

مطالعه بیوانفورماتیکی ژن های منتخب دخیل در مسیر بیوسنتز فلاونوئیدها در زیره سبز

فرشته لطفی¹، سید محمد مهدی مرتضویان²، علی ایزدی دربندی²، حسین رامشینی²

1-بترتیب دانشجوی سابق بیوتکنولوژی کشاورزی

2-دانشیاران گروه علوم زراعی و اصلاح نباتات دانشکدهگان ابوریحان، دانشگاه تهران، پاکدشت

نویسنده مسول: mortazavian@ut.ac.ir

چکیده

زیره سبز گیاهی یکساله و علفی و متعلق به خانواده جعفری یا چتریان¹ است. در تحقیق حاضر پنج ژن منتخب موثر در مسیر بیوسنتز فلاونوئیدها که با روش NGS در زیره سبز شناسایی شده و بیشترین تغییر بیانی را داشتند مورد مطالعه بیوانفورماتیکی قرار گرفت. از این میان، دو قطعه ژنی DN32640 و DN1196 که دارای بیشترین فاصله تکاملی و از دو خانواده مختلف بودند، جهت شناسایی بیشتر، مورد بررسی قرار گرفتند. توالی DN1196 از خانواده GT6، یک گلیکوزیل ترانسفراز در مسیر بیوسنتز ترکیبات فنلی، و توالی DN32640 از خانواده BHLH، یک فاکتور رونویسی در مسیر بیوسنتز آنتوسیانین می باشد. پروتئین کد شده توالی DN1196، با شباهت بیش از 85 درصدی با پروتئینی از گیاه هویج، در مرحله پایانی سنتز فلاونول، با انتقال یک مولکول گلوکز، آن را گلیکوزیله می کند. در این فرایند دو پروتئین فاکتور رونویسی با این پروتئین در تعامل می باشند. ساختار ثانویه پیش بینی شده محصول این ژن، منطبق با توالی پروتئینی هویج مشخص شد. پروتئین کد شده توالی DN32640، با شباهت بیش از 90 درصدی با پروتئین هویج، یک فاکتور رونویسی از خانواده BHLH و جزئی از کمپلکس MBW است. این فاکتور رونویسی در تنظیم مراحل پایانی مسیر منتهی به بیوسنتز آنتوسیانین ها و تانن ها دخالت دارد. ساختار ثانویه محصول این ژن نیز بر اساس توالی پروتئینی شناسایی شده هویج مشخص شد. با توجه به اهمیت این دو ژن و برهمکنش محصول آن ها با سایر پروتئین ها می توان از آن دو در دستورزی های ژنتیکی جهت تولید بیشتر ترکیبات فنلی خصوصا آنتوسیانین در زیره سبز استفاده نمود.

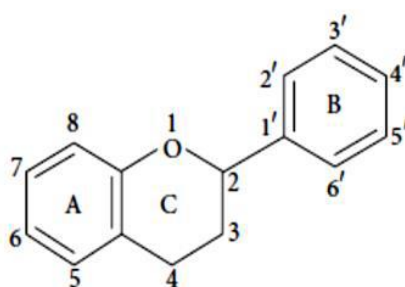
واژگان کلیدی: بیوانفورماتیک، پیش بینی، زیره سبز، ساختار ثانویه، فلاونوئیدها.

¹ Apiaceae

مقدمه

می‌کنند. گروه اصلی متابولیت‌های ثانویه، ترکیبات فنلی از جمله فلاونوئیدها هستند (Singh et al., 2021). فلاونوئیدها علاوه بر گیاهان تولید کننده این ترکیبات، برای حیواناتی که از این ترکیبات تغذیه می‌کنند نیز دارای فعالیت‌های مختلف بیولوژیکی هستند (Saito et al., 2013). تحقیقات نشان می‌دهد، مسیر اصلی بیوسنتز فلاونوئیدها در گیاهان مشترک و کاملاً حفظ شده است (Casati, 2012). فلاونوئیدها دسته‌ای از رنگیزه‌های گیاهی می‌باشند که محل تجمع آن‌ها در واکوئل سلول است. خاصیت آنتی‌اکسیدانی قوی، مهمترین اثر فلاونوئیدها، در رابطه با بیماری‌هایی مانند پیری و سرطان است که در اثر تنش اکسیداتیو ایجاد می‌شوند (درافشان و همکاران، 1398). فعالیت آنتی‌اکسیدانی فلاونوئیدها وابسته به ساختار آن‌هاست. هرگونه تغییر در ساختار آن‌ها از قبیل آلکیلاسیون، گلیکوزیلاسیون، محیط فیزیکی و شیمیایی ذخیره سازی آن‌ها، مانند نور، اکسیژن، pH، دما و تاریکی باعث ایجاد تغییرات قابل توجهی در میزان فعالیت آنتی‌اکسیدانی آن‌ها می‌شود (Casati, 2012) (شکل 1).

گیاهان دارویی، دسته‌ای از گیاهان هستند که به علت داشتن خواص دارویی از زمان‌های قدیم مورد استفاده انسان بوده و تعدادی از آن‌ها به دلیل داشتن طعم مطبوع به عنوان ادویه در غذا بکار می‌رفته‌اند. این خواص دارویی و ادویه‌ای به دلیل وجود ترکیبات مختلفی است که در طول واکنش‌های متابولیکی، در آن‌ها انباشته می‌شود (Ghasemi et al., 2019). متابولیت‌های ثانویه ترکیبات فعال بیولوژیکی هستند که از متابولیت‌های اولیه مشتق می‌شوند و نقشی در متابولیسم‌های اولیه فتوسنتزی و تنفسی نداشته و بیشتر خاصیت دارویی دارند (Alqethami and Aldhebiani, 2021). طیف وسیعی از تنش‌های محیطی زیستی و غیرزیستی مانند: دما، خشکی، شوری، قلیایی بودن، اشعه ماوراءبنفش، عوامل بیماری‌زای گیاهی مانند قارچ‌ها، باکتری‌ها و ویروس‌ها و نیز گیاه‌خواران، می‌توانند آسیب جدی به گیاهان وارد کنند (Jan et al., 2021). متابولیت‌های ثانویه مجموعه‌ای از واکنش‌های دفاعی مستقیم و غیرمستقیم فیزیکی و شیمیایی را در گیاهان ایجاد کرده و این دفاع را تسهیل



شکل 1- ساختار اصلی فلاونوئید (Amer, 2018)

فلاونوئیدها می‌توانند آزاد (آگلیکون) و یا مرتبط با قند باشند. مولکول آنتوسیانیدین پس از گلیکوزیله شدن به مولکول آنتوسیانین تبدیل می‌شود که این امر موجب تغییر در رنگ، جذب جانوران و به دنبال آن افزایش گرده افشانی می‌گردد. همچنین گلیکوزیلاسیون موجب

خاصیت آنتی‌اکسیدانی این مولکول‌ها به علت کمبود الکترون در ساختار مولکولی آن‌ها و استعداد واکنش با رادیکال‌های آزاد اکسیژن بوده و در خنثی نمودن تنش‌های اکسیداتیو نقش حفاظتی را در بدن ایفا می‌کنند (کاظم زاده بانه و همکاران، 1393).

افزایش حلالیت، توزیع و متابولیسم از طریق تسهیل انتقال از غشاء می‌شود (Silva, 2021). UFGT (UDP-glucose-flavonoid3-O-glycosyltransferase) آنزیمی است که مولکول‌های قند را به گروه C3-OH آنتوسیانیدین متصل کرده و آنتوسیانین‌های مسئول رنگ‌های مختلف تولید می‌کند. مولکول‌های آنتوسیانین به دلیل الگوهای مختلف گلیکوزیلاسیون از رنگ زرد تا بنفش را ایجاد می‌کنند. ژن تنظیم کننده MYB با اتصال به پروموتور ژن UFGT فعالیت آن را تنظیم می‌کند (Kaur et al., 2021).

ترکیب خاصی از عوامل رونویسی با تنظیم دقیق مکانی و زمانی مسیر بیوسنتزی، تولید و توزیع فلاونوئیدها را بر عهده دارند (Bogs et al., 2011) و با اتصال به نواحی پروموتور ژن‌های هدف سرعت رونویسی را از طریق RNA پلی‌مراز تغییر می‌دهند (Jan et al., 2021). تنظیم ژن‌های رمزگذار آنزیم‌هایی که به طور خاص در مراحل پایانی مسیر منتهی به بیوسنتز آنتوسیانین‌ها و تانن‌های متراکم هستند، بر عهده کمپلکس MBW است (Bogs et al., 2011) کمپلکس MBW نیز مانند مسیر فلاونوئید در تمام گیاهان به شدت حفظ شده است (Kaur et al., 2021). پروتئین‌های MYC یا BHLH‌ها که به 12 خانواده تقسیم می‌شوند؛ همگی در دامین حفظ شده و مناطق خاص خارج از دامین مشترک می‌باشند. این دامنه با نزدیک به 60 اسید آمینه دارای 19 اسید آمینه حفظ شده است که 5 اسید آمینه در ناحیه پایه، 5 اسید آمینه در مارپیچ اول، یک اسید آمینه در حلقه و 8 اسید آمینه آن در مارپیچ دوم قرار دارد (Bogs et al., 2011).

گیاه دارویی زیره^۲ متعلق به خانواده جعفری یا چتریان^۳ گیاهی یکساله و علفی است که در حالت دیپلوئید دارای فرمول $2n = 2x = 14$ و از قدیمی‌ترین و اقتصادی‌ترین گیاهان دارویی می‌باشد (سورنی و همکاران، 1391). مطالعات نشان می‌دهد بسیاری از

ادویه‌جات مورد استفاده در طب هندی، مانند زیره، زنجبیل و میخک به علت وجود مواد پلی‌فنل مانند فلاونوئیدها دارای خاصیت آنتی‌اکسیدان، ضد التهاب و جهش زاوی و سرطان می‌باشند (Das and Savage, 2012). در تحقیقی روی زیره سبز و چند گیاه دارویی دیگر، با توجه به طیف وسیع متابولیت‌های ثانویه از جمله فنول‌ها و فلاونوئیدها، آلکالوئیدها و ترپنوئیدهای آن‌ها، زیره سبز بالاترین پتانسیل مهار بیماری‌های باکتریایی (90٪) را از خود نشان داد (Sybiya, Vasantha Packiavathy et al., 2012). بنابراین شناسایی ژن‌های کلیدی و موثر در مسیر بیوسنتز فلاونوئیدها به عنوان ترکیبات دارویی چند وجهی، در گیاه زیره سبز ضروری به نظر می‌رسد. در تحقیقی که در دانشکده‌گان ابوریحان انجام شده است، نخستین پروفایل توالی رونوشت در گیاه زیره سبز به روش NGS گزارش شده است که بر اساس آن می‌توان ژن‌های دخیل در مسیر بیوسنتزی متابولیت‌های ثانویه را شناسایی کرده و در تحقیقات بعدی جهت اصلاح این گیاه ارزشمند مورد استفاده قرار داد. هدف از این تحقیق مطالعه بیوانفورماتیک ژن‌های منتخب درگیر در مسیر بیوسنتز ژن‌های فلاونوئیدی است که براساس مطالعات گذشته (صادقی و همکاران، 1396) بیان متفاوتی در شرایط تنش خشکی و طبیعی در زیره سبز نشان داده‌اند.

مواد و روش‌ها

جهت بررسی فلاونوئیدها در خانواده چتریان از بخش ژن سابت NCBI، 9 ژن، مربوط به هویج با نام علمی *Daucus carota subsp sativus*، به دست آمد. با استفاده از نرم‌افزار Megalign 5.00، هم‌ردیفی توالی FASTA هر یک از این ژن‌ها به روش CLUSTALW انجام و درخت فیلوژنتیکی رسم شد. پنج ژن در مسیر بیوسنتز فلاونوئیدها در زیره سبز از میان ژن‌های شناسایی شده در تحقیق گذشته با روش NGS که بیشترین تفاوت بیان را در شرایط تنش

³ Apiaceae

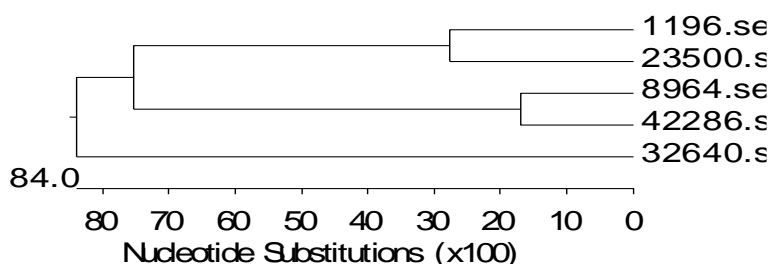
² با نام علمی *Cuminum Cyminum L*

نرم افزار (MEGALIGN، DNASTAR 5.00) پس از هم ردیفی توالی‌های منتخب، به روش

نتایج و بحث

از آنجایی که در تحقیق انجام شده در زیره سبز به روش NGS، یکی از مسیرها با بیشترین تغییر بیان در تنش خشکی مربوط به ژن‌های مسیر بیوسنتز فلاونوئیدها بود، بررسی ژن‌های منتخب در مسیر متابولیسم ترکیبات فنلی به خصوص فلاونوئیدها، دنبال شد. با استفاده از نرم افزار رویگو از میان 81 ژن با بیشترین تغییر بیان در مسیر بیوسنتز فلاونوئیدها در اکوتیپ‌های 3 (سیوند فارس) و 18 (کوهبان کرمان)، پنج قطعه ژن با بلندترین طول و بالاترین تغییر بیان انتخاب شدند. فاصله ژنتیکی این 5 توالی با نرم افزار مگالاین به روش ClustalW به دست آمد (شکل 2

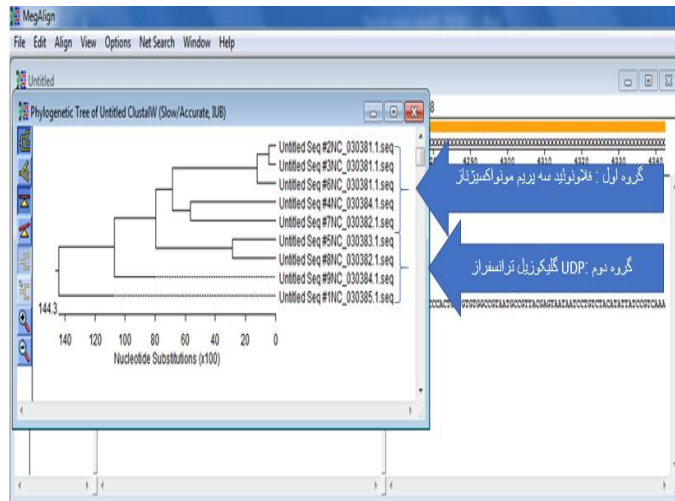
خشکی نسبت به شرایط آبیاری طبیعی داشتند، با استفاده از نرم‌افزار REVIGO جدا شدند. با استفاده از CLUSTALW، درخت فیلوژنتیکی رسم شد. پس از هم‌ترازی ژن‌های هویج با ژن‌های زیره و رسم درخت فیلوژنتیکی بین آن‌ها، دو توالی که دورترین فاصله را در این دو درخت داشتند (توالی DN32640 و DN1196) جهت انجام مطالعات بیوانفورماتیکی انتخاب شدند. جهت اطمینان از کد کنندگی پروتئین این دو توالی از نرم‌افزار ORFfinder در سایت NCBI استفاده شد. پس از هم ردیفی با توالی هویج به عنوان نزدیک‌ترین توالی در یونی پرات، ساختار ثانویه این دو توالی در سایت swissmodel و تعاملات پروتئینی آن‌ها در پایگاه داده STRING بر اساس توالی هویج بدست آمد.
(.



شکل 2-فاصله ژنتیکی پنج ژن شامل بلندترین توالی‌ها در مسیر بیوسنتز فلاونوئیدها در زیره سبز

خانواده فلاونوئید 3' مونواکسیژناز و گروه دوم از خانواده UGT بودند. گروه اول به علت تمایز خانوادگی با توالی‌های منتخب زیره حذف شدند و فقط از گروه دوم در بقیه مراحل استفاده شد (شکل 3).

نتیجه بررسی ژن‌های فلاونوئیدی در سایت NCBI در خانواده چتریان، 9 ژن از هویج بود که طبق رابطه خانوادگی در دو گروه قرار گرفتند. گروه اول از

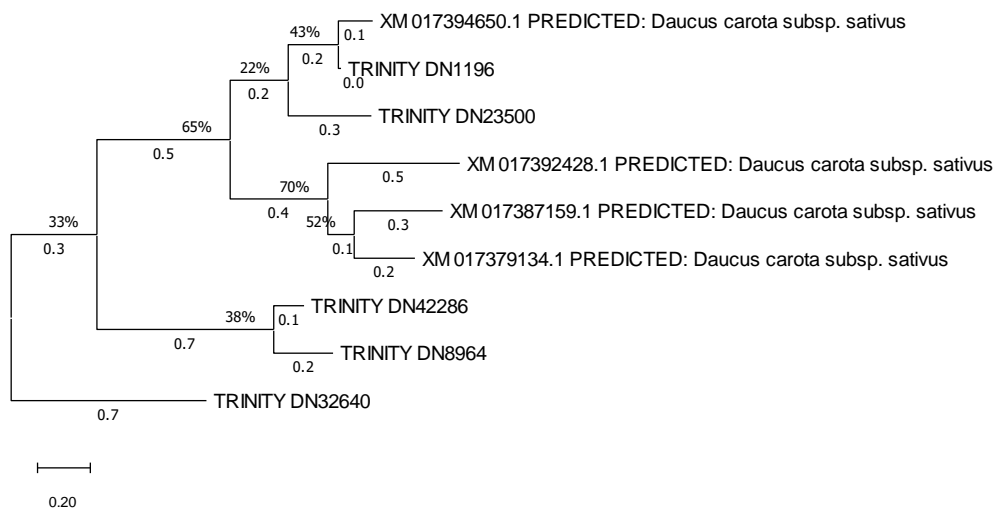


شکل 3- دسته بندی ژن های فلاونوئیدی در خانواده هویج

با نرم افزار Mega X فاصله ژنتیکی 4 ژن هویج و 5 ژن زیره سبز به دست آمد. توالی DN32640 بیشترین فاصله را با بقیه توالی ها داشت. از این توالی تجزیه و تحلیل ها با استفاده از مدل حداکثر درست‌نمایی ترکیبی انجام شد

جدول 1- تخمین واگرایی تکاملی بین توالی ها

	1	2	3	4	5	6	7	8	9
[50]									
[59]	1.446								
[34]	1.292	0.556							
[28]	1.415	0.873	0.870						
[40]	1.843	2.822	2.077	2.249					
[96]	0.141	1.203	1.098	1.234	1.895				
[86]	1.801	1.852	1.876	2.225	1.965	1.814			
[64]	1.933	2.159	2.011	2.360	1.815	1.871	0.339		
[00]	0.567	1.141	1.261	1.484	2.242	0.573	1.781	1.903	
substitutions per site									



شکل 4- تاریخچه تکاملی 4 ژن هویج و 5 ژن زیره سبز در مسیر بیوسنتز فلاونوئیدی با روش همسایه-پیوستن

هترودایمر هستند و در محل G-box به پروموتور ژن‌های هدف متصل می‌شوند. (Zhang et al., 2020). برای اطمینان از داشتن چهارچوب خوانش باز (ORF) جهت کد کردن پروتئین مربوطه این دو توالی، از نرم‌افزار ORFfinder در سایت NCBI استفاده شد (با شرایط: کمترین توالی 70 bp، استاندارد 1، Any sense codon، هر رمزی که اولین اسید آمینه یعنی متیونین (met) را کد می‌کند و Ignore nested ORFs). در توالی DN1196، 10 ORF وجود دارد. بلندترین آن‌ها با 1101 نوکلئوتید، پروتئینی با 366 اسید آمینه را کد می‌کند. در همین صفحه با اتصال به پایگاه یونی پرات، بلاست blastp توالی پروتئین کد شده این توالی، با گیاهان گلدار flowering plants (taxid:3398) انجام شد. بلاست با خانواده چتریان نتیجه‌ای نداشت. بالاترین امتیازها از 100 مورد، مربوط به گیاه جینسینگ و توت فرنگی با امتیاز بالای 400 بود. جهت بررسی پروتئین کد شده احتمالی، گیاه جینسینگ انتخاب شد. پروتئین احتمالی شناخته شده مربوطه، یک UDP-glucosyltransferase 103 با نام مختصر UGTPg103 است. میزان هم‌پوشانی در توالی پروتئینی جینسینگ با 472 اسید آمینه، با شباهت 59

DN32640 در هردو درخت فیلوژنتیکی توالی-های DN1196 و دورترین فاصله را دارند. فواصل تکاملی با استفاده از روش حداکثر درست‌نمایی مرکب⁴ محاسبه (جدول 1) و بر حسب واحد تعداد تعویض‌های باز⁵ در هر جایگاه است. تاریخچه تکاملی با استفاده از روش همسایه-پیوستن⁶ (شکل 4) استنباط شد. طول شاخه‌ها همان واحدهای فاصله‌های تکاملی است که برای استنتاج درخت فیلوژنتیک استفاده می‌شود. توالی DN32640 بیشترین فاصله را نسبت به بقیه داراست و این به دلیل عملکرد این ژن می‌باشد. با بررسی این ژن در یونی پرات مشخص شد؛ محصول این ژن یک عامل رونویسی است و پروتئینی تولید می‌کند که دارای دامین BHLH بوده و از خانواده BHLH می‌باشد. زانگ در نقش فاکتور رونویسی BHLH SmBHLH92 در تنظیم منفی بیوسنتز اسیدهای فنولیک در مریم گلی⁷ معتقد است، تنظیم بیوسنتزی این ترکیبات زیست فعال توسط مجموعه‌ای از فاکتورهای رونویسی (TFs) کنترل می‌شود. این فاکتورهای رونویسی دارای یک ناحیه بازی برای اتصال DNA در انتهای آمین و یک دامنه HLH در انتهای کربوکسیل برای تشکیل همودایمر یا

⁶ Neighbor-Joining Method

⁷ Salvia Miltiorrhiza

⁴ Maximum Composite Likelihood Method

⁵ Substitutions Per Site

UniProtKB/Swiss-Prot:A0A0K0PVL3.1
 مربوط به ژن 103 UDP-glycosyltransferase
 بصورت زیر است (شکل 5): (از شماره 107 تا 472 از
 توالی اصلی)

درصد و طول پروتئین زیره با 369 اسید آمینه، 99
 درصد می‌باشد، به طوری که رنگ میزان هم‌پوشانی، در
 بالاترین درجه یعنی قرمز می‌باشد. توالی این پروتئین
 در گیاه جینسینگ با شماره دسترسی

1	estrivglla dilfvdifdi adefnvptyv yspagtgflg lafhqltln dkkqdvtefr
61	nsdtellvps fanpvpaefs psiflekdgr hdvllslyrr creakgiavn tfeeleyai
121	nslrmdsmip piypvgpiln lngegqnsde aavilgwldd qppssvvlc fgsfgsfpen
181	qvkeiamgle rsghrflwsl rpcisegett lqlkysnlel pagfldrtsc vrkvigwapq
241	mavlaheavg gfvshcgwns vlesvwygmp vatwpmyeq qlnafemvke lglaveievd
301	yrneynksdf ivkadeietk ikklmmdgkn skirkkvkem keksrvamse ngssytslak
361	lfeeim

شکل 5- توالی، مربوط به UDP-glycosyltransferases, UDPGT با کد A0A0K0PVL3.1

glycosyltransferase است که یک مولکول قند را به
 یک فلاونول منتقل می‌کند.

WAPQLAVLSHRAVGGFVSHCGWN
 SILESVWFGVPMATWPSYAEQ

شکل 6- دامین توالی پروتئینی 3-UDP-glucose flavonoid
 O-glucosyltransferase در هویج

توالی دامین در هویج به علت تفاوت سوبسترا کمی
 با توالی این دامین در جینسینگ متفاوت است.

WapQmaVLahavgGFVSHCGwnSyleSVwyg
 mPvatwPmygEQ

شکل 7- توالی دامین پروتئین UDP-glucose flavonoid
 3-O-glucosyltransferase در جینسینگ

برای به دست آوردن ساختار ثانویه یک پروتئین
 باید از پروتئین‌هایی که ساختار ثانویه آن‌ها مشخص
 شده و بیشترین تشابه را با توالی پروتئینی ناشناخته
 دارند استفاده کرد. جهت به دست آوردن ساختار ثانویه
 توالی پروتئینی احتمالی DN1196، از سایت NCBI
 استفاده شد. به دلیل تشابه بیش از 80 درصد توالی
 پروتئینی DN1196 با هویج، از توالی هویج جهت به
 دست آوردن ساختار ثانویه پروتئین احتمالی زیره سبز
 استفاده شد. در قسمت BLASTP، در بخش DATA
 BASE، گزینه pdb انتخاب شد، تا جستجو فقط در
 میان پروتئین‌های گیاهی که ساختار سه بعدی آن‌ها به

این پروتئین در سایت
<https://prosite.expasy.org/cgi-bin/prosite>
 دارای 366 اسید آمینه و یک دامین از شماره 237 تا
 280 می‌باشد که مربوط به UDPGT، UDP-
 glycosyltransferases با کد 57PS003 می‌باشد.
 توالی دامین مورد نظر به صورت زیر می‌باشد:

WapQmaVLahavgGFVSHCGwnSvleSVwyg
 mPvatwPmygEQ

بنابراین می‌توان احتمال داد که توالی DN1196 در
 زیره سبز یک توالی کد کننده پروتئین با دامین مذکور
 باشد. توالی DN1196 در یونی پرات با 58/8 درصد
 تشابه، در جینسینگ آسیایی یا جینسینگ چینی
 (*panax ginseng*) یکی از گونه‌های یازده‌گانه سرده
 جینسینگ، هم وجود دارد. محصول این ژن در
 جینسینگ UDP- α -D-glucose است. این آنزیم در
 جینسینگ در واکنشی یک گلوکز به مولکول (20S)-
 protopanaxatriol اضافه کرده و آن را به (20S)-
 ginsenoside Rh1 تبدیل می‌کند که یک ترپنوئید و
 متابولیت ثانویه می‌باشد. توالی DN1196 در یونی پرات
 با ژن هویج (*Daucus carota* subsp.sativus)
 85/8 درصد تشابه دارد و یک UDP-

فیتلاکا دکاندر⁸ با نام سرخاب کولی نام یک گونه از سرده سرخاب کولی از گونه *P. Americana* می باشد (شکل 8).

دست آمده است انجام شود. پس از BLASTP، چهارگزینه بدست آمد. چهارمین گزینه با 33/61 درصد تشابه و 97٪ همپوشانی به دلیل داشتن یک ساختار انتخاب شد. این پروتئین با 489 اسید آمینه متعلق به

```

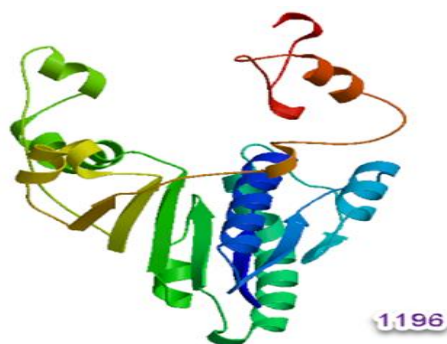
Model_02 MKASITFPAPVRGHLISMVLAKLILSRNQSISVTITFDKFPYDTGTVNSVIESDETNE 60
template_upload.1.A--AFELVYFSPSPG(GHLIFLVEFAKLVSR)F--HFVYSSTLPTTAQPTCAQTLLDGLPS 56
Model_02 HPRIALTEVPISSSEAYKFE SYVTMIPDFIDSHVTFVKNVAVNQPDRAVAALVDVFCG 119
template_upload.1.AVSRHLLPTIV(CARLPVA(EVYIISHTH)HSCSIRAAAGSIA---QAVVALD)LFGT 113
Model_02 SMIDMAKQDNIISYVYPTSGASVGLVLFHIGGIDVYENEDIYFYKDSDAEISVCCFN--- 176
template_upload.1.ACYTVARLGIIFYVYPTSTAMCLLFLHLLPLLEDDVYSCCECRDMP---EHLVLPGCVFPH 170
Model_02 -----N-----AVPAKVLVETICIDKGGQSLV-----ADDKVPISVY 208
template_upload.1.AGDFVYDPRDQRQDQAYHVLIDHYK(V)AEVSTNTF(CD)CGAKITIQD)EDPNVPLVY 230
Model_02 VGPNIIMLTSARQNSDPEVAI(LKKN)DDGKPVSSVVFICFGSFGSFTVQVGFGTREMKNAELI 268
template_upload.1.AVGPIIQSGL-----DDDSI(LKWLDRQPSGGSV)VFVFGSGGILSCEQ----- 272
Model_02 FVPAPAMGHLISMVELAKLILSRDESLSVTICIKFPPYDMGVTSYVESLSNNPTPRLSTL 328
template_upload.1.A-----
Model_02 EIPPSDSESYKQSHTTVFYNFMKSHVTINVRDQVVAMTRADQSTRIAAAFILDIFAVSMID 388
template_upload.1.A-----
Model_02 MAKELNIPSYVYSTSGADYLGILLHLQAMKDYENKDMGEYKKNPDAELSVPYFRNQVPAKV 448
template_upload.1.A-----
Model_02 LPAMCLDQGGSEITLLWGRKFRKTLGLIVNTFVELEAYVVESLMADDRIPFVYSVGPNLN 508
template_upload.1.A-----
Model_02 LSSRGQDSDEAADILKWLDEKPLSSVVFLCFGTFGGFPEDQVKEIALALEHSGHPFLWSL 568
template_upload.1.A-----
Model_02 RPVKKEGHLISKYTNLDDILPPGFLERTSGTKIIGWAPQLAVLSHRAVGGFVSHCGWNSI 628
template_upload.1.A-----
Model_02 LESVWFGVPMATWPSYAEQQLNAFEMVSELGLAVDICMDYAHHGFSKCELIVEAEIISK 688
template_upload.1.A-----
Model_02 INEVMKDGSDIRKRVAEIKGKSRAA VAENGSSVYVNLGRLIEDILAT 734
template_upload.1.A-----

```

شکل 8- هم ترازوی هویج اهلی با توالی سرخاب کولی در swissmodel.

پروتئینی هویج کپی و ساختار دانلود شده آپلود شد (شکل 9).

فلشها صفحات بتا و بیضیها مارپیچ آلفا هستند. ساختار پیشنهادی با روش x-ray و با شماره JEN6 PDB ID: در سایت swissmodel توالی



شکل 9- ساختار ثانویه توالی پروتئینی احتمالی DN1196 براساس ساختار هویج تعداد مارپیچ آلفا: 16 زیگزاگ صورتی، تعداد صفحات بتا: 13 فلش سبز

تا پروتئینهای در تعامل با این توالی مشاهده شود. ولی پروتئینی در تعامل با این توالی به دست نیامد. جستجو در گیاهان انجام شد. توالی توت فرنگی وحشی نام یک

توالی پروتئینی هویج در پایگاه داده STRING و در خانواده چتریان مورد بررسی قرار گرفت. اولین نتیجه به دست آمده در هویج با 100 درصد تشابه انتخاب شد

گونه از سرده توت‌فرنگی با نام علمی *Fragaria vesca* با 50٪ درصد تشابه انتخاب شد. در توت‌فرنگی (*Fragaria vesca*) با شماره دسترسی 918.XP_004303956.175 تعامل با دو پروتئین فاکتور رونویسی با کد زیر گزارش شده است:

Transcription factor HY5-like XP_004291469.1
Transcription factor HY5-like 046.17XP_00429

در گزارش صادقی این توالی از خانواده GT6 در توت‌فرنگی گزارش شده است. بیان این ژن در تنش خشکی در هر دو 3 (سیوند فارس) و 18 (کوهبان کرمان) دو برابر حالت معمول بوده است (داریوش صادقی و همکاران، 1396).

توالی DN32640 دارای ORF 9 است که بلندترین آن‌ها ORF7 بر روی زنجیره منفی و فریم 3 می‌باشد. شروع از نوکلئوتید 1069 و پایان آن 305 است و پروتئینی با 254 اسید آمینه را تولید می‌کند. بلاست BLASTP توالی پروتئینی مورد نظر با گیاهان خانواده چتریان نتیجه‌ای نداشت ولی با گیاهان گلدار 86 نتیجه داشت. بالاترین امتیاز مربوط به کیوی از خانواده Actinidiaceae، با 55/16٪ تشابه بود. دو توالی اول دارای بیشترین امتیاز هم‌پوشانی بودند که هر دو در گیاهان گلدار دولپه‌ای قرار دارند. توالی اول با امتیاز 264 و هم‌پوشانی 98 درصدی، پروتئینی در کیوی، دارای دامین BHLH42 بوده و یک عامل رونویسی می‌باشد. توالی دوم با 251 امتیاز و هم‌پوشانی 98 درصدی، ژنی در نخود فرنگی، دارای دامین BHLH بوده و طبق گزارش صادقی یک عامل رونویسی است (داریوش صادقی و همکاران، 1396). پروتئین‌های BHLH دارای دامین BHLH هستند. نقش این پروتئین‌ها در تمام موجودات از مخمر تا انسان، تنظیم رونویسی می‌باشد (سایت NCBI). طول توالی مورد نظر در کیوی 689 با شماره دسترسی A0A2R6QE26.1 و در نخود فرنگی 652 با شماره دسترسی E3SXU4.1 می‌باشد. جهت بررسی پروتئین

کد شده احتمالی، گیاه کیوی انتخاب شد. این پروتئین مربوط به فاکتور رونویسی¹¹ BHLH42 به بیان دیگر basic helix-loop-helix protein 42 در گیاه کیوی chinensis¹² با نام مختصر AcBHLH42 می‌باشد. طول این پروتئین 689 اسید آمینه است که از اسید آمینه 414 تا 689، با توالی زیره سبز هم ترازی دارد. در این هم ترازی 155 اسید آمینه (55٪) شباهت وجود دارد. این پروتئین دارای یک دامین basic helix-loop-helix (BHLH) جهت اتصال به DNA بوده و به عنوان یک فاکتور رونویسی عمل می‌کند. این دامنه در پروتئین TT8 در آرابی‌دوپسیس تالیانا دیده شده است. خانواده BHLH در آرابی‌دوپسیس¹³ شامل چندین فاکتور رونویسی مانند TT8، EGL1 و GL3 است. TT8 که AtBHLH42 یا EN32 نیز نامیده می‌شود، در کنترل رنگدانه فلاونوئید نقش دارد و نقش کلیدی در تنظیم لکوانتوسیانیدین ردوکتاز (BANYULS) و دی هیدروفلاونونول-4-ردوکتاز (DFR) دارد. توالی مورد نظر در آرابی‌دوپسیس به صورت شکل 10 می‌باشد. دامین basic helix-loop-helix (BHLH) از اسید آمینه 87-161 به صورت شکل 11 است. موتیف اتصال به DNA (DNA binding) به صورت شکل 12 است: (102..103,106..107,109..110,114,122..123, 126..127,130,132..133,136..137)

توالی این پروتئین در کیوی با 276 اسید آمینه با نام BHLH42 و شماره 000044807 با زیره سبز هم‌پوشانی دارد. اسید آمینه 70-140 دامین (BHLH) است. موتیف پایه¹⁴ در بخش 70-83 می‌باشد و اسید آمینه‌های (105..104, 87, 82..83, 79..80, 75..76) بخش‌های اتصال به DNA¹⁵ می‌باشند. اسید آمینه‌های 84..119 بخش HLH

¹⁴ Basic Motif

¹⁵ Nucleotide Binding

¹¹ Transcription Factor

¹² *Actinidia Chinensis* Var

¹³ *Arabidopsis thaliana*

(85..86,89,92..93,96,105..106,109..110,113,116,119..120,123,126..127,129..130,133,136..137,140)

motif است که در این منطقه اسید آمینه‌های زیر محل تشکیل دایمر و یا رابط دایمر است^{۱۶}. (شکل 14)

1	mevsrfgnak araragarrs afqrwdyarl tkpphtkqgs hqqqmlktsl lsiprlftvr
61	rgidehdeql sgnamesaet adghiqrqdg rhmmaerkrr eklnrdrvtl rslvpyvskq
121	dkvsllgdai dfikdlqrqv eelesrrkis enpskprvei tvennravfe isspwrqdll
181	iailetvgt hmqv edvaak vskdtfkatl kakvtssgid ddddedkks ikqiretlls
241	vvngnnsptf eklc
	شکل 10

1	mevsrfgnak araragarrs afqrwdyarl tkpphtkqgs hqqqmlktsl lsiprlftvr
61	rgidehdeql sgnamesaet adghiqrqdg rhmmaerkrr eklnrdrvtl rslvpyvskq
121	dkvsllgdai dfikdlqrqv eelesrrkis enpskprvei tvennravfe isspwrqdll
181	iailetvgt hmqv edvaak vskdtfkatl kakvtssgid ddddedkks ikqiretlls
241	vvngnnsptf eklc
	شکل 11

1	mevsrfgnak araragarrs afqrwdyarl tkpphtkqgs hqqqmlktsl lsiprlftvr
61	rgidehdeql sgnamesaet adghiqrqdg rhmmaerkrr eklnrdrvtl rslvpyvskq
121	dkvsllgdai dfikdlqrqv eelesrrkis enpskprvei tvennravfe isspwrqdll
181	iailetvgt hmqv edvaak vskdtfkatl kakvtssgid ddddedkks ikqiretlls
241	vvngnnsptf eklc
	شکل 12

1	mevsrfgnak araragarrs afqrwdyarl tkpphtkqgs hqqqmlktsl lsiprlftvr
61	rgidehdeql sgnamesaet adghiqrqdg rhmmaerkrr eklnrdrvtl rslvpyvskq
121	dkvsllgdai dfikdlqrqv eelesrrkis enpskprvei tvennravfe isspwrqdll
181	iailetvgt hmqv edvaak vskdtfkatl kakvtssgid ddddedkks ikqiretlls
241	vvngnnsptf eklc
	شکل 13

1	yssqsfkw plrpsdhhhq adatsqcllk yilftvpflh skyrdnnsps ksataadsas
61	rfrkptpqde lsanhvlaer rreklnerf iilrslvpfv tkmdkasilg dtieyvkqlr
121	kniqdleara rqieidqrsr ssgdpqrga kaatdkrklr iveasggakg kavdsavat
181	ppppappqv agvgvqvqs iiesdalvel qchreglll dvmvvlrdhr vevtavqssl
241	tngvfvaelr akvkenvngk kpsivevkra ihqiip
	شکل 14

e lsanhvlaer rreklnerf iilrslvpfv tkmdkasilg dtieyvkqlr kniqdleara rqieidqrsr

بنابراین توالی DN32640 یک توالی کد کننده پروتئین در زیره سبز با دامین احتمالی زیر می‌باشد.

e lsanhvlaer rreklnrf iilrslvpfv tkmdkasilg
dtieyvkqlr kniqdleara rjeidqrsr
نقش‌های این ژن در مسیر پیام رسانی با واسطه
اسید جاسمونیک- تنظیم مثبت فرآیند بیوسنتزی
آنتوسیانین- تنظیم فرآیند بیوسنتزی ماکرومولکول
سلولی- تنظیم فرآیند بیوسنتزی پروآنتوسیانیدین-
توسعه پوشش بذر و تمایز تریکوم می‌باشد. پس از
بلاست توالی DN32640 در یونی پرات نزدیک ترین
توالی، توالی هویج اهلی بود. از این توالی پروتئینی جهت
به دست آوردن ساختار ثانویه توالی DN32640 در
سایت NCBI استفاده شد. با توالی هویج 5 توالی از
آراییدوپسیس نتیجه داد. طول تمام این توالی‌ها زیر
200 اسید آمینه بود. بنابراین فقط بخش کوچکی از
توالی هویج قابل پیش بینی ساختار ثانویه قرار گرفت
(شکل 15).

در یونی پرات توالی DN32640 با 250 توالی دیگر
ازدولپه‌ای‌ها 48.3٪ تا 91.4٪ یکسانی را نشان داد.
تمامی این موارد پروتئین‌هایی هستند که دارای دامین
BHLH هستند. دو توالی اول که بیش از 90٪ شباهت
را دارا هستند، از گروه هویج خوراکی و هویج وحشی
می‌باشند که هردو از خانواده چتریان هستند. اولین
پروتئین، در هویج خوراکی *Daucus carota subsp.*
(Carrot) sativus با کد A0A162B508 است که
91.4٪ یکسانی را نشان می‌دهد. طول این توالی در
هویج 1161 و در زیره 1071 می‌باشد. بخش‌های
یکسان از اسید آمینه 308 تا 634 می‌باشد که در این
میان 96.9٪ تشابه کامل وجود دارد. دامین مشترک
زیره سبز و کیوی در زیر آمده ، بخش زرد دامین
مشترک کیوی، زیره سبز و هویج است.

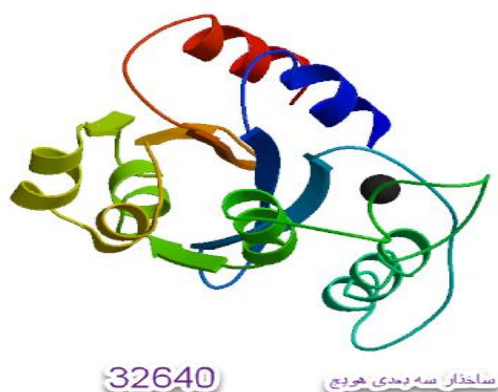
```

Model_01 MASPPPYKSNAH QNMLQAAYQAV--NNTYSLFQQLS-----DEKQRLVLRGDDGPNNGA 51
template_upload.1.A-----QQLQALITSSAGE NWTYALFWQISHRY>SSTGENTVILRWGDDGYVKG- 50
Model_01 ITRKMTVNSVEVSTEEASLERSQQLRELYDSISTGDLNPNRRRPAALSPEDITGSEWFWYL 111
template_upload.1.A-----EEDTAEQPEHRKEVIREL-NSL>SGD-----EEVITTEWFFIL 85
Model_01 MCVSFSFPLGVGLPGKAYVPROHLVYKGANFASRKHFTRAILN-----KTVVCTPDLDSY 166
template_upload.1.A VSMTCSEYAGVGLPGRSFI NSRVTWLTGSGALITGSGCRAAGGQYGLKIMVCTIAYNGV 145
Model_01 VETITTDKVEEKISFIEQVKSRIINCNNQATHPPKPALSEHSTSNLISPSRLVEMDDDDDE 226
template_upload.1.A VELIG>EVISQSSDGMKRVNLIK-CCNNQATHPPKPALSEHSTSNLISPSRLVEMDDDDDE 168
Model_01 NDDDDDDDDAGEEDDDDEGDGIDVEAEAAARTPTPFPGSVDPNRNVVGDIEGEPRIQLVMLHMH 286
template_upload.1.A-----
Model_01 SNDNIHLGSPESGSGNNLSDSFHLLKVSQPGIPIDPYVPGSKWIHAPLTAARITAEDELLEE 346
template_upload.1.A-----
Model_01 HNHYSQTITITILHNRNFSGRSSGTSSAVYSIHSPQSSFSICTAAARTSRSHRHSNRISQYL 406
template_upload.1.A-----
Model_01 LKYILFTVPYLHKSRLDNTAEDSVMLQLKKSASQEELSANHVLAERRRRREKLNRFVILR 466
template_upload.1.A-----
Model_01 SLVPFVTKMDKASILGDTIEYVKQLAKKIHELEAREHVNEDADCVDQRLGNRGDSRRDK 526
template_upload.1.A-----
Model_01 RKMRVVEQGGRRVVKAAEVQOETTMEVSIIEENDGLVEIKCPYKEGLLDVMMKLRFRFME 586
template_upload.1.A-----
Model_01 ITAAHSSSSNGGFVAELRAKVKETVNGKKASITEVKKAMKNIIPSHF 634
template_upload.1.A-----

```

شکل 15- هم‌ترازی هویج اهلی با توالی RQW آراییدوپسیس در سوئیس مدل.
فلش‌ها صفحات بتا و بیضی‌ها مارپیچ آلفا هستند

در ساختار پیش بینی شده هویج بر اساس توالی
ARQW در آراییدوپسیس 6 صفحات و 7 مارپیچ آلفا
وجود دارد (شکل 16).



شکل 16- ساختار ثانویه توالی پروتئینی احتمالی DN32640 زیره سبز براساس ساختار هویج 6 بتا شیت و 7 آلفا هلیکس

متعلق به دو خانواده متفاوت پروتئینی هستند. در بررسی‌های بیوانفورماتیکی مشخص شد، هردو توالی DN32640 و DN1196 دارای چهارچوب خوانش باز ORF هستند. بلندترین خوانش توالی DN1196 کد کننده پروتئینی با 366 اسید آمینه و توالی DN32640 کد کننده پروتئینی با 354 اسید آمینه می‌باشد. پروتئین احتمالی توالی DN1196 از خانواده GT6 در هویج با تشابه بیش از 85 درصد، انتقال دهنده یک مولکول گلوکز به یک مولکول فلاونول و گلیکوزیله کننده آن می‌باشد. ساختار ثانویه این پروتئین براساس توالی پروتئینی هویج و ساختار ثانویه مدل گیاه سرخاب‌کولی دارای 10 صفحه بتا و 16 مارپیچ آلفا می‌باشد. این توالی در زیره سبز با بیش از 50 درصد تشابه با توالی پروتئینی توت فرنگی احتمالاً با دو پروتئین فاکتور رونویسی HY5 در تعامل می‌باشد. بیان این ژن در زیره سبز در تنش خشکی در اکوتیپ‌های 3 (سیوند فارس) و 18 (کوهبان کرمان) دو برابر حالت معمول بوده است؛ بنابراین می‌تواند در آزمایشات بعدی جهت افزایش تولید گلیکوزیل فلاونول که مهمترین عضو خانواده فلاونوئیدها می‌باشد مورد بررسی قرارگیرد. توالی DN32640 کد کننده پروتئینی احتمالی با دامین بسیار حفظ شده BHLH، دارای یک موتیف بازی جهت اتصال به DNA و یک موتیف HLH جهت تشکیل دایمر پروتئینی متصل شونده به پروموتور DNA هدف به منظور تنظیم رونویسی از ژن هدف در مسیر

جهت شناسایی پروتئین هایی که در تعامل با این پروتئین در هویج هستند در سایت STRING توالی پروتئینی هویج وارد شد. در میان خانواده چتریان جستجو انجام شد. نتیجه تماماً در هویج به دست آمد. اولین گزینه با 100 درصد تطابق همان توالی داده شده بود و بقیه زیر 35 درصد تطابق داشتند. گزینه اول انتخاب شد. طبق تحقیق صادقی بیان ژن DN32640 در تنش خشکی در اکوتیپ 3 (سیوند فارس) نصف و در اکوتیپ 18 (کوهبان کرمان) دو برابر شده است که نشان از اهمیت این ژن در شرایط تنش خشکی در افزایش یا کاهش تولید آنتوسیانین، در بین اکوتیپ‌های مختلف می‌باشد (داریوش صادقی و همکاران، 1396).

نتیجه گیری

در تحقیق حاضر دو توالی DN1196 و DN32640 از پنج ژن منتخب پیشنهادی در زیره سبز در مسیر بیوسنتز ترکیبات فنلی و فلاونوئیدی که بیشترین تغییر بیان را در تنش خشکی در تحقیقات قبلی داشتند، مورد بررسی بیوانفورماتیکی قرار گرفت و پس از بررسی همردیفی هردو توالی با توالی هویج از خانواده چتریان بیش از 85 درصد تشابه نشان دادند. بنابراین از آنجایی که اطلاعات ژنتیکی کمی در باره زیره سبز وجود داشت، جهت شناسایی پروتئین احتمالی این ژن‌ها از توالی هویج استفاده شد. این دو توالی که بیشترین فاصله ژنتیکی را در میان 5 توالی منتخب زیره سبز و توالی ژن‌های فلاونوئیدی در هویج داشتند،

MBW فعالیت می‌کنند، تعداد بالای پروتئین‌های در تعامل توجیه می‌شود. از این 10 پروتئین، یک فاکتور TT1 از خانواده WD40، یک فاکتور ELF6 از خانواده ADOX1، 3 فاکتور از خانواده MYB، یک لیگاز VI، یک ARF، یک نوع ویتامین K و یک پروتئین از خانواده HOX از عوامل رونویسی و دخیل در مسیر بیوسنتز متابولیسم‌های ثانویه می‌باشند. کاهش و افزایش بیان این ژن در تنش خشکی در اکوتیپ‌های 3 و 18، نشان از تاثیر دو برابری این ژن در افزایش و کاهش دو برابری تولید آنتوسیانین در زیره سبز دارد. بنابراین می‌تواند گزینه مناسبی در بررسی‌های آینده در تولید آنتوسیانین در زیره سبز باشد.

بیوسنتز آنتوسیانین می‌باشد. این پروتئین احتمالی از خانواده TFS و یک فاکتور رونویسی است. براساس تشابه بیش از 90 درصدی با پروتئین هویج، این پروتئین یک تنظیم کننده مثبت فرایند بیوسنتز آنتوسیانین و پروآنتوسیانین می‌باشد. با توجه به تشابه بیش از 90 درصدی با توالی هویج و بالاترین تشابه با ساختار ثانویه پروتئین مشابه در آرکیدوپسیس، ساختار ثانویه این پروتئین احتمالی در زیره سبز دارای 6 صفحه بتا و 7 مارپیچ آلفا خواهد بود. براساس توالی هویج احتمالاً این پروتئین با 10 پروتئین دیگر در تعامل خواهد بود. از آنجایی که این پروتئین یک فاکتور رونویسی می‌باشد و فاکتورهای رونویسی در گروه

منابع

1. سورنی، ج. رستمی احمدوندی، ح. کهیزی، د. معصومی، م. چقامیرزا، ک. کیانی، س. (1391) بررسی ارتباط نشانگرهای پروتئینی و مورفولوژیکی درتوده‌های گیاه دارویی زیره سبز (*Cuminum cyminum L.*). انجمن ژنتیک ایران، دوازدهمین کنگره ژنتیک ایران.
2. صادقی، د. مرتضویان، م. بختیاری زاده، م. (1396) ارزیابی توالی رونوشت گیاه دارویی زیره سبز (*Cuminum cyminum L.*) با استفاده از RNA-seq. مجله بیوتکنولوژی کشاورزی، دوره 9، شماره 4، ص 116-101.
3. درافشان، م. سلطانی حویزه، م. شریعتی، و. (1398) شناسایی ژن‌های مسیر بیوسنتزی اسکلت ترپنوئید در میوه گیاه دارویی هندوانه ابوجهل (*Citrullus colocynthis (L.) Schrad*) نشریه علمی تحقیقات گیاهان دارویی و معطر ایران، جلد 35، شماره 4، ص 691-702.
4. کاظم زاده بنه، ه. کرامتی، م. مهنا، ن. عابدی، ح. تیزفهم، پ. (1393) ژنتیک و بیوشیمی بیوسنتز آنتوسیانین در گیاهان. اولین کنگره زیست شناسی و علوم طبیعی ایران، ص 1-8.
5. Alqethami, A., and A.Y. Aldhebani. 2021. Saudi Journal of Biological Sciences Medicinal plants used in Jeddah , Saudi Arabia : Phytochemical screening. Saudi J. Biol. Sci. 28(1): 805–812. doi: 10.1016/j.sjbs.2020.11.013.
6. Amer, A. 2018. BIOTECHNOLOGY APPROACHES FOR IN VITRO PRODUCTION OF FLAVONOIDS. : 457–468. doi: 10.15414/jmbfs.2018.7.5.457-468.
7. Bogs, J., C. Kappel, S. Delrot, and V. Lauvergeat. 2011. Recent advances in the transcriptional regulation of the flavonoid biosynthetic pathway. 62(8): 2465–2483. doi: 10.1093/jxb/erq442.
8. Casati, P. 2012. Flavonoids : biosynthesis , biological functions , and biotechnological applications. 3(September): 1–15. doi: 10.3389/fpls.2012.00222.
9. Das, S.G., and G.P. Savage. 2012. Total and Soluble Oxalate Content of Some Indian Spices. Plant Foods Hum. Nutr. 67(2): 186–190. doi: 10.1007/s11130-012-0278-0.
10. Ghasemi, S., H.H. Kumleh, and M. Kordrostami. 2019. Changes in the expression of some genes involved in the biosynthesis of secondary metabolites in *Cuminum cyminum L.* under UV stress. Protoplasma 256(1): 279–290. doi: 10.1007/s00709-018-1297-y.
11. Jan, R., S. Asaf, M. Numan, and K. Kim. 2021. Plant Secondary Metabolite Biosynthesis and Transcriptional Regulation in Response to Biotic and Abiotic Stress Conditions. : 1–31.

in Plant-Herbivore Interactions.

16. Sybiya Vasantha Packiavathy, I.A., P. Agilandeswari, K.S. Musthafa, S. Karutha Pandian, and A. Veera Ravi. 2012. Antibiofilm and quorum sensing inhibitory potential of *Cuminum cyminum* and its secondary metabolite methyl eugenol against Gram negative bacterial pathogens. *Food Res. Int.* 45(1): 85–92. doi: 10.1016/j.foodres.2011.10.022.
17. Zhang, J., H. Lv, W. Liu, A. Ji, X. Zhang, et al. 2020. bHLH transcription factor SmbHLH92 negatively regulates biosynthesis of phenolic acids and tanshinones in *Salvia miltiorrhiza*. *Chinese Herb. Med.* 12(3): 237–246. doi: 10.1016/j.chmed.2020.04.001.
- 18.
12. Kaur, R., L. Aslam, S. Hussain, N. Kapoor, and R. Mahajan. 2021. Flavonoid Biosynthetic Pathway: Genetics and Biochemistry. 18(June): 271–286.
13. Saito, K., K. Yonekura-sakakibara, R. Nakabayashi, Y. Higashi, M. Yamazaki, et al. 2013. Plant Physiology and Biochemistry The flavonoid biosynthetic pathway in *Arabidopsis*: Structural and genetic diversity. *Plant Physiol. Biochem.* (February): 1–14. doi: 10.1016/j.plaphy.2013.02.001.
14. Silva, A.M.S. 2021. Plant Flavonoids: Chemical Characteristics and Biological Activity. : 1–16.
15. Singh, S., I. Kaur, and R. Kariyat. 2021a. The Multifunctional Roles of Polyphenols

"Bioinformatics study of selected genes involved in the biosynthesis of flavonoids in cumin"

Fereshte Lotfi – Seyed Mohammad Mahdi Mortazavian - Ali Izadi Darbandi - Hossein Ramshini

Abstract

Cumin is an annual and herbaceous plant belonging to the parsley family. In this research, five selected genes effective in the biosynthesis of flavonoids, which were identified by NGS method in cumin and had the most expression changes, were studied bioinformatically. Among these, two gene fragments, DN32640 and DN1196, which had the greatest evolutionary distance and were from two different families, were investigated for further identification. The sequence of DN1196 is from the GT6 family, a glycosyltransferase in the biosynthesis pathway of phenolic compounds, and the sequence of DN32640 is from the BHLH family, a transcription factor in the anthocyanin biosynthesis pathway. The coded protein of the sequence DN1196, with more than 85% similarity with a protein from the carrot plant, glycosylates it in the final step of flavonol synthesis by transferring a glucose molecule. In this process, two transcription factor proteins interact with this protein. The predicted secondary structure of the product of this gene was identified in accordance with the carrot protein sequence. The encoded protein of the DN32640 sequence, with more than 90% similarity to the carrot protein, is a transcription factor of the BHLH family and a component of the MBW complex. This transcription factor is involved in regulating the final stages of the path leading to the biosynthesis of anthocyanins and tannins. The secondary structure of this gene product was determined based on the identified protein sequence of carrot. Considering the importance of these two genes and the interaction of their product with other proteins, they can be used in genetic manipulations to produce more phenolic compounds, especially anthocyanins in cumin.

Key words: bioinformatics, prediction, cumin, secondary structure, flavonoids.