

## مطالعه بیوانفورماتیکی و بررسی الگوی بیان چند ژن مهم دخیل در سنتز گلیسیریزین شیرین بیان (*Glycyrrhiza glabra L.*) بومی ایران طی دو فصل پاییز و بهار

عاطفه خاکپور<sup>۱</sup>، مریم ذوالفقاری<sup>۲\*</sup> و کریم سرخه<sup>۳</sup>

۱- دانشجوی کارشناسی ارشد، گروه علوم باغبانی، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز

۲- استادیار، گروه علوم باغبانی، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز

۳- استادیار، گروه مهندسی تولید و ژنتیک گیاهی، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز

(تاریخ دریافت: ۱۳۹۷/۰۴/۱۲ - تاریخ پذیرش: ۱۳۹۷/۰۹/۱۷)

### چکیده

شیرین بیان یکی از گیاهان دارویی با ارزش و در معرض انقراض است. شناسایی و معرفی توده‌هایی که ماده مؤثره گلیسیریزین بیشتری دارند از برنامه‌های اصلاحی مهم است. روش‌های ژنومیکس کارکردی مانند تجزیه و تحلیل توالی‌های EST، امکان شناسایی خانواده‌های ژنی حفاظت‌شده بین گونه‌های مورد مطالعه و بررسی بیان ژن و تفسیر ژنوم را فراهم آورده است. در این پژوهش به منظور شناسایی جنبه‌های مولکولی و تحلیل کارکردی ژنوم و شبکه ژنی دخیل در سنتز گلیسیریزین، ۵۵۹۶۰ توالی EST مربوط به دو گونه مختلف این گیاه و هم‌چنین توالی‌های پروتئینی مورد تجزیه و تحلیل بیوانفورماتیکی قرار گرفتند. هم‌چنین جهت اعتبارسنجی نتایج، تغییرات بیان نسبی چهار ژن مهم در مسیر بیوسنتز گلیسیریزین شامل اسکوالن سنتاز (SQS)، بتا-آمرین سنتاز (BAS)، بتا-آمرین ۱۱-اکسیداز (CYP88D6) و UDP-گلوکوکورونوسیل ترانسفراز (UGT) مورد بررسی قرار گرفت. پس از پیرایش و کیفیت‌سنجی اولیه توالی‌ها، ۶۴۲۷ توالی کانتیگ و ۳۰۸۹۵ سینگل تون (۳۷۳۲۲ یونی ژن) ایجاد گردید که در مجموع ۲۶۸۸۴۶۶۶ جفت‌باز (۷/۰۶ درصد) از ژنوم شیرین بیان را پوشش دادند. فعالیت عملکردی ژنومی نشان داد که اکثر ژن‌ها در فعالیت کاتالیزوری و فرآیندهای سلولی و متابولیکی نقش داشتند و معلوم شد که این ژن‌ها در درون سلول و اندامک‌های درون سلولی فعالیت می‌کنند. مکان‌یابی این دسته از ژن‌ها نشان داد که ژن‌های مهم مسیر سنتز گلیسیریزین در شبکه آندوپلاسمی متمرکز شده‌اند. نتایج اعتبارسنجی با استفاده از qRT-PCR نشان داد که در فصل پاییز و در بافت ریزوم، ژن‌های BAS، CYP88D6، UGT و SQS دارای بیان نسبی بیشتری بودند. نتایج حاصل از این مطالعه می‌تواند برای توالی‌یابی ژنوم، حاشیه‌نویسی (گروه‌های کارکردی)، تنوع ژنتیکی و نیز مطالعات عملکرد مولکولی و ژنومیکسی این گیاه ارزشمند باشد.

**واژگان کلیدی:** ژنوم کارکردی، شبکه ژنی، شیرین بیان، گلیسیریزین، qRT-PCR

\* نویسنده مسئول، آدرس پست الکترونیکی: m.zolfaghari@scu.ac.ir

## مقدمه

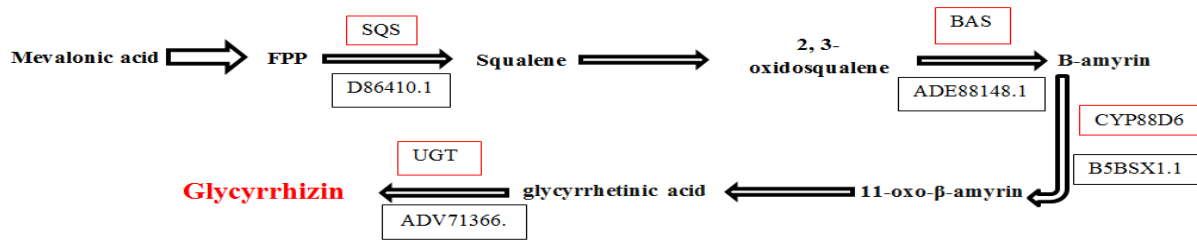
شیرین بیان متعلق به خانواده‌ی بقولات (Fabaceae) است که شامل گونه بومی جنوب اروپا، هندوستان و بخشی از آسیا (*Glycyrrhiza. glabra L.*) و گونه بومی چین (*G. uralensis F.a*) می‌باشد (Rebhun *et al.*, 2015). سال‌هاست این گیاه در سراسر دنیا به‌عنوان یک گیاه دارویی و شیرین‌کننده طبیعی مورد استفاده قرار می‌گیرد. گلیسیریزین<sup>۱</sup> ماده اصلی موجود در ریشه و ریزوم شیرین بیان، عامل شیرین شدن این گیاه است که به‌عنوان ماده شیرین‌کننده در بسیاری از محصولات تجاری استفاده می‌شود (Pandey and Ayangla, 2017). علاوه بر گلیسیریزین ترکیبات فعال دیگری نیز از قبیل فلاونوئیدها، کومارین‌ها، تانن‌ها، پلی‌ساکاریدها، پکتین‌ها، قندهای ساده، آمینواسیدها و نمک‌های معدنی در ریشه‌ی شیرین بیان وجود دارند (Bayati zadeh *et al.*, 2013). از جمله از فلاونوئیدهای مهم ریشه‌ی شیرین بیان کوئرستین<sup>۲</sup> و گلابریدین<sup>۳</sup> می‌باشد (Fukai *et al.*, 1996). خواص دارویی مانند اثرات ضدالتهابی، ضدویروسی، ضد میکروبی، آنتی‌اکسیدان، ضدآسم و نیز فعالیت‌های ضدسرطانی برای شیرین بیان در نظر گرفته شده است (Hosseinzadeh and Nassiri-Asl, 2015) و تا به حال داروهای ضدالتهابی، ضدسرطانی و ضدویروسی از شیرین بیان ساخته شده‌اند (Tang *et al.*, 2015). گلیسیریزین یک تری‌ترپنوئید<sup>۴</sup> ساپونینی<sup>۵</sup> می‌باشد که در قسمت‌های چوبی ریشه‌های ضخیم شده و ریزوم شیرین بیان وجود دارد (Pandey and Ayangla, 2017). مسیر پیشنهادی برای بیوسنتز گلیسیریزین و تعیین ژن‌های دخیل در این مسیر در شکل ۱ نشان داده شده است. این مسیر از چرخه اصلی موالونیک اسید<sup>۶</sup> شروع شده که در ابتدا فارنسیل دی‌فسفات<sup>۷</sup> (FDP) از طریق آنزیم اسکوالن سنتاز<sup>۸</sup> (SQS) به اسکوالن و در مرحله بعد توسط آنزیم اسکوالن اپوکسیداز<sup>۹</sup> (SQE) به 2, 3-

oxidosqualene تبدیل می‌شود. یکی از مهم‌ترین آنزیم‌هایی که در سنتز گلیسیریزین نقش دارد بتا-آمرین سنتاز<sup>۱۰</sup> (BAS) می‌باشد که باعث تبدیل شدن 2, 3-oxidosqualene به بتا-آمرین شده و در آخر بتا-آمرین توسط آنزیم‌های سیتوکروم P450s<sup>۱۱</sup> (CYP88D6) و CYP72A154 (CYP72A154) تبدیل به اسیدگلیسیرینیک شده و در انتها از طریق آنزیم مهم UDP-گلاکوسیل ترانسفراز (UGT<sup>۱۲</sup>) به گلیسیریزین تبدیل می‌شود.

بررسی‌های بیوانفورماتیکی ساختار اسیدنوکلئیک و پروتئین‌ها و مقایسه، طبقه‌بندی و پیش‌بینی ساختار آن‌ها می‌تواند در پیش‌بینی موقعیت سلولی پروتئین، بازسازی مسیرهای متابولیکی و شبیه‌سازی و تحلیل عملکردی آن‌ها مؤثر واقع باشد (Zarabi *et al.*, 2013). ESTها<sup>۱۳</sup> توالی‌های کوتاه cDNA (با میانگین طول ۵۰۰ جفت‌باز) هستند که به‌وسیله‌ی توالی‌یابی دقیق انتهای ۵' یا ۳' همسانه‌های<sup>۱۴</sup> (کلون‌های) cDNA ایجاد می‌شوند. کاربرد اولیه‌ی ESTها کشف ژن است. اما از کاربردهای پیچیده‌تر آن‌ها می‌توان به شناسایی خانواده‌های ژنی حفاظت‌شده بین گونه‌های مورد بررسی، شناسایی گوناگونی‌های ویرایش رونوشت ژن و تعیین ساختار ژن از طریق ترادف‌سازی<sup>۱۵</sup> و انطباق رونوشت‌ها، بررسی بیان ژن و تفسیر ژنوم<sup>۱۶</sup>، شناسایی ژن‌های جدید، نقشه‌یابی ژنوم و شناسایی مناطق کدکننده اشاره کرد (AliAkbari *et al.*, 2013; Zarabi *et al.*, 2013). داده‌های خام کتابخانه‌ها و جزئیات مربوط به هر کدام را می‌توان از پایگاه‌های داده مختلف همچون NCBI<sup>۱۷</sup> و PlantGDB<sup>۱۸</sup> به‌دست آورد. وجود این منابع غنی از اطلاعات در کنار در دسترس بودن سرویس‌ها و نرم‌افزارهای مختلف بیوانفورماتیک راه را برای مطالعه و بررسی تغییرات ژنومی هموار نموده است (Ogata and Suzuki, 2011).

10-  $\beta$ -amyrin synthase  
11- Cytochrome P450  
12- UDP-glucosyltransferases  
13- Expressed Sequenced Tags  
14- Clone  
15- Alignment  
16- Genome annotation  
17- National center for biotechnology information (<https://www.ncbi.nlm.nih.gov>)  
18- Plant genome database

1- Glycyrrhizin  
2- Quercetin  
3- Glabridin  
4- Triterpenoids  
5- Saponin  
6- Mevalonic acid  
7- Farnesyl diphosphate  
8- Squalene synthase  
9- Squalene epoxidase



شکل ۱- مسیر بیوسنتز گلیسیریزین

Figure 1. Glycyrrhizin biosynthesis pathway

SQS: اسکوالن سنتاز؛ BAS: بتا-آمرین سنتاز؛ CYP88D6: بتا-آمرین ۱۱-اکسیداز؛ UGT: UDP-گلوکورونوسیل ترانسفراز  
 SQS: squalene synthase; BAS:  $\beta$ -amyrin synthase; CYP88D6:  $\beta$ -amyrin 11-oxidase; UGT: UDP-glucuronosyltransferase

کیفی توالی‌های EST و حذف وکتورها، توالی‌های تکرار شونده، توالی‌های کلروپلاستی، میتوکندریایی از پایگاه اطلاعاتی Egassembler<sup>4</sup> استفاده گردید. خروجی این سرویس شامل یک فایل دارای دسته‌هایی با دو یا تعداد بیشتری توالی EST ادغام شده و با طول مشخص است که برای یک ژن خاصی بیان می‌شوند و در پاره‌ای از موارد حفاظت شده‌اند که به آن‌ها کانتینگ می‌گویند و فایل دیگر شامل توالی ESTهای منفرد است که وارد هیچ کانتیگی<sup>5</sup> نمی‌شوند و به آن‌ها سینگل تون<sup>6</sup> می‌گویند.

**یافتن توالی‌های مشابه:** برای بررسی میزان همولوژی توالی‌های EST مربوط به ژنوم شیرین بیان از نرم‌افزار CLC Genomics Workbench 9 برای تمامی توالی‌های کانتینگ‌ها با در نظر گرفتن حداکثر E-value برابر با ۱۰ در مقابل بانک توالی نوکلئوتیدی گیاه *Arabidopsis thaliana* به‌عنوان گیاه مرجع به‌صورت Local Blast استفاده شد.

**دریافت توالی‌های پروتئینی:** توالی‌های پروتئینی همانند توالی‌های EST مربوط به گونه‌های ذکر شده گیاه شیرین بیان از پایگاه اطلاعاتی NCBI به فرمت Fasta دریافت گردید. برای تعیین فعالیت عملکردی، ترسیم نقشه‌ی ژنتیکی، تعیین نقطه‌ی ایزوالکتریک<sup>7</sup>، وزن مولکولی<sup>8</sup> و مکان‌یابی سلولی<sup>9</sup>، توالی‌های پروتئینی هر دو

با توجه به اهمیت نقش ساپونین‌ها و نیز گلیسیریزین به‌عنوان ترکیبات با ارزش دارویی، لازم است که تحقیقات بیشتری از نظر ژنتیکی، مولکولی و بیوانفورماتیکی به‌منظور آشکارسازی مسیر بیوسنتزی آن‌ها صورت گیرد؛ بنابراین داشتن اطلاعات کافی در زمینه ژنتیکی و مولکولی برای درک بهتر مکانیسم و چگونگی تنظیم و بیان این ژن‌های مهم و تأثیر آن‌ها بر تولید مقدار این ترکیب مهم دارویی در شیرین بیان ضروری به نظر می‌رسد (Kim *et al.*, 2011; Mochida *et al.*, 2017; Nasrollahi *et al.*, 2014). هدف از مطالعه حاضر واکاوی کتابخانه‌های EST مربوط به سه گونه مختلف شیرین بیان؛ شناسایی عملکرد مولکولی<sup>1</sup>، فرآیندهای زیستی<sup>2</sup> ژن‌ها و اجزای سلولی<sup>3</sup> مرتبط با آن‌ها و ترسیم نقشه‌ی ژنی و مشخص کردن فعالیت عملکردی توالی پروتئینی می‌باشد. در ادامه به‌منظور اعتبارسنجی نتایج به‌دست آمده از واکاوی ژن‌های دخیل در بیوسنتز گلیسیریزین در طی دو فصل بهار و پاییز، از فن qRT-PCR استفاده شد.

## مواد و روش‌ها

### مطالعات بیوانفورماتیکی

**دریافت کتابخانه‌های بیانی EST ژن‌های بیوسنتز گلیسیریزین و کیفیت‌سنجی توالی‌های EST:** ۵۵۹۶۰ توالی EST مربوط به کتابخانه‌ی دو گونه‌ی *Glycyrrhiza uralensis* F.a و *G. glabra* L. از پایگاه داده NCBI به فرمت Fasta دریافت شد. به‌منظور سنجش و ارزیابی

4- <http://egassembler.hgs.jp>  
 5- Contig  
 6- Singleton  
 7- Isoelectric point  
 8- Molecular weight  
 9- Cell location

1- Molecular function  
 2- Biological processes  
 3- Cellular components

گردیدند. در مرحله بعد عمل سنتز cDNA طبق دستورالعمل کیت prime script RT reagent (تاکارا، ژاپن) و با استفاده از آغازگرهای Oligo dt و Random انجام شد.

طراحی آغازگرهای رفت و برگشتی مربوط به ژنهای مهم و ژن اکتین جهت آنالیز بیان ژن: با استفاده از نرم‌افزار آنالیز Primer<sup>4</sup> آغازگرهای مرتبط با هریک از ژنهای اسکوالن‌سنتاز (SQS) (D86410.1)، بتا-آمرین‌سنتاز (BAS) (ADE88148.1)، بتا-آمرین-اکسیداز (CYP88D6) (B5BSX1.1) و UDP-گلوکورونوسیل‌ترانسفراز (UGT) (ADV71366.1) و ژن اکتین (EU190972.1) به‌عنوان ژن مرجع طراحی گردیدند. لیست آغازگرها در جدول ۱ آورده شده است.

**بررسی میزان بیان ژن‌ها با استفاده از واکنش qRT-PCR:** جهت بررسی میزان بیان ژنهای مورد نظر و مقایسه آن‌ها در دو فصل مختلف، از واکنش زنجیره‌ای پلیمراز در زمان واقعی استفاده شد. این واکنش با استفاده از مخلوط واکنش تاکارا حاوی SYBR-Green و آغازگرهای اختصاصی مرتبط با ژن‌ها انجام گردید. جهت بررسی اختصاصی بودن آغازگرهای مورد استفاده از نمودار نقطه ذوب در طیف دمای ۹۵ درجه سانتی‌گراد استفاده شد و به ازای هر سیکل ۰/۳ درجه سانتی‌گراد کم شد تا به دمای ۶۰ درجه سانتی‌گراد (دمای اتصال آغازگرها) برسد. برای تجزیه و تحلیل نتایج از ژن اکتین به‌عنوان ژن کنترل داخلی استفاده شد. آنالیز نتایج حاصل از qRT-PCR با استفاده از روش  $2^{-\Delta\Delta Ct}$  انجام گردید. معادله  $2^{-\Delta\Delta Ct}$  زمانی استفاده می‌شود که تنها هدف تشخیص بیان نسبی هر یک از ژن‌ها و نسبت به ژن کنترل داخلی می‌باشد (Schmittgen and Livak, 2008). آزمون t-test بر روی داده‌ها برای تشخیص تفاوت‌ها و معنی‌دار بودن یا نبودن از نظر آماری، استفاده شد. در نهایت پردازش داده‌ها و رسم نمودارها با استفاده از نرم‌افزار Excel انجام شد.

### نتایج و بحث

گونه باهم ادغام گردیدند و توالی‌های اضافی حذف شدند.

**ترسیم نقشه‌ی ژنی و مشخص کردن فعالیت عملکردی توالی پروتئینی:** تعیین نقشه‌ی ژنی برای شناسایی ارتباط بین ژن‌های مربوط به گونه‌های *Glycyrrhiza* با استفاده از پایگاه String<sup>1</sup> در مقابل بانک توالی پروتئینی گیاه *Arabidopsis thaliana* ترسیم گردید. هم‌چنین به‌منظور تعیین فعالیت عملکردی توالی‌ها (که شامل عملکرد مولکولی، فرآیندهای بیولوژیکی و اجزای سلولی می‌باشد) از همین سرویس در مقابل بانک توالی پروتئینی گیاه *A. thaliana* استفاده شد.

**تعیین نقطه‌ی ایزوالکتریک (Pi)، وزن مولکولی (Mw) و مکان‌یابی سلولی توالی‌های پروتئینی:** به‌منظور تعیین نقطه‌ی ایزوالکتریک (Pi) و وزن مولکولی (Mw) مربوط به توالی‌های پروتئینی از سرویس بیوانفورماتیکی Expasy<sup>2</sup> compute pi/mw tool استفاده گردید. مکان‌یابی سلولی توالی‌های پروتئینی توسط پایگاه softberry<sup>3</sup> انجام گردید.

### مطالعات آزمایشگاهی

#### جمع‌آوری نمونه‌های گیاهی شیرین‌بیان، استخراج

**RNA کل و سنتز cDNA:** ریزوم‌ها و برگ‌های تازه شیرین‌بیان (*Glycyrrhiza glabra*) از منطقه منصوریه واقع در شهرستان بهبهان (در ۳۰ درجه و ۲۶ دقیقه عرض شمالی و ۵۰ درجه و ۱۶ دقیقه طول شرقی در منتهی‌الیه جنوب شرقی استان خوزستان و با ۳۲۰ متر ارتفاع از سطح دریا) طی دو فصل پاییز و بهار در تاریخ ۲۱ آبان سال ۱۳۹۶ و ۲۱ اردیبهشت سال ۱۳۹۶ جمع‌آوری و تا زمان استخراج RNA در ازت مایع نگه‌داری گردیدند. ۵۰-۲۵ میلی‌گرم از نمونه‌های ریزوم و برگ برای استخراج RNA، توسط ازت مایع پودر گردیدند. استخراج RNA کل از بافت‌ها با استفاده از کیت Biobasic کانادا مطابق دستورالعمل شرکت انجام شد. کمیت و کیفیت نمونه‌های RNA با استفاده از اسپکتروفتومتری بررسی شد. نمونه‌های RNA جهت حذف DNA ژنومی باقی‌مانده با آنزیم DNase I (Takara, Ltd., Japan) تیمار

1- <https://string-db.org>  
2- <https://web.expasy.org>  
3- <http://www.softberry.com>

4- <http://primer3.ut.edu>

ژن‌های قبل و بعد خود در این مسیر ارتباط نزدیک‌تری در مقایسه با ژن‌های دیگر دارند. فهرست برخی از این ژن‌های مهم به‌همراه سایر اطلاعات مربوط به آن‌ها در جدول ۲ آورده شده است. ژن اسکوالن‌ستناز (SQS) کدکننده آنزیمی است که در واکنش تبدیل فارنسیل‌دی‌فسفات (FDP) (به‌عنوان یکی از محصولات چرخه موالونات که شروع کننده چرخه سنتز تری‌ترپنوئیدها می‌باشد) به اسکوالن شرکت می‌کند (Shirazi et al., 2017). ژائو و همکاران (Zhao et al., 2017) گزارش نمودند که این آنزیم در گیاهان نقش مهمی در تنظیم بیوسنتز فیتواسترول‌ها و تری‌ترپنوئیدها دارد. این ژن در شبکه ژنی موجود با ژن (FPS1) فارنسیل‌دی‌فسفات‌ستناز ۱ ارتباط بسیار نزدیک و قوی دارد. ژن بتا-آمرین‌ستناز از دیگر ژن‌های مهم و کلیدی سنتز گلیسیریزین است که متعلق به گروهی از آنزیم‌های مهم کاتالیزکننده به‌نام اکسیدواسکوالن‌سیکلازها<sup>۳</sup> می‌باشد (Hosseini Panahi et al., 2016) که در چرخه سنتز تری‌ترپنوئیدها در واکنش تبدیل اکسیدواسکوالن به بتا-آمرین نقش دارد. این ژن در شبکه ژنی با ژن SQS ارتباط نزدیک دارد.

سیتوکروم‌ها آنزیم‌های بسیار پراهمیت و کلیدی هستند که درون شبکه ژنی موجود، سه نوع ژن CYP93D1، CYP716A1 و AT3G32047 وجود دارد که کدکننده آنزیم مهم سیتوکروم P450 هستند و نقش مهمی در واکنش‌های اکسیداتیو دارد. این آنزیم به‌صورت متوالی دو مرحله اکسیداسیون تبدیل بتا-آمرین به ۱۱-اکسو-بتا-آمرین<sup>۴</sup> را کاتالیز می‌کند (Seki et al., 2008). برای این ژن نیز درون شبکه ژنی ارتباط نزدیکی با ژن BAS در نظر گرفته شده است. ژن UGT73B3 کدکننده آنزیم UDP-گلوکزیل‌ترنسفرآز است که در مرحله نهایی تولید گلیسیریزین نقش کلیدی و اختصاصی ایفا کرده و اسیدگلیسیرینیک را به گلیسیریزین تبدیل می‌کند. UGTs، چرخه گلایکوسیل‌اسیون<sup>۵</sup> را با استفاده از یک دهنده قند

آنالیز توالی‌های EST: پس از ارزیابی اولیه ۵۵۹۶۰ توالی EST، برای هر دو گونه شیرین‌بیان در مجموع ۶۴۲۷ توالی کانتینگ با میانگین طول جفت‌باز ۶۰۰/۹۴۴ و ۳۰۸۹۵ توالی سینگل‌تون باقی ماند. تعداد کل Unigenes (که شامل کانتینگ‌ها و سینگل‌تون‌ها می‌باشد) ۳۷۳۲۲ مشخص گردید و در مجموع، این توالی‌ها، ۲۶۸۸۴۶۶۶ جفت‌باز از ژنوم شیرین‌بیان را پوشش دادند که این مقدار چیزی در حدود ۷/۰۶ درصد از کل ژنوم تخمین زده شده برای شیرین‌بیان را در بر می‌گیرد. با توجه به ژنوم بسیار بزرگ شیرین‌بیان و این‌که ESTها فقط نماینده بخش کدشونده ژنوم می‌باشند، این مقدار وسیع کاملاً مناسب است (جدول ۲).

خروجی به‌دست آمده از BLAST در NCBI با استفاده از نرم‌افزار CLC Genomics Workbench 9 به‌صورت HSP<sup>۱</sup> می‌باشد و HSP شامل دو قطعه توالی دلخواه و یکنواخت است که هم‌ترازی<sup>۲</sup> آن حداکثر می‌باشد. با به‌نتایج به‌دست آمده از انجام BLAST برای کل کانتینگ‌ها حداکثر HSP در مقابل توالی‌های نوکلئوتیدی گیاه مرجع آرابیدوپسیس ۲۰۱۸ تا و کم‌ترین HSP صفر به‌دست آمد. در میان ۶۴۲۷ توالی کانتینگ تنها تعداد ۱۴ HSP صفر نشان دادند و مابقی کانتینگ‌ها حداقل یک Hit داشتند که این به‌معنای داشتن شباهت زیاد توالی ژنوم شیرین‌بیان با توالی‌های گیاه مرجع آرابیدوپسیس است.

تعیین شبکه ژنی مربوط به ژن‌های مهم دخیل در سنتز گلیسیریزین: با توجه به شکل ۲، کاملاً مشخص می‌باشد که شبکه ژنی پیچیده‌ای بین ژن‌های مهم دخیل در مسیر سنتز گلیسیریزین وجود دارد. البته بسیار واضح است که این شبکه تنها بخشی از پیچیدگی‌های کارکردی ژنوم شیرین‌بیان را در بر می‌گیرد. به‌دلیل حضور ژن‌هایی که در مسیر سنتز گلیسیریزین هستند، این بخش بسیار حائز اهمیت است. از بین ژن‌های دخیل در سنتز گلیسیریزین تقریباً اکثر ژن‌های مهم و کلیدی درون شبکه موجود هستند. با توجه به شکل ۲ می‌توان دریافت که به شکل جالبی ژن‌هایی که در تولید گلیسیریزین نقش دارند با

3- Oxidoscoalen cyclase  
4- 11- Exo-beta amrin  
5- Glycosylation

1- High-scoring segment pair  
2- Alignment

اجزای سلولی برای این ژن‌ها در نظر گرفته شده است درون کمپلکس سنتز ATP برای منتقل کننده‌ی پروتون است. در پژوهش انجام شده توسط لی و همکاران (Li et al., 2010)، برای به دست آوردن یک نمای کلی از ویژگی‌های ژن‌های مربوط به *Glycyrrhiza uralensis*، با توجه به آنالیزهای هستی‌شناسی ژن‌ها، ویژگی‌های ژن‌های درگیر و شناخته شده در فرآیندهای متنوعی قرار گرفتند. تعداد زیادی از آن‌ها به عنوان هیدرولازها، کینازها و ترانسفرازها شناسایی شدند. بیشترین فعالیت آن‌ها در فرآیند بیولوژیکی مربوط به فرآیندهای سلولی بود و نیز برای عملکرد اجزای سلولی آن‌ها بیشترین فعالیت در اجزای سلولی، سلول و سیتوپلاسم نشان داده شد. این مشاهدات با نتایج به دست آمده در این مطالعه از واکاوی ژن‌های دخیل در سنتز گلیسیریزین مطابقت دارد.

**نقطه‌ی ایزوالکتریک (Pi)، وزن مولکولی (Mw) و مکان‌یابی سلولی توالی‌های پروتئینی:** جهت تعیین نقطه ایزوالکتریک، وزن مولکولی و نیز مکان سلولی ژن‌های دخیل در سنتز گلیسیریزین، در ابتدا پروتئین‌های اضافی حذف گردیده و همان‌طور که در جدول ۴ و ۵ مشاهده می‌شود پروتئین‌های مهم مربوط به مسیر بیوسنتز گلیسیریزین مشخص گردید. این آنزیم‌ها اکثراً از نظر مکان سلولی در شبکه آندوپلاسمی<sup>۱</sup> قرار گرفته‌اند. ساپونین‌های تری‌ترپنوئیدی در سیتوزول<sup>۲</sup> تولید می‌شوند و با توجه به این‌که گلیسیریزین هم نوعی ساپونین تری‌ترپنوئیدی است در سیتوزول ساخته شده و چون ترکیبی آب‌دوست می‌باشد برای ذخیره به واکوئل منتقل می‌شود (Wink, 2010).

هم‌چنین آنزیم‌های کاتالیزگر در شبکه آندوپلاسمی قرار دارند و برای فعالیت خود از طریق انتقال‌دهنده‌هایی به سیتوزول منتقل می‌شوند. دیگر آنزیم‌های دخیل در مسیر سنتز گلیسیریزین (اسکوآل‌سنتاز، لوپئول‌سنتاز، سیکلوآرتنول‌سنتاز، بتا-آمرین‌سنتاز، سیتوکروم P450 و UDP-گلیکو‌ترانسفراز) نیز در شبکه آندوپلاسمی قرار گرفته‌اند.

مانند UDP-گلوکز، UDP-گالاکتوز، UDP-اکسیلوز و UDP-گلوکورونیک‌اسید کاتالیز می‌کند. به نظر می‌رسد که تغییرات در تعداد زنجیره‌های قند، ترکیب و موقعیت آن‌ها در ساختار تری‌ترپنوئیدها تأثیر زیادی بر انتقال و ذخیره داخل و خارج سلولی در گیاهان دارد (Seki et al., 2015). هم‌چنین این ژن در پاسخ‌های دفاعی و تنشی هم درگیر می‌شود (جدول ۳).

**هستی‌شناسی ژن (Gene ontology) و فعالیت عملکردی توالی‌های مربوط به ژنوم شیرین‌بیان:** برای پیش‌بینی فعالیت عملکردی در میان توالی‌های پروتئینی مرتبط به ژنوم شیرین‌بیان از پایگاه String استفاده گردید. که شامل سه دسته‌ی عملکرد مولکولی، فرآیندهای بیولوژیکی و اجزای سلولی می‌باشند (شکل ۳). همان‌طور که در شکل ۳ مشاهده می‌شود، این ژن‌ها از نظر عملکرد مولکولی نقش بسیار زیادی در فعالیت کاتالیزوری ایفا می‌کنند و نیز به‌طور قابل‌توجهی در اکسیدوردوکتاز فعالیت دارند. کم‌ترین کاربرد برای آن‌ها در نظر گرفته شده است در فعالیت oxidosqualene cyclase است. بیشترین سهم افزایشی برای ژن‌های مربوطه در فرآیندهای زیستی، مربوط به فرآیندهای سلولی و متابولیکی می‌باشند و کم‌ترین نقش این ژن‌ها از نظر فرآیندهای زیستی در فسفوریلاسیون اکسیداتیو و فرآیند بیوسنتز استیل کوآنزیم A می‌باشد.

در مطالعات کیم و همکاران (Kim et al., 2011) نیز گزارش شده است که در فرآیندهای بیولوژیکی، ژن‌ها در فرآیندهای متابولیکی و فرآیندهای متابولیسم سلولی به‌طور قابل‌توجهی نقش دارند و عملکرد مولکولی ژن‌ها بیشتر روی اتصالات پروتئینی و یون‌ها و فعالیت کاتالیزوری متمرکز شده‌اند. این ژن‌ها از نظر اجزای سلولی اکثراً درون سلول و اندامک‌های درون سلولی فعالیت زیادی دارند که این را می‌توان با سنتز ترکیبات تری‌ترپنوئیدی (گلیسیریزین) که اکثراً تولید آن‌ها در سیتوزول هست تطبیق داد. آنزیم‌های درگیر در این مسیر سنتز از شبکه‌ی آندوپلاسمی به سیتوزول منتقل و در آنجا فعالیت می‌کنند (Wink, 2010). کمترین فعالیتی که از نظر

1- Endoplasmic reticulum  
2- Cytosol

جدول ۱- آغازگرهای رفت و برگشتی طراحی شده مربوط به ژن‌های مورد مطالعه و ژن اکتین

Table 1. Designed forward and reverse primers for the studied genes and actin gene

نام آغازگر	توالی‌های نوکلئوتیدی (5' → 3')	دمای اتصال (سانتی‌گراد)	اندازه محصول (جفت‌باز)
Name of primer	Nucleotide sequences (5' → 3')	Annealing temperature (°C)	Product size (bp)
SQS-forward	GCGAGGTCTAACTGCCAAAG	60.5	225
SQS- reverse	CCGTTCCCATTCCTCAATAA	58.4	
bAS-forward	TGCGGACATTGTAGTTGAGC	58.4	220
bAS- reverse	ACCAAGTGCGAACCAAGAAC	58.4	
CYP88D6-forward	TCGGCGATCTATTGTCCTTC	58.4	230
CYP88D6- reverse	GGGTCTACATCGTGCCAAC	60.5	
UGTs-forward	CCCAACGACTTCAACCTCAT	58.4	217
UGTs- reverse	ACCCAAGTTTTGTGCTCCAC	58.4	
Actin-forward	TCCTTCGTCTGGATCTTGCT	58.4	248
Actin- reverse	GCAACGGAATCTCTCAGCTC	60.5	

جدول ۲- نتایج پالایش و همگذاری توالی‌های EST

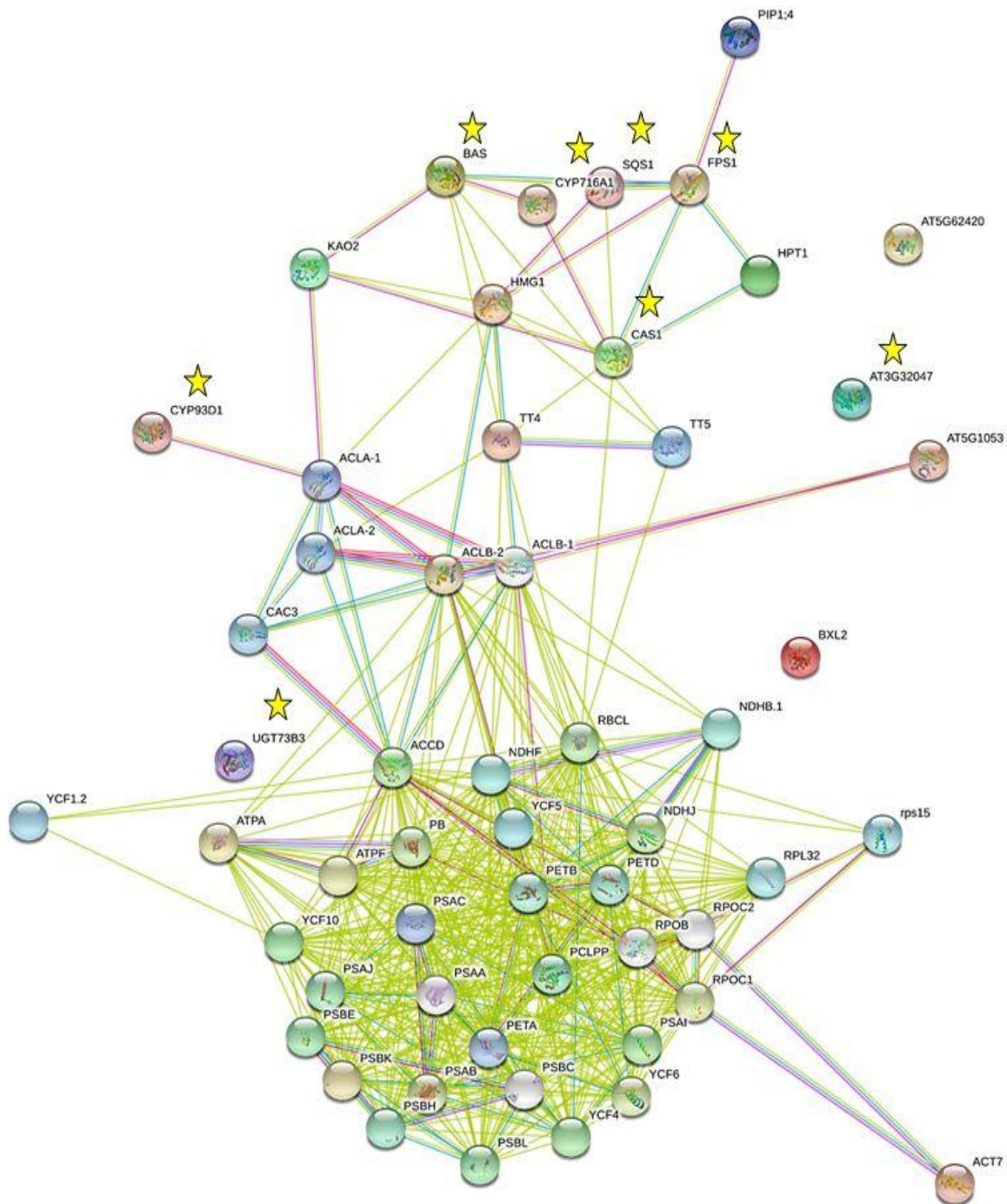
Table 2. Results of ESTs cleaning and assembling

طول کل پوشش داده شده از ژنوم (جفت‌باز)	تعداد کل یونی ژنها	تعداد سینگلتون	میانگین طول توالی‌های (جفت‌باز) کانتیگ	تعداد کانتیگ	تعداد کل توالی‌های EST
Total length of the genome covered (bp)	Total number of unigenes	Number of singelton	Average length of contig sequences (bp)	Number of contig	Total number of sequences EST
26884666	37322	30895	600.944	6427	55960

جدول ۳- ژن‌های مهم مسیر سنتز گلیسریریزین درون شبکه ژن

Table 3. Important genes for synthesis of glycyrrhizin within the gene network

گره	شناسه ژن	حاشیه‌نویسی
Node	Gene ID	Annotation
BAS	AT1G78950.1	Beta-amyrin synthase; Converts oxidosqualene to beta-amyrin
CAS1	AT2G07050.1	Cycloartenol synthase; Converts oxidosqualene to cycloartenol. Involved in plastid biogenesis
UGT73B3	AT4G34131.1	UDP-glucosyl transferase 73B3; Involved in stress or defense responses
SQS1	AT4G34640.1	Squalene synthase
CYP93D1	AT5G06900.1	Member of CYP93D
AT3G32047	AT3G32047.1	Cytochrome P450 superfamily protein
CYP716A1	AT5G36110.1	Member of CYP716A

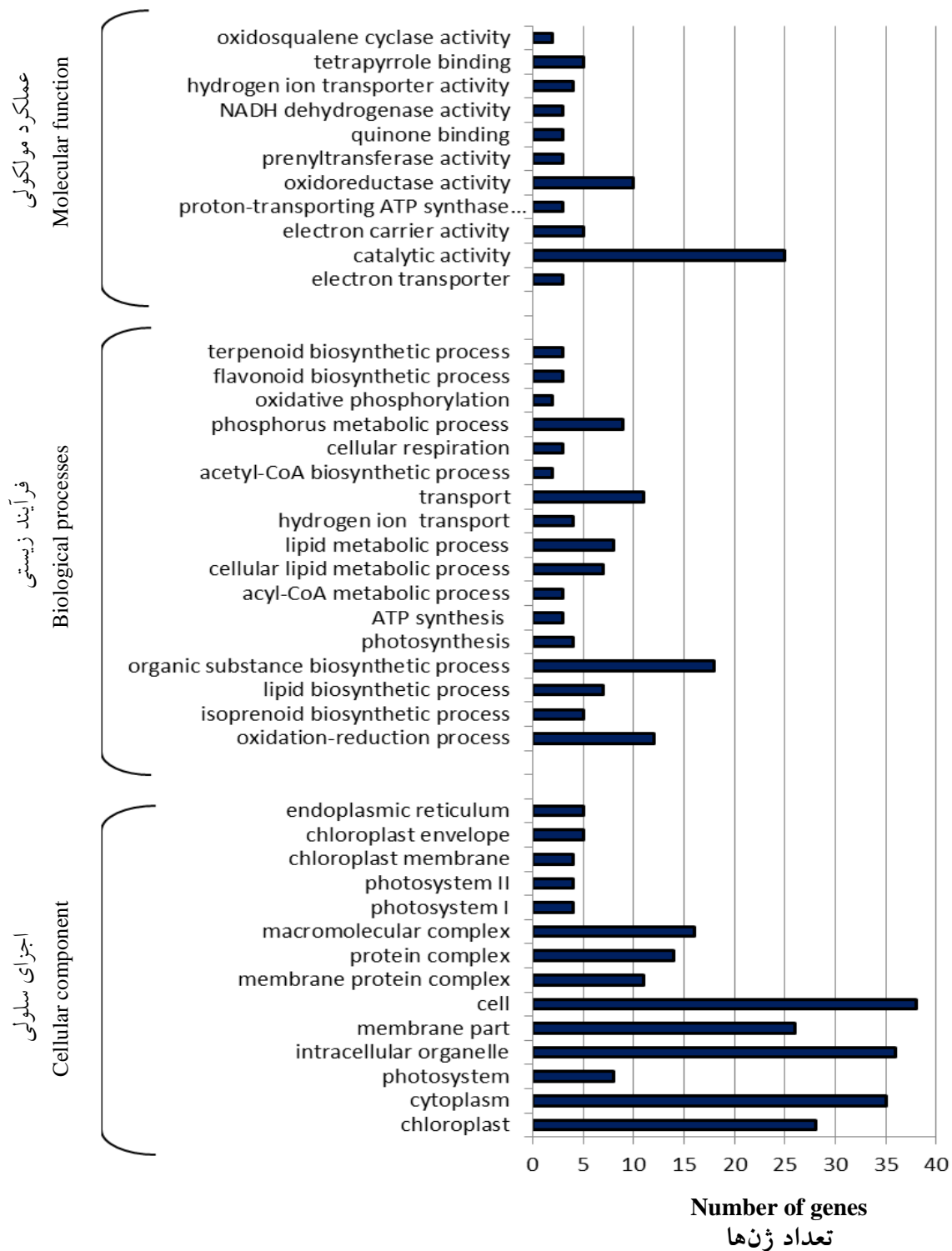


شکل ۲- شبکه ژنی ژنهای مهم مسیر سنتز گلیسیریزین در شیرین بیان در مقابل ژنوم *A. thaliana*  
 Figure 2. Gene network of genes involved in synthesis of glycyrrhizin against *A. thaliana* genome

نسبی هر چهار ژن، تفاوت معنی داری بین دو فصل پاییز و بهار در سطح ۱ درصد نشان دادند. نتایج نشان داد که فصل پاییز در مقایسه با فصل بهار باعث افزایش در سطح بیان نسبی برخی از ژن‌ها، در هر دو بافت شده است (شکل ۴ و ۵).

اعتبارسنجی بیان ژن‌ها با استفاده از واکنش **qRT-PCR** با توجه به آزمون آماری **t-test**، در بافت ریزوم تغییرات بیان نسبی ژن **BAS**، بین دو فصل پاییز و بهار تفاوت معنی داری در سطح ۵ درصد نشان داد اما برای ژن‌های **UGT** و **SQS**، **CYP88D6** تفاوت معنی داری بین این دو فصل مشاهده نشد. در بافت برگ، تغییرات بیان





شکل ۳- هستی‌شناسی ژن (Gene ontology) و فعالیت عملکردی توالی‌ها

Figure 3. Gene ontology and functional activity of sequences

جدول ۴- نتایج تعیین نقطه‌ی ایزوالکتریک (Pi)، وزن مولکولی (Mw) و مکان‌یابی سلولی توالی‌های پروتئینی مربوط به گونه

*G. glabra*

Table 4. Results of determination of isoelectric point (Pi), molecular weight (Mw) and cellular location of protein sequences related to *G. glabra*

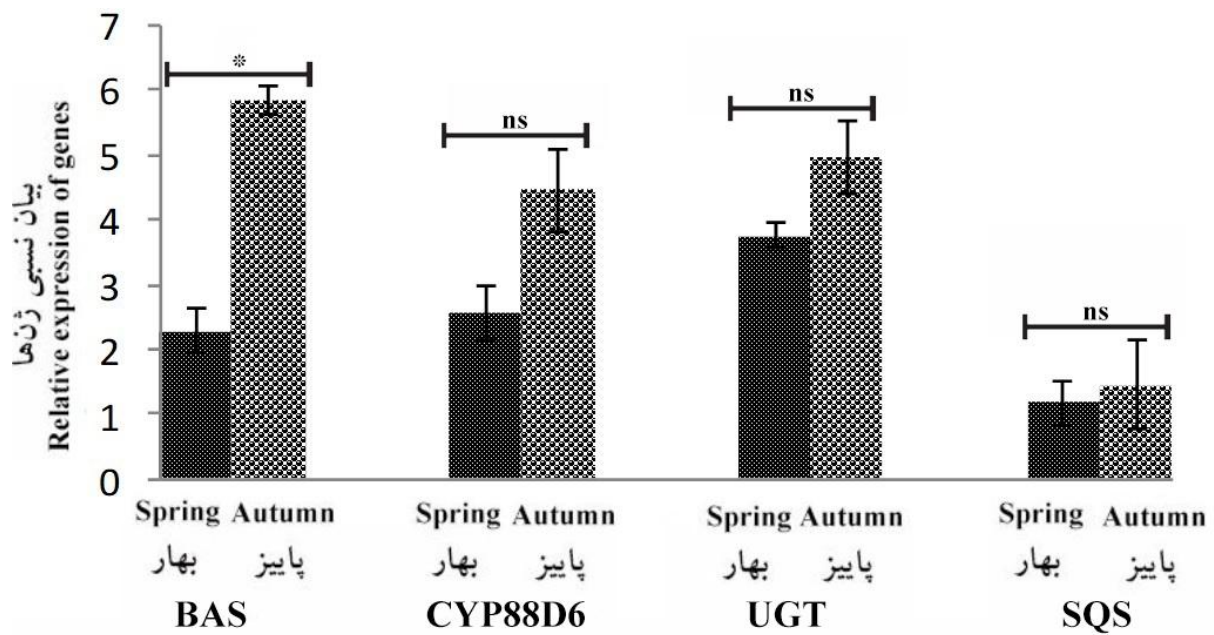
شماره دسترسی Accession number	نام پروتئین Protein name	مکان سلولی Cell location	نقطه ایزوالکتریک Pi	وزن مولکولی Mw
AMR98504.1	squalene synthase 1	Endoplasmic reticulum	8.18	47252.02
BAD08587.1	lupeol synthase	Endoplasmic reticulum	6.61	86688
BAA76902.1	cycloartenol synthase	Endoplasmic reticulum	6.35	86453.96
ARB43805.1	beta-amyrin synthase	Endoplasmic reticulum	6.19	87072.74
AKM97308.1	cytochrome P450 88D6	Endoplasmic reticulum	8.51	56266.28
YP_009027303.1	cytochrome b6/f complex	Extracellular	6.56	17472.83
YP_009027302.1	cytochrome b6	Extracellular	8.89	24126.47
AIN25418.1	beta-amyrin 24-hydroxylase CYP93E6	Plasma membrane	8.7	57905.47
BAA89815.1	beta-amyrin synthase	Endoplasmic reticulum	6.13	87516.3

جدول ۵- نتایج تعیین نقطه‌ی ایزوالکتریک (Pi)، وزن مولکولی (Mw) و مکان‌یابی سلولی توالی‌های پروتئینی مربوط به گونه

*G. uralensis*

Table 5. Results of determination of isoelectric point (Pi), molecular weight (Mw) and cellular location of protein sequences related to *G. uralensis*

شماره دسترسی Accession number	نام پروتئین Protein name	مکان سلولی Cell location	نقطه ایزوالکتریک Pi	وزن مولکولی Mw
ADG36724.1	squalene synthase 2	Endoplasmic reticulum	8.21	47084.66
ACV21067.1	beta-amyrin synthase	Endoplasmic reticulum	6.19	87026.69
BAW34647.1	cytochrome P450 716A179	Plasma membrane	8.06	55285.2
ADE88148.1	beta-amyrin synthase	Endoplasmic reticulum	6.19	87072.74
ACS66750.1	squalene synthase	Endoplasmic reticulum	8.5	47276.08
ADE18771.1	farnesyl-diphosphate synthase	Cytoplasmic	5.56	39359.12
ARB43808.1	beta-amyrin synthase	Endoplasmic reticulum	6.13	87130.78
AAW73054.1	cytochrome P450	Endoplasmic reticulum	4.7	4366.93
ANJ03631.1	UDP-glycosyltransferase	Plasma membrane	5.88	53868.89
BAL41371.1	lupeol synthase	Endoplasmic reticulum	6.61	86790.09
BAG68930.1	cytochrome P450 monooxygenase	Plasma membrane	8.78	58154.54
BAG68929.1	cytochrome P450 88D6	Endoplasmic reticulum	8.62	56372.49

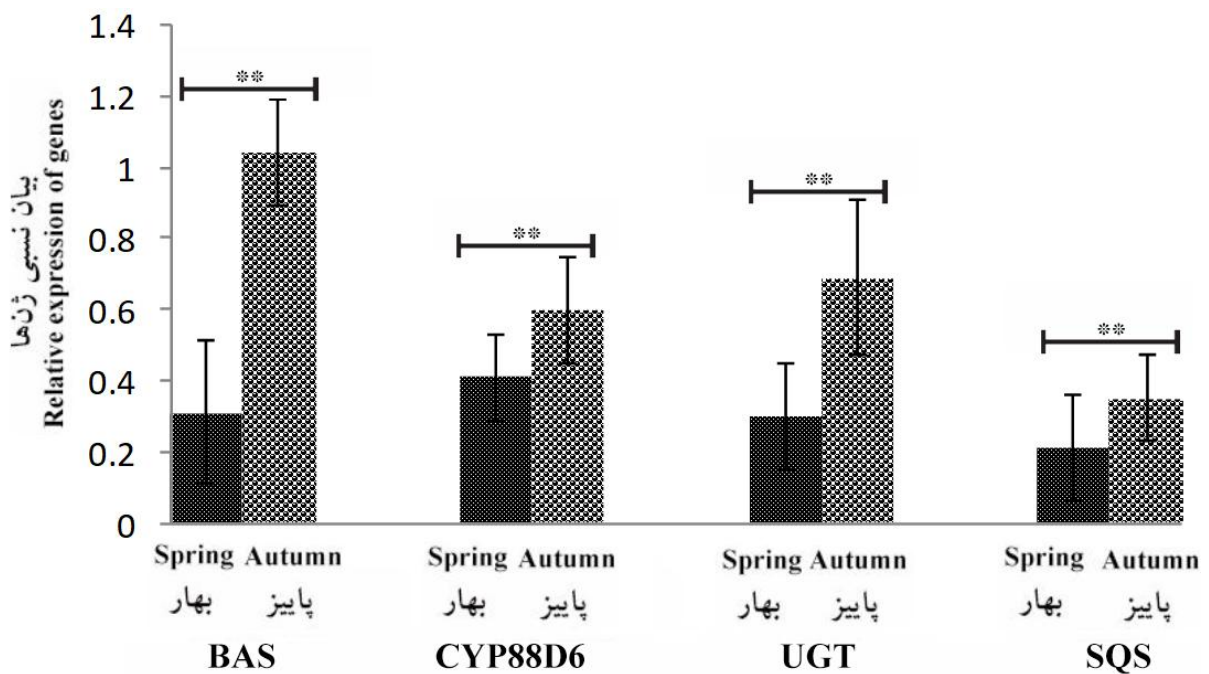


شکل ۴- مقایسه تغییرات بیان ژن‌ها بین دو فصل پاییز و بهار مربوط به بافت ریزوم

Figure 4. Comparison of gene expression changes between autumn and spring seasons related to rhizome tissue

<sup>ns</sup> و \* : به ترتیب غیرمعنی‌دار و معنی‌دار در سطح احتمال ۵ درصد

<sup>ns</sup> and \* : Non-significant and significant at 5% probability level, respectively



شکل ۵- مقایسه تغییرات بیان ژن‌ها بین دو فصل پاییز و بهار مربوط به بافت برگ

Figure 5. Comparison of gene expression changes between autumn and spring seasons related to leaves tissue

\*\* : معنی‌دار در سطح احتمال ۱ درصد

\*\* : Significant at 1% probability level

قرار دارد و در اکتبر و نوامبر بیان این ژن شروع به کاهش کرد (Liu and Liu, 2012).

تغییرات اقلیمی و فصول مختلف می‌تواند به‌طور اساسی بر روی میزان و ترکیبات عصاره گیاهان دارویی تأثیر بگذارد. تحقیقات نشان داده است که میزان مواد مؤثره و در راستای آن بیان ژن‌های تولید کننده آن‌ها در اندام‌ها و گیاهان هیچ‌گاه ثابت نیست و متناسب با مراحل رشد گیاه، شرایط محیطی و فصل‌های مختلف قابل تغییر است. میراندا و همکاران (Miranda et al., 2012) گزارش نمود که شرایط محیطی مانند ویژگی‌های اقلیمی و تغییرات فصلی می‌تواند ترکیبات ساپونینی گیاهان را نیز تحت تأثیر قرار دهد. سومجن و همکاران (Somjen et al., 2004) نشان دادند که تولید متابولیت‌های ثانویه تحت کنترل ژن‌ها بوده، ولی میزان تولید آن‌ها به‌طور قابل‌توجهی تحت تأثیر شرایط محیطی قرار می‌گیرند. در مجموع، بر این اساس نتایج به‌دست آمده از این پژوهش ارتباط عامل تأثیرگذار ژن‌ها و شرایط محیطی را به‌خوبی نشان می‌دهد.

این تفاوت در بیان نسبی ژن‌ها را می‌توان به روند تولید ماده مؤثره گلیسیریزین مرتبط دانست. تغییرات بیان نسبی ژن‌ها در هر دو بافت به‌صورت تنظیم مثبت بوده است و در بافت ریزوم در مقایسه با بافت برگ بیان بیشتری برای ژن‌ها دیده شد و این را می‌توان به محل تولید و تجمع ماده مؤثره گلیسیریزین در بافت ریزوم شیرین بیان دانست. این تفاوت‌ها را می‌توان به شرایط آب و هوایی و جذب رطوبت نیز مرتبط دانست. همچنین تولید مواد مؤثره در گیاهان دارویی معمولاً تحت تأثیر ژنوتیپ و عوامل محیطی قرار دارد (D'antuono et al., 2002). در مطالعه‌ای با هدف اثر بافت و زمان نمونه‌برداری بر روی بیان ژن بتا-آمرین سنتاز در *G. uralensis*، گزارش شد که بخش‌های زیرزمینی در مقایسه با بخش‌های هوایی و به‌ویژه در قسمت بالایی ریزوم‌ها نسبت به قسمت‌های پایینی، افزایش در بیان این ژن دیده شده است. از نظر زمان نمونه‌برداری، بیان ژن BAS را می‌توان به چهار مرحله تقسیم کرد. از دسامبر تا فوریه بیان این ژن تحت محدودیت قرار می‌گیرد و از مارس تا می بیان ژن صورت می‌گیرد. از ماه می تا سپتامبر، بیان این ژن در سطح بالایی

## References

- Aliakbari, M., Shamloo-Dashtpagerdi, R. and Ebrahimie, A. (2013). Bioinformatic analysis of saffron (*Crocus sativus* L.) stigma est sequences to determining functional genome orientation and gene network. *Journal of Saffron Agronomy and Technology*, **(1)1**: 40-45 (In Persian).
- Bayati Zadeh, J., Moradi Kor, Z. and Karimi Goftar, M. (2013). Licorice (*Glycyrrhiza glabra* Linn) as a valuable medicinal plant. *International Journal of Advanced Biological and Biomedical Research*, **1(10)**: 1281-1288 (In Persian).
- D'antuono, L.F., Moretti, A. and Lovato, A.F. (2002). Seed yield, yield components, oil content and essential oil content and composition of *Nigella sativa* l. and *Nigella damascene* l. *Industrial Crops and Products*, **15(1)**: 59-69.
- Fukai, T., Sheng, C.B., Horikoshi, T. and Nomura, T. (1996). Isoprenylated flavonoids from underground parts of *Glycyrrhiza glabra*. *Phytochemistry*, **43(5)**: 1119-1124.
- Hossein Panahi, Z.H., Maroufi, A. and Bahramnejad, B. (2016). Isolation and expression pattern of  $\beta$ -amyrin synthase in chicory (*Cichorium intybus* L.). *Genetics Novin Journal*, **3(11)**: 399-409 (In Persian).
- Hosseinzadeh, H. and Nassiri- Asl, M. (2015). Pharmacological effects of *Glycyrrhiza* Spp. and its bioactive constituents: update and review. *Phytotherapy Research*, **29(12)**: 1868-1886 (In Persian).
- Kim, D.W., Kim, R.N., Choi, S.H., Kim, D.W., Nam, S.H., Choi, H.S. and Kang, A. (2011). EST analysis predicts putatively causative genes underlying the pharmaceutical application of *Glycyrrhiza uralensis* fisch. *Plant Molecular Biology Reporter*, **29(4)**: 814-824.
- Li, Y., Luo, H.M., Sun, C., Song, J.Y., Sun, Y.Z., Wu, Q. and Chen, S.L. (2010). EST analysis reveals putative genes involved in glycyrrhizin biosynthesis. *Bmc Genomics*, **11(1)**: 268.

- Liu, Y., and Liu, C.S.** (2012). Study on the spatial and temporal expression of  $\beta$ -AS gene of *Glycyrrhiza uralensis*. *Chin Mater Med*, **35(4)**: 528-531.
- Miranda, M., Vega-Gálvez, A., Quispefuentes, I., Rodríguez, M.J., Maureira, H. and Martínez, E.A.** (2012). Nutritional aspects of six quinoa (*Chenopodium quinoa* Willd.) Ecotypes from three geographical areas of Chile. *Chilean Journal of Agricultural Research*, **72(2)**:175-181.
- Mochida, K., Sakurai, T., Seki, H., Yoshida, T., Takahagi, K., Sawai, S., and Saito, K.** (2017). Draft genome assembly and annotation of *Glycyrrhiza uralensis*, a medicinal legume. *The Plant Journal*, **89(2)**: 181-194.
- Nasrollahi, V., Mirzaie-Asl, A., Piri, K., Nazeri, S. and Mehrabi, R.** (2014). The effect of drought stress on the expression of key genes involved in the biosynthesis of triterpenoid saponins in licorice (*Glycyrrhiza glabra*). *Phytochemistry*, **103**: 32-37.
- Ogata, Y. and Suzuki, H.** (2011). Plant expressed sequence tags databases: practical uses and the improvement of their searches using network module analysis. *Plant Biotechnology*, **28(4)**: 351-360.
- Pandey, D.K. and Ayangla, N.W.** (2017). Biotechnological aspects of the production of natural sweetener glycyrrhizin from *Glycyrrhiza* sp. *Phytochemistry Reviews*, **17**: 1-34.
- Rebhun, J.F., Glynn, K.M. and Missler, S.R.** (2015). Identification of Glabridin as a bioactive compound in licorice (*Glycyrrhiza glabra* L.) extract that activates human peroxisome proliferator-activated receptor gamma (Ppar $\gamma$ ). *Fitoterapia*, **106**: 55-61.
- Schmittgen, T.D. and Livak, K.J.** (2008). Analyzing Real-time PCR data by the comparative c<sub>t</sub> method. *Nature Protocols*, **3(6)**: 1101-1108.
- Seki, H., Ohyama, K., Sawai, S., Mizutani, M., Ohnishi, T., Sudo, H. and Muranaka, T.** (2008). Licorice  $\beta$ -amyryn 11-oxidase, a cytochrome p450 with a key role in the biosynthesis of the triterpene sweetener glycyrrhizin. *Proceedings of the National Academy of Sciences*, **105(37)**: 14204-14209.
- Seki, H., Tamura, K. and Muranaka, T.** (2015). P450s and UGTs: key players in the structural diversity of triterpenoid saponins. *Plant and Cell Physiology*, **56(8)**: 1463-1471.
- Shirazi, Z., Aalami, A., Tohidfar, M. and Sohani, M.** (2017). Cloning, bioinformatics study and gene expression evaluation of squalene synthase 1 in Iranian native licorice. *Genetic Engineering and Biosafety Journal*, **1(6)**: 37-48 (In Persian).
- Somjen, D., Knoll, E. and Vaya, J.** (2004). Estrogen-like activity of licorice root constituents: glabridin and glabrene, in vascular tissues in vitro and in vivo. *Journal of Steroid Biochemical Molecular Biology*. **91**: 147-155.
- Tang, Z.H., Li, T., Tong, Y.G., Chen, X.J., Chen, X.P., Wang, Y.T. and Lu, J.J.** (2015). A systematic review of the anticancer properties of compounds isolated from licorice (gancao). *Planta Medica*, **81(18)**: 1670-1687.
- Wink, M.** (2010). Introduction: biochemistry, physiology and ecological functions of secondary metabolites. *Annual Plant Reviews*, **40**: 1-19.
- Zarabi, M., Fardoe, R. and Maleki-Tabrizi, N.** (2013). *Principle of Bioinformatics*. Ayiizh Publisher, Tehran, IR (In Persian).
- Zhao, H., Tang, Q., Mo, C., Bai, L., Tu, D. and Ma, X.** (2017). Cloning and characterization of squalene synthase and cycloartenol synthase from *Siraitia grosvenorii*. *Acta Pharmaceutica Sinica B*, **7(2)**: 215-222.

## Bioinformatics Study and Investigation of the Expression Pattern of Several Important Genes Involved in Glycyrrhizin Synthesis of *Glycyrrhiza glabra* L. in Autumn and Spring Seasons

Atefeh Khakpour<sup>1</sup>, Maryam Zolfaghari<sup>2,\*</sup> and Karim Sorkheh<sup>3</sup>

1- M.Sc. Student, Department of Horticultural Sciences, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran

2- Assistant Professor, Department of Horticultural Sciences, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran

3- Assistant Professor, Department of Agronomy and Plant Breeding, Shahid Chamran University of Ahvaz, Ahvaz, Iran

(Received: July 3, 2018 – Accepted: December 8, 2018)

### Abstract

Glycyrrhiza is one of the important medicinal plants that is in danger of extinction. Search for finding accessions that have a higher glycyrrhizic acid is very important in breeding programs. Functional genomics methods such as EST sequencing prepare the ability to identify consensus gene families among studied species and interpretation of the genome. In this research, 55960 EST sequences of two different species of this plant along with the protein sequences were analyzed in order to identify the molecular aspects and functional analysis of the genome and the gene network involved in the biosynthesis of glycyrrhizin. Also, in order to validation of results, relative expression of four important genes in the pathway of glycyrrhizin biosynthesis including squalene synthase (SQS),  $\beta$ -amyryn synthase (BAS),  $\beta$ -amyryn 11-oxidase (CYP88D6) and UDP-glucuronyl transferase (UGT) were evaluated. After trimming and qualitative evaluation the sequences, 6427 contig sequences and 30895 singleton (37322 unigene) were generated, which covered a total of 26884666 bp (7.06%) of the licorice genome. Genome functional activity showed that most genes play a role in the catalytic activity and cellular and metabolic processes in which these genes interact within cells and intracellular organelles. Locating this group of genes showed that the genes involved in glycyrrhizin biosynthesis pathway were localized in endoplasmic reticulum. Results of validation using qRT-PCR showed that in the autumn and in the rhizome tissue, the genes of BAS, CYP88D6, UGT and SQS were up-regulated. The results of this study can be valuable for genomic sequencing, functional groups, genetic diversity and functional genomics of this plant.

**Keywords:** Functional genome, Gene network, Licorice, Glycyrrhizin, qRT-PCR

---

\* Corresponding Author, E-mail: m.zolfaghari@scu.ac.ir