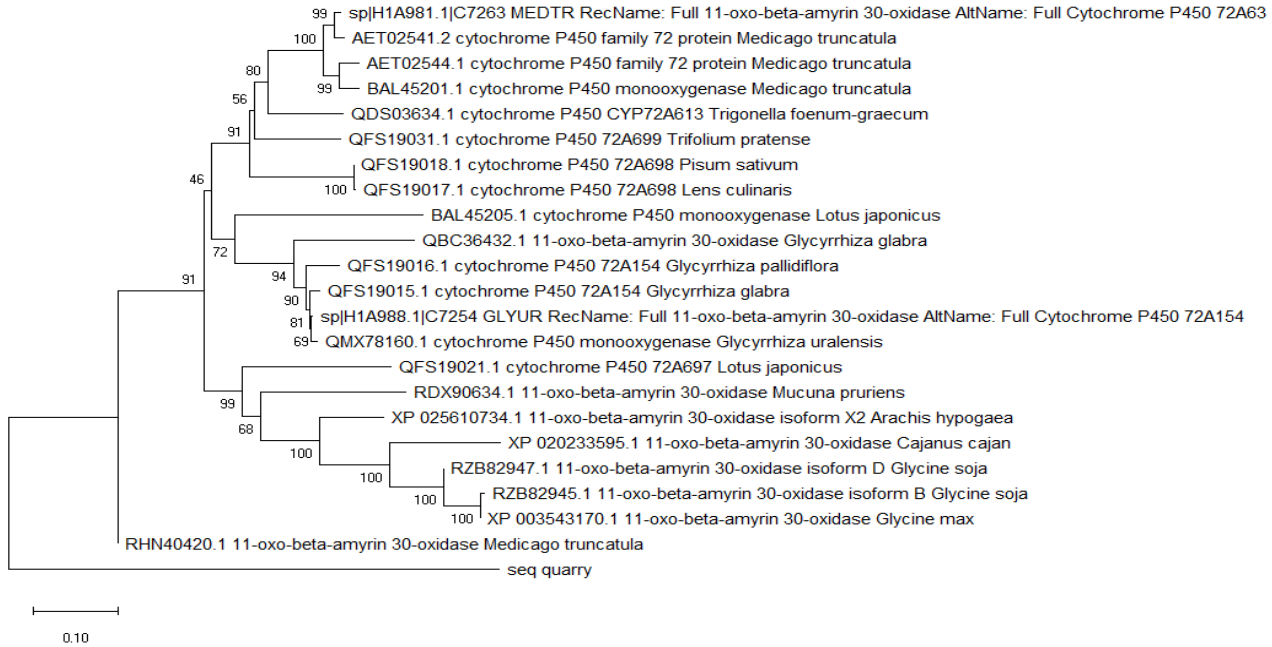


Fig. 4. Phylogenetic relationship between amino acid sequence of *Glycyrrhiza glabra* (seq quarry) and other CYP72A154 proteins

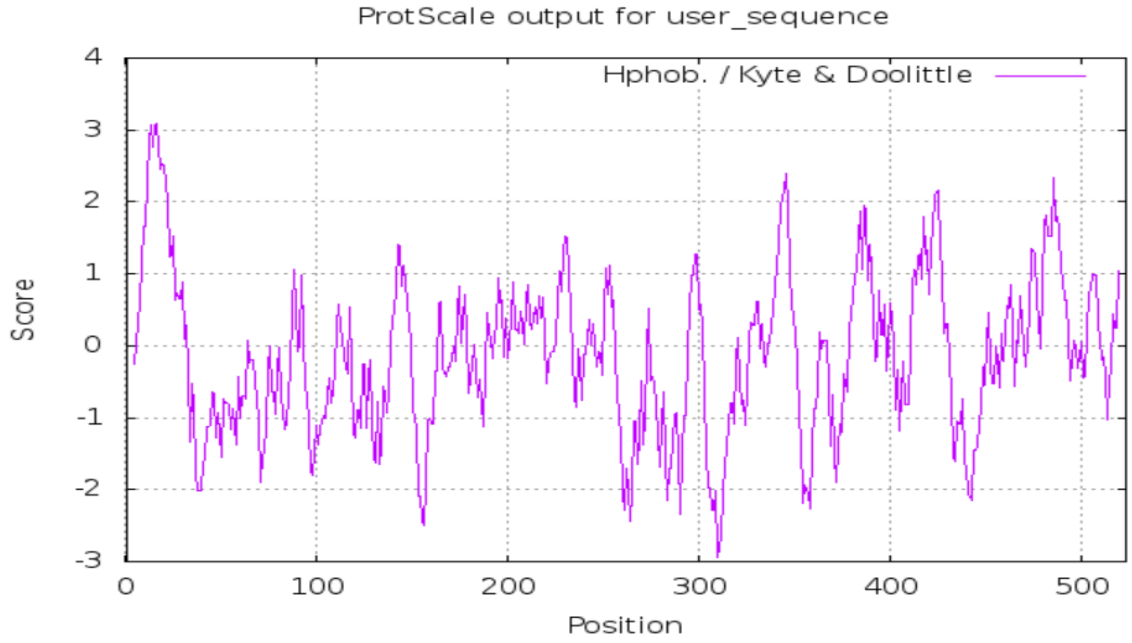


Fig. 5. Hydrophobic-hydrophilic pattern of cytochrome P450 72A154; Hydrophobic domain is located above the zero line and hydrophilic domain is located below the zero line

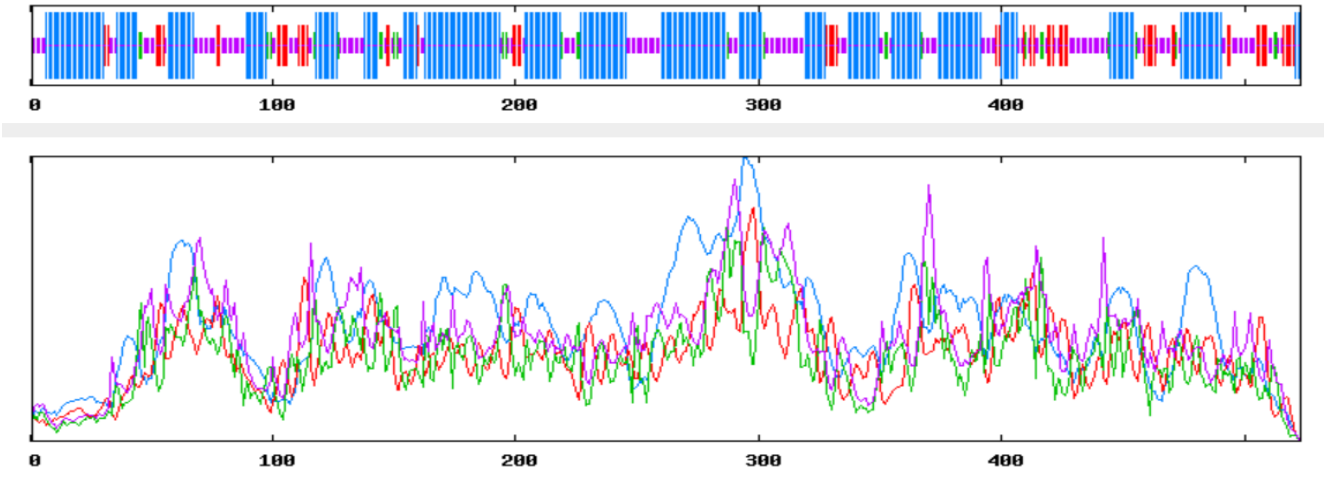


Fig. 6. The diagram of secondary structure of cytochrome P450 72A154 (blue lines), extended strand (red lines), *β*-turn (green line), and random coil (purple lines)

3.3. Expression of CYP72A154 in Iranian licorice

To explore the expression of the CYP72A154 gene using an isolated fragment from the Iranian species, after sequencing and confirming it through bioinformatics analysis, a pair primer was designed based on the sequence of the desired fragment with gene expression analysis also performed. The results indicated that the expression of this gene is comparable to that of the *β*-AS gene and CYP88D6, both being involved in the glycyrrhizin biosynthesis pathway. The mean of fold change was 56.4, 11.08 and 15.6 in b-AS, CYP88D6 and isolated gene from Iranian licorice respectively.

[DOR: 20.1001.1.2717204.2022.21.84.3.5]
[DOI: 10.52547/jmp.21.84.26]

4. Discussion

Licorice grows in different parts of Iran such as Kerman, Sistan and Baluchistan, Isfahan and etc., but the species that grows in the Semirom region of Isfahan has more glycyrrhizin. The Iranian species belonged to *glabra* species. CYP72A154 gene was previously isolated by Seki et al. [12]. In order to isolate this gene from the Iranian species, RNA extraction was performed from the perennial roots, but it did not lead to the isolation of cDNA, so DNA extraction of this plant was performed and according to the designed primers, we were able to isolate a part of CYP72A154 gene. After identifying the ORFs and blasting of protein belonging to this fragment, the results showed that predicted amino acid sequence of CYP72A154 exhibited an extreme similarity with the corresponding enzymes from other plants in the Fabaceae species and was closely related to CYP72A154 gene in the *G. glabra*, *G. uralensis*, and *G. pallidiflora* species.

It was also highly homologous to the *β*-amyrin Synthase gene in this species as well as in *Medicago truncatula*. Most triterpenoid saponins, which are mainly derived from five different triterpenoid aglycones based on the *β*-amyrin skeleton, have been detected from roots and/or aerial parts of the model legume *M. truncatula* [29].

Phylogenic tree revealed that the predicted amino acid belonged to segregated cluster, but had the highest similarity to CYP72A154 and *β*-AS in other plants from Fabaceae family. Seki et al [12] showed that CYP72A154 is responsible for C-30 oxidation in the glycyrrhizin pathway. Seki et al [2] also identified CYP88D6 and CYP93E3 from *G. uralensis* as *β*-amyrin 11-oxidase and *β*-amyrin 24-hydoxylase, respectively. Shirazi et al [13] isolated CYP88D6 from *Glycyrrhiza glabra* (Iranian licorice). They found that isolated segment has 99 % similarity to *β*-amyrin sequence of *Glycyrrhiza uralensis*. Their results corresponded to our results for *β*-amyrin sequence of *Glycyrrhiza uralensis* and *Glycyrrhiza glabra*. The predicted amino acid had similarity to 11-oxo-beta-amyrin 30-oxidase from *Medicago truncatula*, *glycyrrhiza soja*, *Arachis hypogaea*, *Cajanus cajan*, *glycine max*, and *Mucuna pruriens* with these plants belonging to Fabaceae family. Bioinformatic analysis showed the predicted amino acid had the greatest similarity to CYP72A154 protein from *G. glabra* and *G. uralensis*.

This suggests the same role of this enzyme in the biosynthesis pathway of glycyrrhizin in the two species. Grand average of hydropathicity (GRAVY) was negative while higher positive score indicates a greater hydrophobicity.

The expected number of amino acids in transmembrane helices was larger than 18 (22.11), thus it is very likely to be a transmembrane protein, or have a signal peptide. The expected number of amino acids in transmembrane helices in the first 60 amino acids of the protein was 20.013, which is larger than 10, therefore a predicted transmembrane helix in the N-term is a signal peptide. There is a 98% probability that the N-term is on the cytoplasmic side of the membrane. TMHMM analysis in CYP88D6 protein also indicated transmembrane helix at positions of amino acids 3-21 in *β*-amyrin 11-oxidase protein [16].

Analysis of gene expression revealed that expression of CYP72A154 (our quarry) was comparable to that of CYP88D6 in *Glycyrrhiza glabra* (Iranian licorice), therefore we could conclude isolated segment is CYP72A154 and is involved in glycyrrhizin biosynthesis pathway. It was also found that CYP72A154 expression in the roots and stolon was at levels

[DOR: 20.1001.1.2717204.2022.21.84.3.5]
[DOI: 10.52547/jmp.21.84.26]

comparable to those of *β*-AS and CYP88D6. They suggested a role for CYP72A154 in glycyrrhizin biosynthesis [12]. *β*-AS gene had higher expression than CYP88D6 and CYP72A154 possibly due to the fact that the *β*-AS gene is involved in both the primary pathway of glycyrrhizin biosynthesis and in the pathway of soyasaponin biosynthesis [10]. There is an inverse relationship between glycyrrhizin and soyasaponin [30]. Soyasaponin was observed in all parts of the plant, and in contrast to glycyrrhizin, was mostly observed in younger parts of stolons [2].

It was isolated *β*-amyrin Synthase Gene (GmBAS3) from soybean where the expression of the gene was detected in 21-day-old seedlings in the hypocotyls, young leaves, and mature leaves but not observed in stem and root tissues [15]. On the other hand, in our study and studies conducted by another researcher *β*-AS gene was expressed in root and stolon in licorice [2-8]. In another study pBI121^{GUS-9}: CYP88D6 construct was transferred to *G. glabra* using Agrobacterium rhizogene but was not observed increase in glycyrrhizin in all transferred hairy root lines, therefore, in order to increase the amount of glycyrrhizin, it is necessary to transfer several genes involved in the biosynthesis pathway [31].

Thus, it could be concluded that the isolated gene (CYP72A154) is also involved in biosynthesis pathway of glycyrrhizin, as with CYP88D6 [2-16] and *β*AS [13]. The results of phylogenetic tree determined the most similar amino acid sequence to the CYP88D subfamily of cytochrome P450.

5. Conclusion

Predicted amino acid sequence of CYP72A154 exhibited an extreme similarity with the corresponding enzymes from other plants in the Fabaceae species and was closely related to CYP72A154 gene in the *G. glabra*, *G. uralensis*, and *G. pallidiflora* species. It was also highly homologous to the *β*-amirin synthase gene in this species as well as in Medicago truncatula. Therefore we could conclude isolated segment is CYP72A154 and is involved in glycyrrhizin biosynthesis pathway.

Author contributions

M.A.: Contributed to the conception of the study, doing laboratory work, data collection and analysis, interpretation of data, drafting the manuscript. M.O.: Supervised the study, assist in the analysis and interpretation of results, M.R.B.: Supervised the study, formal analysis, reviewing and editing the manuscript. A.R.A.: Collaboration in primer design and bioinformatics analysis. B.A.F.: Cooperation in collecting seeds and final editing of the article.

Conflicts of interest

The authors confirm that there are no conflicts of interest.

Acknowledgments

The authors thank the Research Council of the University of Tehran for their financial support. We also wish to thank Prof. Houshang Alizadeh and Dr. Zamani for their kind help in the lab.

References

1. Li Y-L, Yang Y, FU ChH and Yu LJ. Production of Glycyrrhizin in cell suspension of *Glycyrrhiza inflate* batalin cultured in Bioreactor. *Biotech. Biotechn. Equip.* 2012; 26(5): 3231-3235. doi: 10.5504/ BBEQ.2012. 0083.

2. Seki H, Ohyama K, Sawai S, Mizutani M, Ohnishi T, Sudo H, Akashi T, Aoki T, Saito K

[DOR: 20.1001.1.2717204.2022.21.84.3.5]
[DOI: 10.52547/jmp.21.84.26]

and Muranaka T. Licorice *β*-amyrin 11-oxidase, a cytochrome P450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proc. Natl. Acad. Sci. USA Proc. Natl. Acad. Sci. USA* 2008; 105(37): 14204-14209. doi: 10.1073/pnas.0803876105.

3. Matsui S, Matsumoto H, Sonoda Y, Ando K, Aizu-Yokota E, Sato T and Kasahara T. Glycyrrhizin and related compounds down-regulate production of inflammatory chemokines IL-8 and eotaxin 1 in a human lung fibroblast cell line. *Int. Immunopharmacol* 2004; 4(13): 1633-1644. doi: 10.1016/ j.intimp. 2004.07.023.

4. Takahara T, Watanabe A and Shiraki K. Effects of glycyrrhizin on hepatitis B surface antigen: a biochemical and morphological study. *J. Hepatol*. 1994; 21(4): 601-609. doi: 10.1016/S0168-8278(94)80108-8.

5. He J-X, Akao T, Nishino T and Tani T. The influence of commonly prescribed synthetic drugs for peptic ulcer on the pharmacokinetic fate of glycyrrhizin from Shaoyao-Gancao-tang. *Biol. Pharm. Bull.* 2001; 24(12): 1395-1399. doi: 10.1248/bpb.24.1395.

6. Park H-Y, Park S-H, Yoon H-K, Han M-J and Kim D-H. Anti-allergic activity of 18beta-glycyrrhetinic acid-3-O-beta-D-glucuronide. *Arch. Pharm. Res.* 2004; 27: 57-60. doi: 10.1007/BF02980047.

7. Ito M, Sato A, Hirabayashi K, Tanabe F, Shigeta Sh, Baba M, Clercq ED, Nakashima H and Yamamoto N. Mechanism of inhibitory effect of glycyrrhizin on replication of human immunodeficiency virus (HIV). *Antiviral Res.* 1988; 10(11): 289-298. doi: 10.1016/0166-3542 (88)90047-2.

8. Cinatl J, Morgenstern B, Bauer G, Chandra P, Rabenau H, Doerr HW. Glycyrrhizin, an active component of liquorice roots, and replication of SARS-associated coronavirus. Lancet 2003; 361(9374): 2045-2046. doi: 10.1016/s0140-6736(03)13615-x.

9. Chrzanowski J, Chrzanowska A and Grabo´n W. Glycyrrhizin: an old weapon against a novel coronavirus. *Phytother. Res.* 2021; 35(2): 629-636. doi: 10.1002/ptr.6852.

10. Van de Sand L, Bormann M, Alt M, Schipper L, Heilingloh CS, Steinmann E, Todt D, Dittmer U, Elsner C, Witzke O and Krawczyk A. Glycyrrhizin effectively inhibits SARS-CoV-2 replication by inhibiting the viral main protease. *Viruses* 2021; 13(4): 1-10. doi: 10.3390/v13040609.

11. Bao F, Bai H-Y, Wu Z-R and Yang Z-G. Phenolic compounds from cultivated *Glycyrrhiza uralensis* and their PD-1/PD-L1 inhibitory activities. *Natu. Pro. Res.* 2021: 35(4): 562-569. doi: 10.1080/ 14786419. 2019. 1586698.

12. Seki H, Sawai S, Ohyama K, Mizutani M, Ohnishi T, Sudo H, Fukushima EO, Akashi T, Aoki T, Saito K and Muranaka T. Triterpene functional genomics in licorice for identification of CYP72A154 involved in the biosynthesis of glycyrrhizin. *Plant Cell*. 2011; 23(11): 4112-4123. doi: 10.1105/tpc.110.082685.

13. Hayashi H, Huang P, Kirakosyan A, Inoue K, Hiraoka N, Ikeshiro Y, Kushiro T, Shibuya M and Ebizuka Y. Cloning and characterization of a cDNA encoding. *β*-amyrin synthase involved in glycyrrhizin and soyasaponin Biosynthesis in Licorice. *Bio Pharm Bull* 2001; 24(8):912-916. doi: 10.1248/bpb.24.912.

14. Chen H, Liu Y, Zhang X, Zhan X and Liu C. Cloning and characterization of the gene encoding *β*-amyrin synthase in the glycyrrhizic acid biosynthetic pathway in *Glycyrrhiza uralensis*. *Acta Pharmaceutica Sinica B* 2013; 3(6):416-424. doi: 10.1016/j.apsb.2013.09.002.

15. Ali MM, Krishnamurthy P, El-Hadary M.H, Kim J.M, Nawas M.A, Yang S.H and

Chung G. Identification and expression profiling of a new *β*-amyrin synthase gene (GmbAS3) from soybean. *Russ. J. Plant Physio.* 2016; 63: 383-390. doi: 10.1134/ S1021443716020035.

16. Shirazi Z, Aalami1 A, Tohidfar M and Sohani MM. Isolation, cloning and bioinformatics analysis of *β*-amyrin 11-oxidase coding sequence from licorice. *Plant Omics J.* 2016; 9(2): 165-171. doi: 10.21475/poj. 160902.p7778x.

17. Niu Y, Luo H, Sun C, Yang TJ, Dong L, Huang L and Chen S. Expression profiling of the triterpene saponin biosynthesis genes FPS, SS, SE, and DS in the medicinal plant *Panax notoginseng*. *Gene* 2014; 533(1): 295-303. doi: 10.1016/j.gene.2013.09.045.

18. Goyal P, Muzafar Manzoor M, Gupta AP, Pandotra P and Gupta S. Molecular dissection of genes and promoters involved in glycyrrhizin biosynthesis revealed phytohormone induced modulation in *Glycyrrhiza glabra* L. *Gene* 2022; 836: 146682. doi: 10.1016/j.gene. 2022.146682.

19. Agbagw I.O, Datta S, Patil PG, Singh P and Nadarajan N. A protocol for high-quality genomic DNA extraction from legumes. *Genet Mole Res.* 2012; 11(4): 4632-4639. doi: 10.4238/2012.September.14.1.

20. Ausubel FM, et al. Eds, Current protocols in molecular biology, John Wiley & Sons, Inc., NY, 1997. doi: 10.1002/0471142727. mb0500s64.

21. Chang AY, Chau VWY, Landas JA and Pang Y. Preparation of calcium competent escherichia coli and heat-shock transformation. *JEMI methods* 2017; 1: 22-25.

22. Li X, Jin H, Wu Z, Rayner S and Wang B. A continuous process to extract plasmid DNA based on alkaline lysis. *Nature Protocols*. 2008; 3: 176-180. doi: 10.1038/nprot.2007.526.

23. Kumar S, Strecher G, Li M, Knyaz C and Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms, *Mol. Biol. Evol.* 2018; 35(6): 1547-1549. doi: 10.1093/molbev/msy096.

24. Thompson JD, Higgins DG and Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994; 22: 4673-4680. doi: 10.1093/nar/22.22.4673.

25. Zarei A, Zamani Z, Mousavi A, Fatahi R, Karimi Alavijeh M, Dehsara B and Salami SA. An effective protocol isolation of high-quality RNA from Pomegranate seeds. *The Asian and Australasian Journal of Plant Science and Biotechnology* 2012*;* 6(Special Issue 1): 32-37.

26. Chang S, Puryear S and Cairney J. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter* 1993; 11(2): 113-116. doi: 10.1007/ BF02670468.

27. Pfaffl MW**.** A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 2001; 29: 45-49. doi: 10.1093/nar/29.9. e45.

28. Kyte J and Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 1982; 157(1): 105-132. doi: 10.1016/0022-2836(82)90515-0.

29. Huhman DV and Sumner LW. Metabolic profiling of saponins in *Medicago sativa* and *Medicago truncatula* using HPLC coupled to an electrospray ion-trap mass spectrometer. *Phytochemistry* 2002; 59(3): 347-360. doi: 10. 1016/s0031-9422(01)00432-0.

30. Hayashi H and Sudo H. Economic importance of licorice. *Plant Biotech.* 2009; 26(1): 101-104. doi: 10.5511/ plantbiotechnology.26.101.

31. Shirazi Z, Aalami A, Tohidfar M and Sohani MM. Metabolic engineering of glycyrrhizin pathway by over-expression of *β*-amyrin 11-oxidase in transgenic roots of *Glycyrrhiza glabra*. *Mol. Biote.* 2018: 60: 412-419. doi: 10.1007/s12033-018-0082-7.

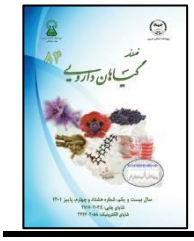
How to cite this article: Allahdou M, Omidi M, Bihamta MR, Abbasi AR, Fakheri BA. Isolation of CYP72A154, a gene involved in glycyrrhizin biosynthesis pathway, in *Glycyrrhiza glabra* L. (Iranian licorice). ***Journal of Medicinal Plants*** 2022; 21(84): 26-38.
doi: 10.52547/jmp.21.84.26

فصلنامه گیاهان دارویی ۱۴۰۱؛۲۱(۸۴): ۲۶-۳۸

فصلنامه گیاهان دارویی

Journal homepage: www.jmp.ir

پژوهشکده گیاهان دارویی
جهاد دانشگاهی



مقاله تحقیقاتی

جداسازی ژن CYP72A154 به عنوان ژن درگیر در مسیر بیوسنتز گلیسیریزین در *Glycyrrhiza glabra* L. (شیرین بیان ایرانی)

مریم الهدو[۱،*]، منصور امیدی[۲]، محمدرضا بی‌همتا[۲]، علیرضا عباسی[۳]، براتعلی فاخری[۴]

[۱] *گروه اصلاح نباتات و بیوتکنولوژی، دانشکده کشاورزی، دانشگاه زابل، زابل، ایران*

[۲] *گروه زراعت و اصلاح نباتات، دانشکده کشاورزی، دانشگاه تهران، تهران، ایران*

[۳] *گروه مستقل بیوتکنولوژی، دانشکده کشاورزی، دانشگاه تهران، تهران، ایران*

[۴] *گروه اصلاح نباتات و بیوتکنولوژی، دانشکده کشاورزی، دانشگاه زابل، زابل، ایران*

اطلاعات مقاله

گل‌واژگان:
شیرین‌بیان
سیتوکروم P450
همسانی
آنالیز بیوانفورماتیک
RT-PCR

چکیده

مقدمه: سیتوکروم P450 نقش مهمی در واکنش‌های اکسیداتیو در طول بیوسنتز متابولیت‌های ثانویه از جمله ترپنوئیدها ایفا می‌کند. **هدف:** هدف از این تحقیق شناسایی ژن CYP72A154 به عنوان ژن دخیل در مسیر بیوسنتز گلیسیریزین در شیرین‌بیان ایرانی بود. **روش بررسی:** ژن CYP72A154 از شیرین‌بیان ایرانی جدا و در ناقل PTG19-T کلون شد. پس از تایید طول قطعه، پلاسمید نوترکیب برای تعیین توالی ارسال شد. از NCBI BLAST برای تجزیه و تحلیل همولوژی توالی نوکلئوتیدی/پروتئینی بین *Glycyrrhiza glabra* و سایر گیاهان استفاده شد. توصیف توالی‌های اسید آمینه پیش‌بینی‌شده مانند همسانی توالی، دومین‌های پروتئینی و مکان‌های عملکردی، با استفاده از InterProscan انجام شد. RT-PCR برای بررسی بیان نسبی این ژن در ریشه شیرین‌بیان انجام شد. **نتایج:** طول کوآری 314 aa بود که پس از بلاست در NCBI حدود ۷۸ تا ۸۰ درصد یکسانی با سیتوکروم P450 72A154 و 11-oxo-beta-amyrin 30-oxidase در گونه‌های *G. glabra*، *G. uralensis* و *G. pallidiflora* و همچنین حدود ۶۴ تا ۶۷ درصد با سایر گونه‌های خانواده Fabaceae داشت. تجزیه و تحلیل TMHMM نشان داد که تعداد Expe اسیدآمینه‌ها در TMH ها ۲۲/۱۱۹۳۱ و تعداد Exp، ۶۰ آمینو اسید اول ۲۰/۰۱۳ بود. نتایج RT-PCR نشان داد که بیان این ژن با ژن b-AS و CYP88D6 که هر دو در مسیر بیوسنتز گلیسیریزین نقش دارند، قابل مقایسه است. **نتیجه‌گیری:** با توجه به تجزیه و تحلیل بیوانفورماتیک و RT-PCR می‌توان بیان کرد که قطعه مورد نظر متعلق به ژن CYP72A154 بوده و در بیوسنتز گلیسیریزین نیز نقش دارد.

مخففها: MVA، موالانات؛ SQS، اسکوالن سنتاز؛ β-AS، بتا آمیرن سنتاز؛ EST، تگ توالی بیان شده

* نویسنده مسؤول: Maryam.allahdou@uoz.ac.ir

تاریخ دریافت: ۱۳ مرداد ۱۴۰۱؛ **تاریخ دریافت اصلاحات:** ۲۳ مهر ۱۴۰۱؛ **تاریخ پذیرش:** ۲۵ مهر ۱۴۰۱

doi: 10.52547/jmp.21.84.26