

شناسایی مقیاس گسترده نشانگرهای مولکولی ریزماهواره در گیاه دارویی

زنیان (*Trachyspermum ammi*) با استفاده از توالی‌یابی RNA

مهدی سلطانی حویزه^۱، سید احمد سادات نوری^{۳*}، وحید شریعتی^۴ و محبوبه امیری پور^۱

۱- دانش‌آموخته دکتری، گروه علوم زراعی و اصلاح‌نباتات، پردیس ابوریحان، دانشگاه تهران، تهران

۲- مربی، گروه ژنتیک و به‌نژادی گیاهی، دانشگاه آزاد اسلامی واحد اهواز، اهواز

۳- استاد، گروه علوم زراعی و اصلاح‌نباتات، پردیس ابوریحان، دانشگاه تهران، تهران

۴- استادیار، گروه بیوتکنولوژی مولکولی و مرکز ژنوم، پژوهشگاه ملی مهندسی ژنتیک و زیست‌فناوری، تهران

(تاریخ دریافت: ۱۳۹۶/۱۰/۱۶ - تاریخ پذیرش: ۱۳۹۷/۰۲/۲۹)

چکیده

گیاه دارویی زنیان یک منبع غنی از ترکیبات فعال دارویی و دارای اثرات مختلف دارویی است. با توجه به این‌که نشانگرهای ریزماهواره دارای نقش کلیدی در ژنوم و مرتبط با ژن‌ها به‌ویژه در بیوسنتز متابولیت‌های ثانویه در گیاهان دارویی هستند، لذا در این مطالعه، از توالی‌یابی ترنسکرپتوم زنیان برای اولین بار جهت شناسایی نشانگرهای ریزماهواره استفاده شد. پس از انجام توالی‌یابی توسط پلتفرم Illumina HiSeq 2000 به صورت دو طرفه، ارزیابی کیفیت خوانش‌ها توسط نرم‌افزار FastQC، تریم کردن با نرم‌افزار Trimmomatic و یکپارچه‌سازی نوپدید با استفاده از نرم‌افزار Trinity انجام گردید. در این پژوهش، ۱۱۴۶۸ توالی یونی‌ترنسکرپت (۷۹۱۳ توالی یونی‌ژن) حاوی ۱۳۵۹۳ ریزماهواره بالقوه یافت شد. فراوان‌ترین نوع ریزماهواره‌ها، دی‌نوکلئوتیدها (۶۷ درصد) و تری‌نوکلئوتیدها (۲۴ درصد) بودند. هم‌چنین تکرارهای شش‌تایی فراوان‌ترین تکرارها بودند و توالی غالب، AG/CT (۳۱ درصد) بود. ۶۵ درصد ریزماهواره‌ها از ریزماهواره کلاس دوم (۱۰ تا ۲۰ نوکلئوتید) و ۳۵ درصد از ریزماهواره کلاس اول (بیش از ۲۰ نوکلئوتید) بودند. فراوانی ریزماهواره‌ها تقریباً یک در هر ۱۰/۱ کیلوباز توالی یکپارچه شده بود. ۵۷/۹ درصد یونی‌ژن‌های حاوی ریزماهواره، با ژنوم هویج بلاست شدند. تعداد ۳۴۳۷ یونی‌ژن (۴۳ درصد) دارای دسته‌بندی کارکردی بودند که از میان آن‌ها تعداد ۲۲۱۹ یونی‌ژن (۶/۶ درصد) به دسته "فرآیند متابولیسم" و ۷۱ یونی‌ژن (۲/۱ درصد) به دسته "فرآیند متابولیسم ثانویه" تخصیص داشت. در این تحقیق ۱۲ ژن در مسیر پایه بیوسنتز ترپنوئیدها شناسایی شدند که ترنسکرپت مربوط به آن‌ها دارای ریزماهواره بود. این ریزماهواره‌ها احتمالاً در بیان ژن‌ها و تولید متابولیت‌ها به‌ویژه متابولیت‌های ثانویه نقش دارند. معرفی این نشانگرها می‌تواند برای مطالعات آینده انتخاب به کمک نشانگر، تنوع ژنتیکی و ساخت نقشه‌های ژنتیکی در این گیاه دارویی مورد بهره‌برداری قرار گیرد.

واژگان کلیدی: آپیاسه، ترپنوئیدها، ترنسکرپتوم، متابولیت‌های ثانویه

* نویسنده مسئول، آدرس پست الکترونیکی: noori@ut.ac.ir

مقدمه

ریزماهوره‌ها (SSRs) توالی‌های تکراری DNA بطول ۱ تا ۶ نوکلئوتید هستند که به صورت آرایه‌های متوالی در تمام ژنوم‌های یوکاریوتی وجود دارند. در مقایسه با سایر انواع نشانگرهای مولکولی، نشانگرهای ریزماهوره مزایای زیادی دارند که از آن جمله می‌توان به فراوانی بالا، توزیع تصادفی در کل ژنوم، محتوای اطلاعاتی بالا، وراثت هم‌بارز و تکرارپذیری اشاره نمود. این نشانگرها در تحقیقات ژنتیکی گیاهان دارویی از پراطاعات‌ترین نشانگرهای مولکولی و چندمنظوره هستند که شناسایی آن‌ها بسیار سخت و هزینه‌بر است (Li *et al.*, 2002; Li *et al.*, 2004). ارزیابی تنوع ژنتیکی با استفاده از نشانگرهای ریزماهوره یکی از رایج‌ترین کاربردهای این نوع نشانگر در مطالعات ژنتیک گیاهی می‌باشد (Shuorvazdi *et al.*, 2014; Karimbeigi *et al.*, 2016).

توالی‌یابی RNA با عمق زیاد می‌تواند مجموعه کل توالی‌های بیان شده موجود در بافت خاص، در نقطه زمانی مشخص و حتی رونوشت‌های کمیاب را نشان دهد. بدین منظور این روش می‌تواند تصویر تقریباً کاملی از حوادث ترنسکرپتوم در یک نمونه زیستی را نشان دهد (Alagna *et al.*, 2009; Barakat *et al.*, 2009; Dassanayake *et al.*, 2011; Bräutigam *et al.*, 2009). توالی‌یابی RNA در گونه‌های غیرمدل که اطلاعات ژنتیکی برای آن‌ها وجود ندارد یا اطلاعات ژنتیکی و داده‌های توالی‌یابی شده‌ی محدودی برای آن‌ها وجود دارد، قابل استفاده می‌باشد. به‌ویژه اینکه در این روش تمرکز توالی‌یابی، به‌جای توالی‌یابی کل ژنوم تنها بر توالی‌یابی نواحی کدشونده است. یکی از مهم‌ترین کاربردهای توالی‌یابی نسل آینده^۱ در ژنتیک جمعیت، شناسایی نشانگرهای مولکولی در مقیاس گسترده می‌باشد (Strickler *et al.*, 2012).

تاکنون در گیاهان دارویی مطالعات زیادی با استفاده از نشانگرهای مولکولی بر روی تنوع ژنتیکی صورت گرفته است (Farajpour *et al.*, 2011; Bahmani *et al.*, 2012;)

(Sadati *et al.*, 2016). نشانگرهای مولکولی تقریباً در تمام گونه‌های گیاهان دارویی می‌توانند از توالی‌های ترنسکرپتوم یا ژنوم در مقیاس وسیع ایجاد شوند. یکی از مزایای شناسایی نشانگرها از توالی‌های ترنسکرپتوم این است که آن‌ها با ژن‌های کارکردی مرتبط خواهند شد (Hiremath *et al.*, 2011). ریزماهوره‌های موجود در نواحی کدکننده ژن‌ها، بیان و عملکرد ژن را کنترل می‌کنند و به‌عنوان یک منبع ارزشمند برای مطالعات ژنتیکی آینده محسوب می‌شوند (Li *et al.*, 2002). ریزماهوره‌های پیوسته با توالی بیان شده کوتاه (EST) نوعاً در نواحی غیرترجمه‌ای^۲ (5' و 3' یوتی‌آرز) یافت می‌شوند (Primmer, 2009). پیش‌بینی می‌شود که این مکان‌های ژنی برای نشان دادن اثرات انتخاب (Vasemagi *et al.*, 2005) دارای احتمال بالاتری (در مقایسه با نشانگرهای ژنتیکی خنثی) بوده و آن‌ها هم‌چنین می‌توانند در عمل تفاوت معنی‌داری در تنظیم بیان و کارکرد ژن ایجاد کنند. تا زمانی که نرخ جهش در توالی کدکننده پروتئین در DNA پایین‌تر از سایر توالی‌های DNA (مانند ایترون و توالی بین‌ژنی) می‌باشد، انتظار می‌رود تعداد ریزماهوره‌ها و چندشکلی‌های درون این نواحی پایین‌تر از توالی‌های غیرکدکننده باشد (Blanca *et al.*, 2011). یکی دیگر از مزایای ریزماهوره‌های بر پایه ژن، قابلیت انتقال بالاتر نشانگرهای ریزماهوره به‌دست آمده از این روش برای سایر گونه‌ها می‌باشد. داده‌های توالی‌یابی نمونه‌های انفرادی یا نمونه‌های درهم آمیخته شده افراد، هر دو برای این هدف مناسب می‌باشند (Futschik and Schlötterer, 2010). توالی‌یابی RNA می‌تواند آنالیز مسیر متابولیکی گیاهان دارویی و جستجوی نشانگرهای مولکولی برای روش‌های اصلاح مولکولی را تسهیل و تسریع بخشد.

زنیان با نام علمی *Trachyspermum ammi* (مترادف: *Carum copticum*) که معمولاً به‌عنوان Ajowan شناخته شده است در کشورهای ایران، هند، پاکستان، افغانستان و مصر پراکنده می‌باشد و در این کشورها به‌صورت وسیع

1- Next generation sequencing (NGS)

2- UTRs

بنابراین توالی‌یابی گسترده ترنسکرپتوم یا ژنوم برای گیاه دارویی زنیان به شدت مورد نیاز بوده و این منابع می‌تواند برای شناسایی نشانگرهای جدید، انجام بهتر مطالعات نشانگرهای مولکولی، مکان‌یابی ژن و مقایسه ژنوم و غیره مورد استفاده قرار گیرد. با توجه به اینکه زنیان یک گیاه دیپلوئید با درجه بالایی از هتروزیگوسیتی است (Noori et al., 2017)، توالی‌یابی کامل ژنوم آن دارای هزینه بسیار بالا است و زمان زیادی طول می‌کشد. اما خوشبختانه ابداع روش توالی‌یابی ترنسکرپتوم جایگزین مناسبی برای توالی‌یابی کامل ژنوم است. مطالعات پیشین انجام شده جهت بررسی تنوع ژنتیکی ژرم‌پلاسم گیاه دارویی زنیان، تنها به نشانگر ISSR (Modareskia et al., 2013; Heidari et al., 2016) و نشانگر RAPD (Sargazi, 2016) محدود بوده است. لذا این تحقیق جهت شناسایی مقیاس گسترده نشانگرهای ریزماهوره گیاه دارویی زنیان با استفاده از داده‌های حاصل از توالی‌یابی RNA، جهت معرفی این نشانگرها برای مطالعات آینده تنوع ژنتیکی و شناسایی ریزماهوره‌های احتمالی مرتبط به ژن‌های مسیر متابولیت‌های ثانویه در این گیاه دارویی طراحی و انجام گردیده است.

مواد و روش‌ها

مواد گیاهی: بذر دو اکوتیپ مختلف گیاه زنیان (اکوتیپ شیراز با کد ۱۲۳۱۳ و اکوتیپ اراک با کد ۱۴۴۹۲) از مؤسسه تحقیقات مراتع و جنگل‌های کشور تهیه گردید و در سینی‌های نشاء دارای بستر کشت (شامل کوکوپیت و پرلیت) در اسفندماه ۱۳۹۲ در گلخانه تحقیقاتی پردیس ابوریحان دانشگاه تهران کشت گردید. دو اکوتیپ انتخاب شده زنیان در مقدار تیمول با هم متفاوت بودند (Mirzahosseini et al., 2017). آبیاری سینی‌های کشت نشاء از زمان کشت تا زمان انتقال به گلدان روزی یک تا دو مرتبه به صورت آبیاری انجام گردید. از هر اکوتیپ دو بوته انتخاب و از بافت گل‌آذین هر بوته به طور جداگانه (پنج روز پس از گرده افشانی) نمونه‌برداری شد. نمونه‌ها

کشت می‌شود. این میوه دارای خواصی هم‌چون محرک، ضد تشنج و ضد نفخ می‌باشد و به طور سنتی به عنوان یک عامل مهم درمانی برای نفخ شکم، سوءهاضمه، اسهال، تومور شکمی، دردهای شکمی، زردی، مشکلات تنفسی، فقدان اشتها، محرک شیر، آسم و آموره (قطع قاعدگی) استفاده می‌شود. با بررسی‌های دارویی، ثابت شده است که میوه این گیاه دارای فعالیت‌های مختلف دارویی مثل داروهای ضدقارچی، آنتی‌اکسیدانی، ضد میکروبی، ضد درد، سیتوتوکسیک^۱، کاهش چربی، ضد فشار خون، ضد تشنج، اتساع نایچه، ضد سنگ‌سازی، ادرارآور، سقط جنین، ضد سرفه، ضد انگل، ضد کرم و ضد فیلاریا^۲ می‌باشد. همچنین تحقیقات انجام شده، حضور ترکیبات مختلف فیتوشیمیایی را در این گیاه نشان می‌دهد که به طور عمده شامل تیمول، گاماترینین، پاراسیمین می‌باشد. مطالعات نشان می‌دهد که زنیان یک منبع مملو از ترکیبات فعال دارویی و اثرات مختلف دارویی می‌باشد، از این رو، پیدا کردن کاربردهای درمانی جدید آن به شدت مورد علاقه محققین است (Bairwa et al., 2012; Mirzahosseini et al., 2017).

علی‌رغم اهمیت بالای گیاهان دارویی و بومی کشور از جمله گیاه دارویی زنیان، تنها گزارش‌های بسیار معدودی از قبیل مطالعات زراعی، شناسایی ترکیبات عطرمایه و اثرات دارویی عطرمایه و عصاره آن در ایران صورت گرفته (Mirzahosseini et al., 2017; Soltani Howyzeh et al., 2018a) و تاکنون هیچ گزارشی در رابطه با شناسایی نشانگرهای مولکولی این گیاه در ایران و جهان منتشر نشده است. در طی دهه‌های اخیر تعداد زیادی از توالی‌های ژنوم و ترنسکرپتوم در بسیاری از موجودات مدل به وجود آمده است (Fu et al., 2013). این در حالی است که برای گیاه دارویی زنیان تاکنون تنها ۲۰ توالی نوکلئوتیدی (تا آذرماه ماه ۱۳۹۶) در پایگاه اطلاعاتی NCBI وجود دارد و هیچ‌گونه توالی EST و یا نشانگر ریزماهوره در این گیاه شناسایی نشده است.

1- Cytotoxic
2- Antifilarial

Trimmomatic (نسخه 0.36) و مقدار $QC \geq 30$ حذف شدند. پس از تریم کردن (Trimming)، نرم افزار FastQC برای ارزیابی مجدد ویژگی های کتابخانه ها و تأیید کارایی تریم کردن مورد استفاده قرار گرفت. خوانش های فیلتر شده با کیفیت بالا برای تحلیل های بعدی مورد استفاده واقع شدند. یکپارچه سازی نوپدید خوانش های با کیفیت بالا با استفاده از نرم افزار Trinity (معرفی شده در تاریخ ۲۰۱۶/۰۳/۱۷) انجام گردید. پس از بهینه سازی، Kmer معادل ۳۲ انتخاب و برای سایر پارامترهای نرم افزار Trinity از تنظیمات استاندارد استفاده شد. توالی های به دست آمده از Trinity یونی ترنسکرپت^۲ نامیده می شود که ایزوفرم های رونوشت مورد قبول را نشان می دهد. کلاستر یونی ترنسکرپت ها، یونی ژن^۳ نامیده می شود که ژن های شناسایی شده مورد قبول را نشان می دهد (Grabherr *et al.*, 2011).

شناسایی نشانگرهای ریزماهوره: نشانگرهای ریزماهوره با استفاده از ابزار شناسایی ریزماهوره به نام MISA^۴ شناسایی شدند. حداقل تکرار برای نشانگرهای مونونوکلوئیدی "ده" و برای نشانگرهای دی-نوکلوئیدی، "شش" و برای نشانگرهای تری، تترا، پنتا و هگزا-نوکلوئیدی "پنج" قرار داده شد (Beier *et al.*, 2017).

تفسیر کارکردی^۵ یونی ژن های یکپارچه شده زیان حاوی ریزماهوره: تفسیر کارکردی ترنسکرپتوم زیان به وسیله فرستگر^۶ تفسیر نرم افزار Trinotate^۷ انجام شد. یونی ژن های یکپارچه سازی شده زیان حاوی ریزماهوره علیه پروتئین های Non-redundant پایگاه های اطلاعاتی NCBI، UniProt، آرآییدوپسیس و هویج با E-Value کمتر از 10^{-5} با استفاده از سرور محلی (Local Server) دارای ۲۵۶ گیگابایت حافظه موقت (RAM) و سی پی یو با ۸۰ هسته و سیستم عامل لینوکس بلاست شدند. مسیرهای متابولیتی و

در ازت مایع قرار گرفته و سپس به آزمایشگاه منتقل و تا زمان استخراج در فریزر $-80^{\circ}C$ درجه سانتی گراد قرار گرفتند. **استخراج RNA، ساخت کتابخانه و توالی یابی:** RNA کل از نمونه های گل آذین گیاه دارویی زیان به وسیله بافر TRIzol reagent (Invitrogen) طبق دستورالعمل شرکت سازنده استخراج گردید. نمونه های RNA با هضم کننده DNA (TURBO DNaseI, Ambion) تیمار شدند. کیفیت و کمیت RNA استخراج شده به ترتیب توسط ژل آگاروز ۱ درصد و دستگاه نانودراپ نورسنج مدل NanoDrop 1000 spectrophotometer (Thermo Scientific)، تعیین گردید. سپس بررسی بیشتر کیفیت توسط دستگاه QC Bioanalyzer مدل (Agilent Technologies, Hørsholm, Denmark) در موسسه Beijing Genomes Institute (BGI) در کشور چین صورت گرفت و نمونه های دارای عدد RIN (RNA integrity number) بیش از ۸ برای فرآیند توالی یابی انتخاب شدند. انتخاب دم پلی A، آماده سازی cDNA، اتصال آداپتورها، تشکیل کلاسترها و توالی یابی در موسسه BGI و طبق توصیه شرکت سازنده، با استفاده از کیت استاندارد ایلومینا انجام شد. توالی یابی توسط پلتفرم Illumina HiSeq 2000 به صورت دوطرفه (Paired-end) و طول ۱۰۰ نوکلئوتید انجام شد. داده های خام چهار ژنوتیپ (دو بوته از هر اکوتیپ) در آرشیو خوانش توالی (SRA) NBCI sequence read archive با شماره دسترسی های SRR5137050، SRR5137051، SRR5137052 و SRR5137053 ثبت شدند.

آماده سازی داده های RNA-Seq و یکپارچه سازی نوپدید: در ابتدا کیفیت خوانش های خام به دست آمده از توالی یابی توسط نرم افزار FastQC سنجیده شد. سپس برای هر نمونه با توجه به نتایج کنترل کیفی، آداپتورها و توالی های کم کیفیت خوانش های خام با استفاده از نرم افزار

1- De novo assembly
2- Unitranscripts
3- Unigenes
4- <http://pgrc.ipk-gatersleben.de/misa>
5- Functional annotation
6- Pipeline
7- <http://trinotate.github.io>

(فلفل قرمز) (Lu et al., 2012)، *Citrullus lanatus* (هندوانه) (Guo et al., 2011)، *Scabiosa columbaria* (گل داوودی) (Angeloni et al., 2011) و *Polygonum cuspidatum* (هفت‌بند ژاپنی) (Hao et al., 2012) بوده است در حالی که ریزماهوره دی‌نوکلئوتیدی، تکرار غالب در *Sonneratia alba* (نوعی شاهپسند) (Chen et al., 2011)، *Sesamum indicum* (کنجد) (Wei et al., 2011)، *S. miltiorrhiza* (Li et al., 2010) و *P. notoginseng* (Luo et al., 2011) و *G. Biloba* (Lin et al., 2011) بود. در این پژوهش، در مجموع ریزماهوره‌های با تکرارهای شش‌تایی فراوان‌ترین بودند و پس از آن‌ها به ترتیب ریزماهوره‌هایی با هفت، پنج و هشت تکرار فراوان‌ترین بودند (جدول ۲). در گیاه کنجد نتایج کاملاً مشابه‌ای در مورد بیشترین تکرارهای ریزماهوره‌ها به دست آمد (Wei et al., 2011). در گیاه ارکیده دندروبیوم (*Dendrobium officinale*) نیز ریزماهوره‌های با تکرارهای شش‌تایی فراوان‌ترین بودند و پس از آن‌ها به ترتیب ریزماهوره‌هایی با پنج، هفت و هشت تکرار قرار گرفتند (Xu et al., 2017). فراوان‌ترین ریزماهوره دی‌نوکلئوتیدی به دست آمده در ترنسکرپتوم گیاه دارویی زینان AG/CT با ۴۲۹۳ تکرار (۳۱ درصد) بود. در میان ریزماهوره تری‌نوکلئوتیدی نیز به ترتیب AAT/ATT، AAG/CTT، ATC/GAT و AGC/GCT بیشترین فراوانی را به خود اختصاص دادند (شکل ۲). در گیاهان ارکیده دندروبیوم و کنجد نتایج مشابهی گزارش شده است. در این گیاهان نیز فراوان‌ترین ریزماهوره دی‌نوکلئوتیدی AG/CT و فراوان‌ترین ریزماهوره تری‌نوکلئوتیدی AAG/CTT بود (Wei et al., 2011; Xu et al., 2017). این در حالی است که در گیاه بامبو (*Dendrocalamus latiflorus*) فراوان‌ترین ریزماهوره دی‌نوکلئوتیدی AG/CT گزارش گردید اما از سوی دیگر فراوان‌ترین ریزماهوره تری‌نوکلئوتیدی CCG/CGG بود (Bhandawat et al., 2016).

توصیف کارکردی برای هر یک از یونی‌ژن‌های زینان به وسیله دایره‌المعارف ژن‌ها و ژنوم کیوتو (KEGG^۱) با استفاده از سرور تفسیر اتوماتیک KEGG به نام KAAS^۲ بر اساس تخصیص عدد KO^۳ به آن‌ها به دست آمد. تخصیص عدد KO بر اساس روش SBH^۴ انجام گرفت. پایگاه KAAS تفسیر توالی‌های ارسال شده را با استفاده از شناسه‌های KO به صورت خودکار انجام می‌دهد. هر KO معادل یک گروه ارتولوگ از ژن‌هایی است که به صورت مستقیم به یک رکورد در مسیرهای پایگاه KEGG لینک شده (Moriya et al., 2007) و در نتیجه مرجع اطلاعات برای ارتباط دادن ژنوم‌ها به سیستم‌زیستی است. طبقه‌بندی کارکردی ژن شناسی (GO) برای همه یونی‌ژن‌های یکپارچه‌شده به وسیله نرم‌افزار تحت وب WEGO^۵ صورت گرفت.

نتایج و بحث

جستجوی نشانگرهای ریزماهوره با استفاده از نرم‌افزار MISA در مجموعه داده‌های یکپارچه شده ترنسکرپتوم گیاه دارویی زینان (جدول ۱)، ۱۱۴۶۸ توالی یونی‌ترنسکرپت (۷۹۱۳ توالی یونی‌ژن) حاوی ۱۳۵۹۳ ریزماهوره بالقوه را نشان داد.

ریزماهوره‌های دی‌نوکلئوتیدی فراوان‌ترین ریزماهوره‌ها (۶۷ درصد) در مجموعه داده این پژوهش بودند. نتایج مشابهی در برخی گیاهان دولپه‌ای دیگر مانند آرییدوپسیس، انگور و بادام‌زمینی گزارش شده است (Kumpatla and Mukhopadhyay, 2005). پس از دی‌نوکلئوتیدها به ترتیب تری‌نوکلئوتیدها (۳۲۳۲ عدد، ۲۴ درصد)، مونونوکلئوتیدها (۶ درصد)، تترانوکلئوتیدها (۲ درصد)، هگزانوکلئوتیدها (حدود یک درصد) و پنتانوکلئوتیدها (کمتر از یک درصد) قرار داشتند (شکل ۱).

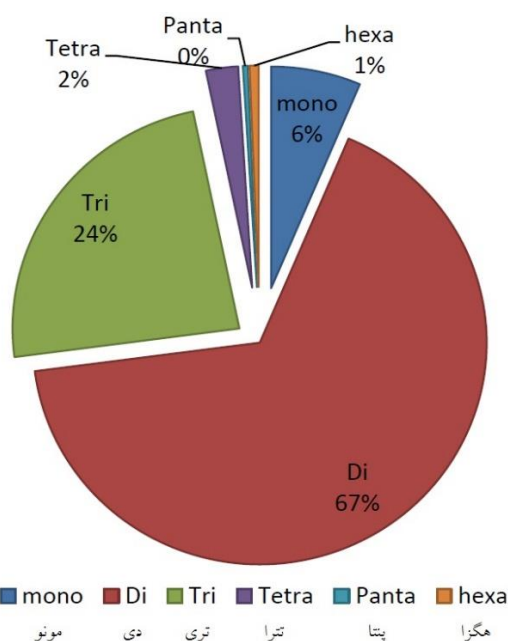
توالی‌های تکراری کوتاه تری‌نوکلئوتیدی، نوع غالب تکرار در مجموعه داده ترنسکرپتوم *A. tridentate* (درمنه معمولی) (Bajgain et al., 2011)، *Capsicum annuum*

1- Kyoto encyclopedia of genes and genome
2- KEGG automatic annotation server,
<http://www.genome.jp/kegg/kaas>
3- KEGG orthology
4- Single-directional best hit
5- <http://wego.genomics.org.cn/cgi-bin/wego/index.pl>

جدول ۱- گزارش آماری کانتیگ‌های یکپارچه شده توسط نرم افزار Trinity

Table 1. Trinity assembly stats report of assembled contigs

شمارش تعداد رونوشت‌ها، ژن‌ها و غیره		
Counts of transcripts, genes, etc.		
تعداد کل ژن‌های ترینیتی	تعداد کل رونوشت‌های ترینیتی	درصد GC
Total trinity genes	Total trinity transcripts	GC-content percentage
123488	151115	38.35
آمار یکپارچه‌سازی ترینیتی		
Trinity assembly stats		
آمار	براساس همه کانتیگ‌های رونوشت	براساس تنها بلندترین ایزوform در هر ژن
Stats	Based on all transcript contigs	Based on only longest isoform per gene
کانتیگ ان ۱۰	3069	2894
Contig N10		
کانتیگ ان ۲۰	2363	2189
Contig N20		
کانتیگ ان ۳۰	1944	1740
Contig N30		
کانتیگ ان ۴۰	1603	1346
Contig N40		
کانتیگ ان ۵۰	1291	994
Contig N50		
میانگ طول کانتیگ	454	385
Median contig length		
متوسط طول کانتیگ	782.02	659.33
Average contig		
کل نوکلئوتیدهای یکپارچه‌سازی شده	118175363	81419608
Total assembled bases		



شکل ۱- فراوانی نسبی نوع ریزماهوره‌های شناسایی شده در ترنسکرپتوم گیاه دارویی زنیان

Figure 1. Relative frequency of microsatellite repeat types identified in medicinal plant transcriptome of *T. ammi*.

جدول ۲- فراوانی ریزماهواره‌های زنیان براساس نوع و تعداد تکرارها

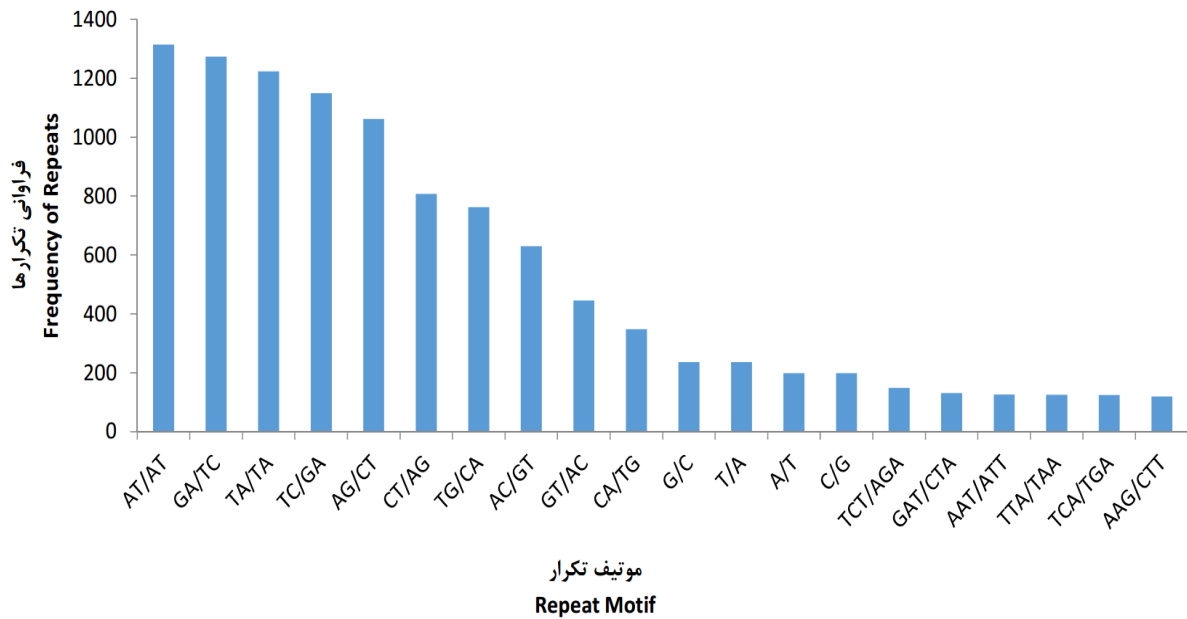
Table 2. Frequency of *T. ammi* microsatellites based on type and number of repeats

تکرار Repeat	نوع Type						مجموع Total
	مونو Mono	دی Di	تری Tri	تترا Tetra	پنتا Penta	هگزا Hexa	
5			1703	199	34	28	1964
6		2473	705	77	6	32	3293
7		1583	380	12	3	17	1995
8		1148	225	8	0	4	1385
9		785	35	7	0	3	830
10	144	519	46	2	1	0	712
11	176	445	23	0	2	0	646
12	114	412	20	4	1	3	554
13	110	195	33	0	0	0	388
14	76	202	19	2	0	0	299
15	62	180	6	1	0	0	249
16	28	162	10	0	0	0	200
17	36	162	6	0	0	0	204
18	6	144	7	0	0	0	157
19	12	141	6	0	0	0	159
≥ 20	106	494	88	0	0	0	608
مجموع Total	870	9045	3232	312	47	87	13593

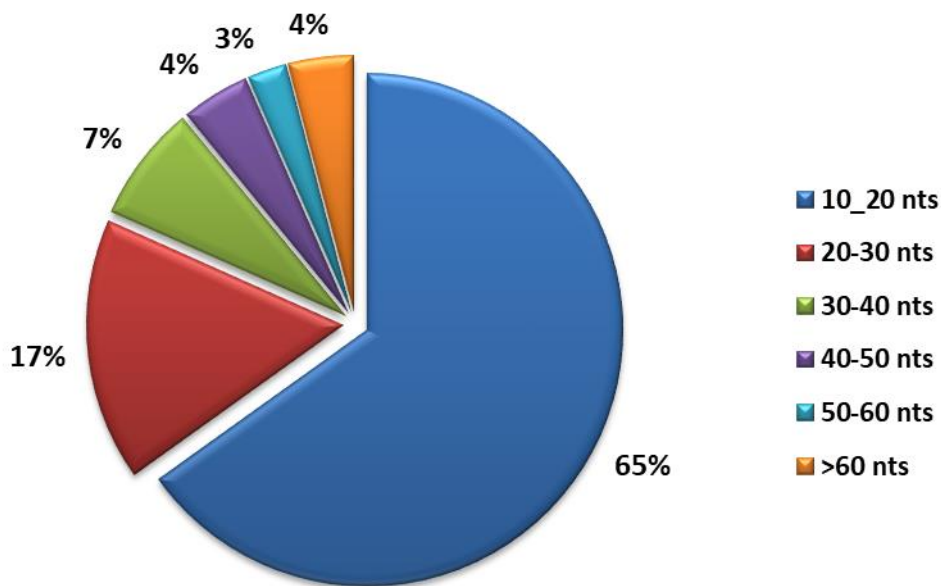
تفسیر کارکردی یونی‌ژن‌های یکپارچه‌شده زنیان حاوی ریزماهواره: تفسیر تک‌ژهای زنیان حاوی ریزماهواره با استفاده از Blastx علیه پایگاه‌های داده KAAS، آرآبیدوپسیس، پروتئین‌های Non-redundant پایگاه NCBI، UniProt و هویج انجام گردید. در مجموع ۱۵۳۱ (۱۹/۳ درصد)، ۴۰۶۶ (۵۱/۳ درصد)، ۳۳۷۸ (۴۲/۶ درصد)، ۴۰۰۶ (۵۰/۶ درصد) و ۴۵۸۳ (۵۷/۹ درصد) یونی‌ژن از همه کتابخانه‌ها به ترتیب به پایگاه‌های داده KAAS، آرآبیدوپسیس، UniProt، پروتئین‌های Non-redundant پایگاه NCBI و هویج گزارمان شدند (شکل ۴). تعداد یونی‌ژن‌های تفسیر شده مشترک بین همه پایگاه‌های اطلاعاتی ۱۴۶۴ بود (شکل ۴). یونی‌ژن‌های زنیان به طور تقریباً مشابهی با پایگاه داده هویج، بیش از سایر پایگاه‌های داده تفسیر شدند. این امر تأییدکننده این مسئله است که زنیان از خانواده آپیاسه^۲ بوده و همانندی بالایی با ژنوم هویج دارد.

در میان نشانگرهای ریزماهواره شناسایی شده در زنیان، ۶۵ درصد از نوع ریزماهواره کلاس دوم (به طول ۱۰ تا ۲۰ نوکلئوتید) و ۳۵ درصد از نوع ریزماهواره کلاس اول (به طول بیش از ۲۰ نوکلئوتید) بودند (شکل ۳). فراوانی تکرارهای ریزماهواره در ترنسکرپتوم گل‌آذین گیاه دارویی زنیان تقریباً یک در هر ۱۰/۱ کیلوباز توالی یکپارچه شده بود. در گیاه ارکیده دندروبیوم فراوانی ریزماهواره‌های نوع کلاس دوم ۹۲ درصد و فراوانی ریزماهواره‌های نوع کلاس اول ۸ درصد بود. در این گیاه فراوانی تکرارهای ریزماهواره تقریباً یک در ۴ کیلوباز توالی یکپارچه شده بود (Xu et al., 2017). فراوانی تکرارهای ریزماهواره در ترنسکرپتوم گل‌آذین گیاه دارویی زنیان کمتر از فراوانی تکرارهای ریزماهواره در گیاهان برنج (یک در ۳/۶ کیلوباز)، سورگوم (یک در ۵/۹ کیلوباز)، جو (یک در ۸/۹ کیلوباز) و ذرت (یک در ۹ کیلوباز) و بیشتر از فراوانی تکرارهای ریزماهواره در ترنسکرپتوم گیاهان گندم (یک در ۱۰/۶ کیلوباز) و نیشکر (یک در ۱۰/۹ کیلوباز) بود (Parida et al., 2010).

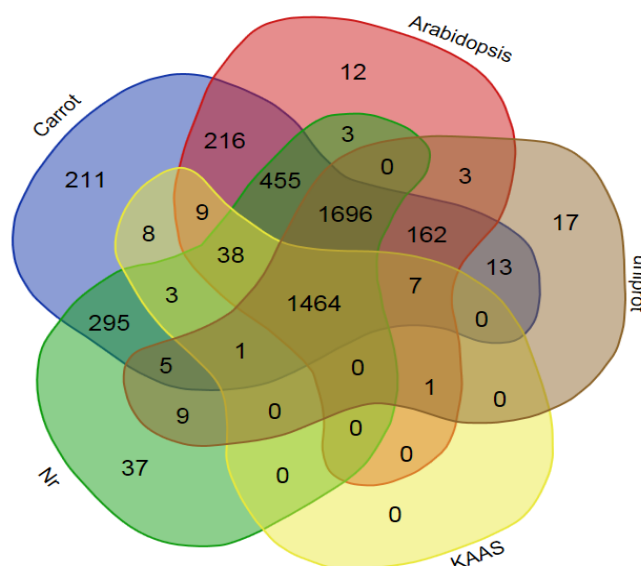
1- Annotation
2- Apiaceae



شکل ۲- فراوانی موتیف تکرارهای ریزماهواره شناسایی شده در ترنسکرپتوم گیاه دارویی زنیان
 Figure 2. Frequency of microsatellite repeat motifs identified in medicinal plant transcriptome of *T. ammi*



شکل ۳- فراوانی طول تکرارهای ریزماهواره شناسایی شده در ترنسکرپتوم گیاه دارویی زنیان
 Figure 3. Frequency of microsatellite repeat length identified in medicinal plant transcriptome of *T. ammi*

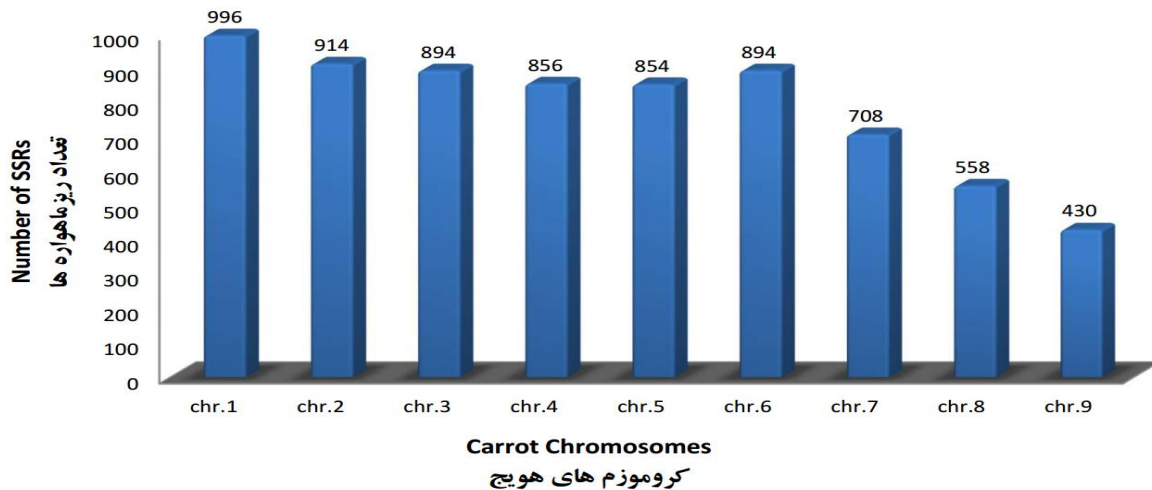


شکل ۴- نمودار ون نتایج تفسیر بلاست یونی‌ژن‌های یکپارچه‌سازی شده حاوی ریزماهوره علیه پایگاه‌های اطلاعاتی مختلف
Figure 4. Venn diagram of BLAST annotation results of assembled unigenes containing microsatellites against different databases

در ۵۲ دسته کارکردی طبقه‌بندی شد (شکل ۶). دستجات "سلول"، "بخش سلولی"، "اندامک"، "اتصال"، "فروکافتی"، "تنظیم زیستی"، "فرآیند سلولی"، "فرآیند متابولیکی"، "رنگدانه‌سازی" و "پاسخ به محرک‌ها" بیشترین درصد تخصیص را دارا بودند (شکل ۶). در طبقه فرآیندهای زیستی (BP) در میان همه یونی‌ژن‌های دسته‌بندی شده ۲۲۱۹ یونی‌ژن (۶۴/۶ درصد) به دسته "فرآیند متابولیکی" متعلق بود که از میان آن‌ها ۷۱ یونی‌ژن (۲/۱ درصد) به دسته "فرآیند متابولیکی ثانویه" تخصیص داشت. فراوانی تخصیص عبارت‌های GO مرتبط به فرآیند متابولیکی به توالی‌های یونی‌ژن حاوی تکرارهای ریزماهوره می‌تواند در شناسایی نشانگرهای ریزماهوره مرتبط به بیان ژن‌ها و تولید و کنترل متابولیت‌ها به‌ویژه متابولیت‌های ثانویه کمک کند. در مطالعه ریزماهوره‌های گیاه کنگد نیز دسته "فرآیند متابولیکی" با ۳۹/۲ درصد گروه غالب در طبقه فرآیندهای زیستی (BP) بود (Wei et al., 2011). نتایج مشابهی در مطالعه ریزماهوره‌های گیاهان بادام و درخت کائوچو به‌دست آمد (Zhang et al., 2012; Li et al., 2012).

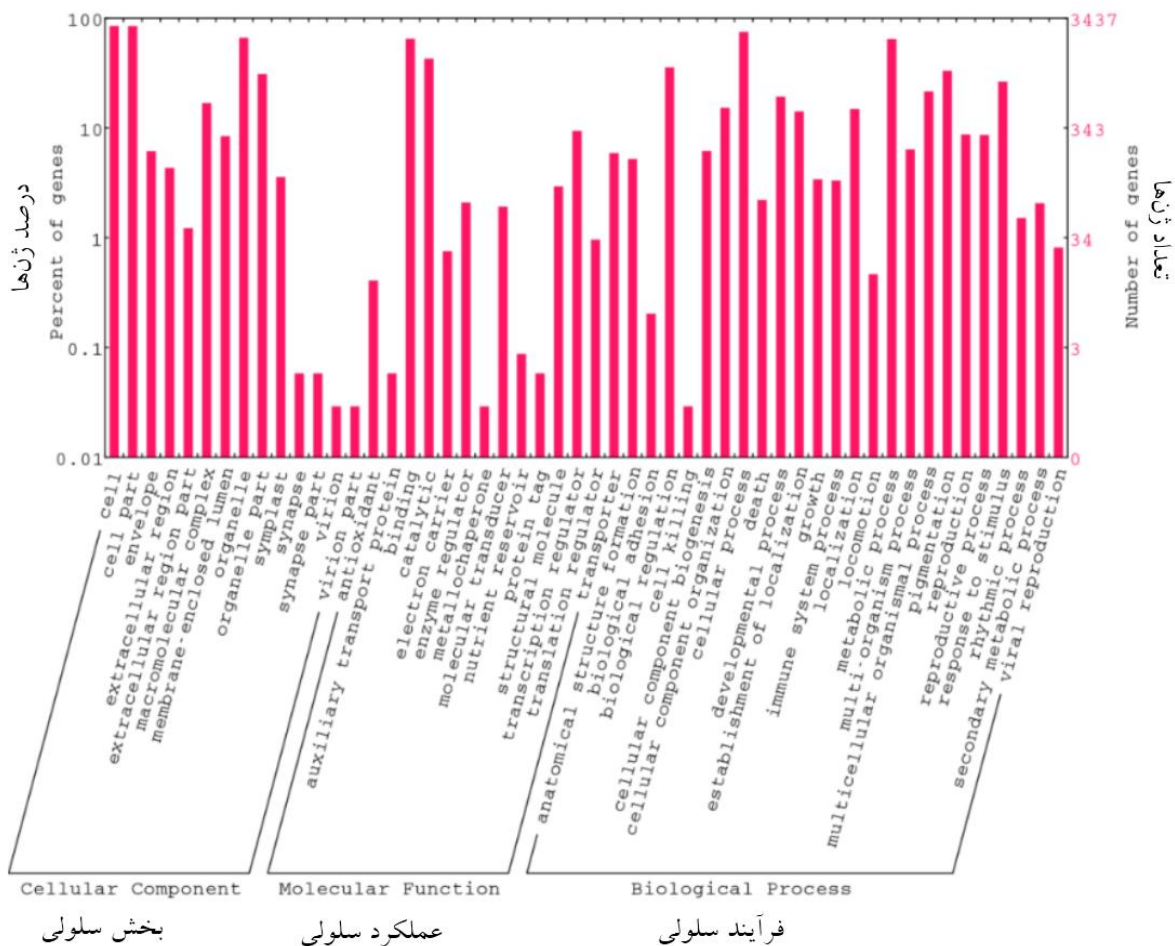
با توجه به اینکه توالی ژنوم گیاه دارویی زنبان در دسترس نیست، نقشه‌یابی فیزیکی ریزماهوره‌ها در کروموزوم‌های این گیاه امکان‌پذیر نمی‌باشد. هرچند با توجه به این‌که حداکثر تعداد توالی‌های مشابه به‌دست آمده از طریق بلاست مربوط به گیاه هویج بود، بنابراین از ژنوم هویج برای شناسایی تعداد ریزماهوره‌ها در کروموزوم‌های گیاه هویج استفاده گردید. نکته لازم به ذکر اینکه گیاه هویج به‌عنوان سردسته و مدل خانواده آپاسه می‌باشد و تعداد کروموزوم آن $(2n = 2x = 18)$ برابر با تعداد کروموزوم گیاه دارویی زنبان است (Noori et al., 2017). بیشترین تعداد ریزماهوره به کروموزوم شماره یک و کم‌ترین تعداد به کروموزوم شماره ۹ اختصاص یافتند (شکل ۵). این نتیجه ممکن است ناشی از تفاوت تعداد ژن‌های کارکردی در کروموزوم‌ها و توزیع گسترده ژنومی ریزماهوره‌ها باشد. پیش‌بینی عملکرد توالی‌های یکپارچه‌سازی شده دارای ریزماهوره: از میان ۷۹۱۳ توالی یونی‌ژن حاوی تکرارهای ریزماهوره، تعداد ۳۴۳۷ توالی یونی‌ژن (۴۳ درصد) دارای دسته‌بندی کارکردی بودند (شکل ۶). نمودار تفسیر ژن‌شناسی اینترنتی (WEGO) نشان داد که همه یونی‌ژن‌ها

1- Catalytic
2- Secondary metabolic process



شکل ۵- نقشه‌یابی مقایسه‌ای *In silico* توالی‌های ریزماهوره‌های زیان‌علیه ژنوم هویج

Figure 5. *In silico* comparative mapping of *T. ammi* microsatellite sequences on carrot genome



شکل ۶- نمودار دسته‌بندی کارکردی یونین‌های یکپارچه‌سازی شده حاوی ریزماهوره

Figure 6. Functional classification of assembled unigenes containing microsatellites

هر میله درصد یونین‌های متصل به گزاره‌های GO را نشان می‌دهد. محور Y در مبنای لگاریتم ۱۰ می‌باشد.

Each bar represents the percentage of unigenes matches to each GO term. Plot with Y axis is in log (10) scale.

نتایج به دست آمده از آنالیز مسیرهای بیوشیمیایی به وسیله سرور تفسیر اتوماتیک KEGG نشان داد که ترنسکرپت‌های یکپارچه‌سازی شده دارای نشانگر ریزماهوره در ۳۳۰ مسیر بیوشیمیایی پایگاه KEGG حضور داشتند. بیشترین ترنسکرپت‌های دارای ریزماهوره، به ترتیب به مسیر متابولیکی (ko01100) و مسیر بیوستنز متابولیت‌های ثانویه (ko01110) با ۲۲۶ و ۱۱۶ ترنسکرپت بودند. در مسیر بیوستنز پایه ترپنوئیدها (Terpenoid backbone biosynthesis) که نقش بسیار مهمی در تولید ترکیبات دارویی گیاه زنیان دارد، ۱۲ ژن از ژن‌های مسیر شناسایی شدند که ترنسکرپت مربوط به آن‌ها دارای ریزماهوره بود (جدول ۳).

تنوع ریزماهوره‌ها در نواحی کدکننده، نواحی غیرترجمه شده در انتهای ۵'، نواحی غیر ترجمه شده در انتهای ۳' و اینترون‌ها منجر به تغییراتی در فرآیندهای رونویسی و ترجمه‌ای شده و در نتیجه از طریق تنظیم بیان ژن منتهی به تغییرات فنوتیپی می‌گردد (Li et al., 2004). این مسأله نشان دهنده اهمیت ریزماهوره‌های واقع شده در نواحی کدکننده است. تنوع در ریزماهوره‌های مناطق ژنی می‌تواند آلل‌های مناسبی برای موفقیت‌های حیاتی در محیط‌های مختلف ایجاد کند (Li et al., 2004). نشانگرهای SSR دارای نقش کلیدی در تحمل به تنش، متابولیسم و تثبیت کربن، بیوستنز متابولیت‌های ثانویه، رشد و فتوستنز شناسایی گردیده‌اند (Bhandawat et al., 2016).

جدول ۳- ژن‌های شناسایی شده در مسیر بیوستنز پایه ترپنوئیدها (ko00900) دارای ریزماهوره در ترانسکرپتوم گل‌آذین گیاه

دارویی زنیان

Table 3. Identified genes in terpenoid backbone biosynthesis pathway (Ko00900) containing microsatellite in *T. ammi* inflorescence transcriptome

ردیف No.	نام ژن Gene name	توصیف ژن Gene description	کد ژن در KEGG KEGG code	کد آنزیم Enzyme code	یونی‌ترنسکرپت Unitranscript	طول (bp) Length (bp)	موقعیت (bp) Position (bp)	موتیف Motif
1	HMGCR	Hydroxymethylglutaryl-CoA reductase (NADPH	K00021	EC:1.1.1.34	DN46370_c0_g1_i1	2031	44 - 58	(CCA)5
2	FDPS	Farnesyl diphosphate synthase	K00787	EC:2.5.1.1 2.5.1.10	DN45915_c0_g1_i1	1544	32 - 46	(CTC)5
3	MK	Mevalonate kinase	K00869	EC:2.7.1.36	DN54851_c0_g1_i6	1536	1483 - 1536	(TG)9aatgagtaat gtaatgca atgtaa(TG)6
4	ispE	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase	K00919	EC:2.7.1.148	DN36615_c0_g1_i1	1492	1 - 12	(G)12
5	ispD	2-C-methyl-D-erythritol 4-phosphate cytidyltransferase	K00991	EC:2.7.7.60	DN42987_c0_g2_i1	1661	248 - 262	(TCT)5
6	dxs	1-deoxy-D-xylulose-5-phosphate synthase	K01662	EC:2.2.1.7	DN30805_c0_g1_i1	2495	171 - 158	(GT)7
7	idi	Isopentenyl-diphosphate Delta-isomerase	K01823	EC:5.3.3.2	DN41437_c0_g1_i1	1325	80 - 94	(GCT)5
8	ispH	4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase	K03527	EC:1.1.7.4	DN56344_c1_g1_i2	1659	63 - 100	(TC)19
9	PCYOX1	Prenylcysteine oxidase / farnesylcysteine lyase	K05906	EC:1.8.3.5 1.8.3.6	DN45249_c0_g1_i1	1910	82 - 99	(CCA)6
10	GGPS	Geranylgeranyl diphosphate synthase	K13789	EC:2.5.1.1 2.5.1.10 2.5.1.29	DN44310_c0_g1_i1	1757	1481 - 1495	(AGA)5
11	GPS	Geranyl diphosphate synthase	K14066	EC:2.5.1.1	DN35063_c0_g1_i1	1804	1764 - 1802	(TTA)13
12	FLDH	NAD+-dependent farnesol dehydrogenase	K15891	EC:1.1.1.354	DN39038_c0_g1_i1	1020	57 - 71	(CAT)5

با پوشش کلی تولید کند. بنابراین تخمین فراوانی تکرارهای ریزماهوره‌ها و ویژگی‌های آن‌ها کمتر دچار اریب می‌گردد. شناسایی تنوع ریزماهوره‌ها در ناحیه آگزونی می‌تواند در فرآیندهای فیزیولوژیکی و نمو گیاه کارایی داشته باشد. نتایج این پژوهش برای نخستین بار توسعه نشانگرهای EST-SSR این گیاه را امکان‌پذیر ساخته است. این نشانگرها می‌توانند برای انجام تجزیه و تحلیل‌های نقشه‌های پیوستگی و ساخت نقشه‌های ژنتیکی با تراکم بالا در زنیان به‌کار روند. شناسایی ریزماهوره‌های موجود در توالی ژن‌های مرتبط به مسیر بیوسنتز متابولیت‌های ثانویه در ترنسکرپت گیاه دارویی زنیان می‌تواند در شناسایی عوامل موثر تنظیم بیان ژن‌های سنتز ترکیبات دارویی این گیاه بسیار موثر باشد.

وجود این ریزماهوره‌ها در ترنسکرپت‌های مرتبط به سنتز متابولیت‌های ثانویه گیاه دارویی زنیان و مطالعه تنوع آن‌ها می‌تواند در شناسایی عوامل مؤثر در سنتز ترکیبات دارویی گیاه بسیار مؤثر باشد. می‌توان از ریزماهوره‌های موجود در توالی‌های ژنومی، در به‌دست آوردن تأثیر مستقیم آن‌ها در سازماندهی ژنومی، نو ترکیبی، تنظیم بیان ژن، تنوع ژنتیکی کمی و تکامل ژن‌ها استفاده کرد (Kaliswamy *et al.*, 2015).

داده‌های ترنسکرپتومی منبع ارزان و روش سریعی جهت شناسایی نشانگرهای ریزماهوره در مقیاس گسترده می‌باشد. در مقایسه با ESTها که به‌طور عمومی بخش کوچکی از ژنوم بیان شده را شامل می‌شود، توالی‌های ترنسکرپتوم می‌تواند تعداد زیادی نشانگرهای ریزماهوره

References

- Alagna, F., D'Agostino, N., Torchia, L., Servili, M., Rao, R., Pietrella, M., Giuliano, G., Chiusano, M.L., Baldoni, L. and Perrotta, G. (2009). Comparative 454 pyrosequencing of transcripts from two olive genotypes during fruit development. *BMC Genomics* **10**: 399.
- Angeloni, F., Wagemaker, C., Jetten, M., Op den Camp, H., Janssen-Megens, E., Francoijs, K.J., Stunnenberg, H. and Ouborg, N. (2011). De novo transcriptome characterization and development of genomic tools for *Scabiosa columbaria* L. using next-generation sequencing techniques. *Molecular Ecology Resources*, **11**: 662-674.
- Bahmani, K., Izadi-Darbandi, A., Jafari, A.A., Noori, S.A.S. and Farajpour, M. (2012). Assessment of genetic diversity in Iranian fennels using ISSR markers. *Journal of Agricultural Science*, **4**: 79.
- Bairwa, R., Sodha, R. and Rajawat, B. (2012). *Trachyspermum ammi*. *Pharmacognosy Reviews*, **6**: 56-60.
- Bajgain, P., Richardson, B.A., Price, J.C., Cronn, R.C. and Udall, J.A. (2011). Transcriptome characterization and polymorphism detection between subspecies of big sagebrush (*Artemisia tridentata*). *BMC Genomics*, **12**: 370.
- Barakat, A., DiLoreto, D.S., Zhang, Y., Smith, C., Baier, K., Powell, W.A., Wheeler, N., Sederoff, R. and Carlson, J.E. (2009). Comparison of the transcriptomes of American chestnut (*Castanea dentata*) and Chinese chestnut (*Castanea mollissima*) in response to the chestnut blight infection. *BMC Plant Biology*, **9**: 51.
- Beier, S., Thiel, T., Münch, T., Scholz, U. and Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics*, **33**: 2583-2585.
- Bhandawat, A., Singh, G., Raina, A.S., Kaur, J. and Sharma, R.K. (2016). Development of genic SSR marker resource from RNA-Seq data in *Dendrocalamus latiflorus*. *Journal of Plant Biochemistry and Biotechnology*, **25**: 179-190.
- Blanca, J., Cañizares, J., Roig, C., Ziarsolo, P., Nuez, F. and Picó, B. (2011). Transcriptome characterization and high throughput SSRs and SNPs discovery in *Cucurbita pepo* (Cucurbitaceae). *BMC Genomics*, **12**: 104.

- Bräutigam, A., Mullick, T., Schliesky, S. and Weber, A.P.** (2011). Critical assessment of assembly strategies for non-model species mRNA-Seq data and application of next-generation sequencing to the comparison of C3 and C4 species. *Journal of Experimental Botany*, **62**: 3093-3102.
- Chen, F., Tholl, D., Bohlmann, J. and Pichersky, E.** (2011). The family of terpene synthases in plants: a mid-size family of genes for specialized metabolism that is highly diversified throughout the kingdom. *The Plant Journal*, **66**: 212-229.
- Dassanayake, M., Haas, J., Bohnert, H. and Cheeseman, J.** (2009). Shedding light on an extremophile lifestyle through transcriptomics. *New Phytologist*, **183**: 764-775.
- Farajpour, M., Ebrahimi, M., Amiri, R., Noori, S., Sanjari, S. and Golzari, R.** (2011). Study of genetic variation in yarrow using inter-simple sequence repeat (ISSR) and random amplified polymorphic DNA (RAPD) markers. *African Journal of Biotechnology*, **10**: 11137-11141.
- Fu, N., Wang, Q. and Shen, H.L.** (2013). De novo assembly, gene annotation and marker development using Illumina paired-end transcriptome sequences in celery (*Apium graveolens* L.). *PLoS One*, **8**: e57686.
- Futschik, A. and Schlötterer, C.** (2010). The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics*, **186**: 207-218.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R. and Zeng, Q.** (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**: 644-652.
- Guo, S., Liu, J., Zheng, Y., Huang, M., Zhang, H., Gong, G., He, H., Ren, Y., Zhong, S. and Fei, Z.** (2011). Characterization of transcriptome dynamics during watermelon fruit development: sequencing, assembly, annotation and gene expression profiles. *BMC Genomics*, **12**: 454.
- Hao, D., Ma, P., Mu, J., Chen, S., Xiao, P., Peng, Y., Huo, L., Xu, L. and Sun, C.** (2012). De novo characterization of the root transcriptome of a traditional Chinese medicinal plant *Polygonum cuspidatum*. *Science China Life Sciences*, **55**: 452-466.
- Heidari, E.F., Rahimmalek, M., Mohammadi, S. and Ehtemam, M.H.** (2016). Genetic structure and diversity of ajowan (*Trachyspermum ammi*) populations based on molecular, morphological markers, and volatile oil content. *Industrial Crops and Products*, **92**: 186-196.
- Hiremath, P.J., Farmer, A., Cannon, S.B., Woodward, J., Kudapa, H., Tuteja, R., Kumar, A., BhanuPrakash, A., Mulaosmanovic, B. and Gujaria, N.** (2011). Large-scale transcriptome analysis in chickpea (*Cicer arietinum* L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. *Plant Biotechnology Journal*, **9**: 922-931.
- Kaliswamy, P., Vellingiri, S., Nathan, B. and Selvaraj, S.** (2015). Microsatellite analysis in the genome of Acanthaceae: An in silico approach. *Pharmacognosy Magazine*, **11**: 152-156.
- Karimbeigi, H., Nazarian-Firouzabadi, F., Khademi, M. and Mousav, E.** (2016) Assessment of genetic diversity among some oilseed rape (*Brassica napus* L.) plants, using single sequence repeats (SSR) molecular markers. *Plant Genetic Researches*, **3(1)**: 45-56 (In Persian).
- Kumpatla, S.P. and Mukhopadhyay, S.** (2005). Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. *Genome*, **48**: 985-998.
- Li, D., Deng, Z., Qin, B., Liu, X. and Men, Z.** (2012). De novo assembly and characterization of bark transcriptome using Illumina sequencing and development of EST-SSR markers in rubber tree (*Hevea brasiliensis* Muell. Arg.). *BMC Genomics*, **13**: 192.
- Li, Y.C., Korol, A.B., Fahima, T. and Nevo, E.** (2004). Microsatellites within genes: structure, function, and evolution. *Molecular Biology and Evolution*, **21**: 991-1007.
- Li, Y., Sun, C., Luo, H., Li, X., Niu, Y. and Chen, S.** (2010). Transcriptome characterization for *Salvia miltiorrhiza* using 454 GS FLX. *Yao xue xue bao. Acta Pharmaceutica Sinica*, **45**: 524-529.
- Li, Y.C., Korol, A.B., Fahima, T., Beiles, A. and Nevo, E.** (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology*, **11**: 2453-2465.

- Lin, X., Zhang, J., Li, Y., Luo, H., Wu, Q., Sun, C., Song, J., Li, X., Wei, J. and Lu, A.** (2011). Functional genomics of a living fossil tree, Ginkgo, based on next-generation sequencing technology. *Physiologia Plantarum*, **143**: 207-218.
- Lu, F.H., Cho, M.C. and Park, Y.J.** (2012). Transcriptome profiling and molecular marker discovery in red pepper, *Capsicum annuum* L. TF68. *Molecular Biology Reports*, **39**: 3327-3335.
- Luo, H., Sun, C., Sun, Y., Wu, Q., Li, Y., Song, J., Niu, Y., Cheng, X., Xu, H. and Li, C.** (2011). Analysis of the transcriptome of *Panax notoginseng* root uncovers putative triterpene saponin-biosynthetic genes and genetic markers. *BMC Genomics*, **12**: S5.
- Mirzahosseini, S.M., Noori, S.A.S., Amanzadeh, Y., Javid, M.G. and Howyzeh, M.S.** (2017). Phytochemical assessment of some native ajowan (*Trachyspermum ammi* L.) ecotypes in Iran. *Industrial Crops and Products*, **105**: 142-147.
- Modareskia, M., Darvishzadeh, R., Hassani, A. and Kholghi, M.** (2013). Molecular diversity within and between Ajowan (*Carum copticum* L.) populations based on inter simple sequence repeat (ISSR) markers. *Journal of Plant Molecular Breeding*, **1**: 51-62.
- Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C. and Kanehisa, M.** (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research*, **35**: W182-W185.
- Noori, S.A.S., Norouzi, M., Karimzadeh, G., Shirkoob, K. and Niazi, M.** (2017). Effect of colchicine-induced polyploidy on morphological characteristics and essential oil composition of ajowan (*Trachyspermum ammi* L.). *Plant Cell, Tissue and Organ Culture (PCTOC)*, **130(3)**: 543-551.
- Parida, S.K., Pandit, A., Gaikwad, K., Sharma, T.R., Srivastava, P.S., Singh, N.K. and Mohapatra, T.** (2010). Functionally relevant microsatellites in sugarcane unigenes. *BMC Plant Biology*, **10**: 251.
- Primmer, C.R.** (2009). From conservation genetics to conservation genomics. *Annals of the New York Academy of Sciences*, **1162**: 357-368.
- Sadati, S., Sadat-Noori, S.A., Ramshini, H., Soltani, E. and Foghi, B.** (2016). Towards conservation and breeding of Ajowan (*Trachyspermum ammi*) by assessing ISSR, morphological traits and germination variability. *Iranian Journal of Genetics and Plant Breeding*, **5**: 8-21.
- Sargazi, A.** (2016). *Genetic diversity of some population of medicinal Ajowan (Trachyspermum copticum) using RAPD marker*. M.Sc. Thesis. Faculty of Agriculture, University of Zabol, Zabol, Iran (In Persian).
- Shuorvazdi, A., Mohammadi, S.A., Norozi, M. and Sadeghzadeh, B.** (2014) Molecular analysis of genetic diversity and relationships of barley landraces based on microsatellite markers. *Plant Genetic Researches*, **1(1)**: 51-64 (In Persian).
- Soltani Howyzeh, M., Sadat Noori, S.A. and Shariati, J.V.** (2018a). Essential oil profiling of Ajowan (*Trachyspermum ammi*) industrial medicinal plant. *Industrial Crops and Products*, **119**: 255-259.
- Soltani Howyzeh, M., Sadat Noori, S.A., Shariati J.V. and Niazi, M.** (2018b). Essential oil chemotype of Iranian Ajowan (*Trachyspermum ammi* L.). *Journal of Essential Oil Bearing Plants*, **21**: 273-276.
- Strickler, S.R., Bombarely, A. and Mueller, L.A.** (2012). Designing a transcriptome next-generation sequencing project for a nonmodel plant species1. *American Journal of Botany*, **99**: 257-266.
- Vasemägi, A., Nilsson, J. and Primmer, C.R.** (2005). Expressed sequence tag-linked microsatellites as a source of gene-associated polymorphisms for detecting signatures of divergent selection in Atlantic salmon (*Salmo salar* L.). *Molecular Biology and Evolution*, **22**: 1067-1076.
- Wei, W., Qi, X., Wang, L., Zhang, Y., Hua, W., Li, D., Lv, H. and Zhang, X.** (2011). Characterization of the sesame (*Sesamum indicum* L.) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. *BMC Genomics*, **12**: 451.
- Xu, M., Liu, X., Wang, J.W., Teng, S.Y., Shi, J.Q., Li, Y.Y. and Huang, M.R.** (2017). Transcriptome sequencing and development of novel genic SSR markers for *Dendrobium officinale*. *Molecular Breeding*, **37**: 18.

Zhang, J., Liang, S., Duan, J., Wang, J., Chen, S., Cheng, Z., Zhang, Q., Liang, X. and Li, Y. (2012). De novo assembly and characterisation of the transcriptome during seed development, and generation of genic-SSR markers in peanut (*Arachis hypogaea* L.). *BMC Genomics*, **13**: 90.

Large Scale Identification of SSR Molecular Markers in Ajowan (*Trachyspermum ammi*) Using RNA Sequencing

Mehdi Soltani Howyzeh^{1,2}, Seyed Ahmad Sadat Noori^{3,*}, Vahid Shariati⁴ and Mahboubeh Amiripour¹

- 1- Former Ph.D. Student, Department of Agronomy and Plant Breeding Sciences, College of Aburaihan, University of Tehran, Tehran, Iran
- 2- Instructor, Department of Genetic and Plant Breeding, Islamic Azad University Ahvaz Branch, Ahvaz, Iran
- 3- Professor, Department of Agronomy and Plant Breeding Sciences, College of Aburaihan, University of Tehran, Tehran, Iran
- 4- Assistant Professor, Department of Plant Molecular Biotechnology and NIGEB Genome Center, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran

(Received: January 6, 2018 – Accepted: May 19, 2018)

Abstract

The medicinal plant, *Trachyspermum ammi* is a rich source of active pharmaceutical ingredients with pharmaceutical effects. Microsatellite markers play a key role in the genome and gene expression, especially in secondary metabolite biosynthesis in medicinal plants. For the first time, transcriptome sequencing of this herb medicine was carried out to identify the microsatellite markers of this species. After pair-end sequencing with the Illumina HiSeq 2000 platform, the quality of the reads was evaluated by FastQC software, trimming was performed by Trimmomatic software and *De novo* assembly was done with Trinity software. In this study, 11,468 unitranscripts (7913 unigenes) were found to contain 13593 potential microsatellites. The most abundant microsatellite types were di-nucleotide (67%) and tri-nucleotide (24%). Also, six repeated SSRs were the most abundant repeats. The predominant sequence was AG / CT (31%). Sixty-five percent of SSRs were belonged to class II (10-20 nucleotides) and 35% to class I (more than 20 nucleotides). The frequency of SSRs found to be approximately one per 10.1 kB of assembled sequence. More than 57 percent of unigenes containing SSRs were blasted with carrot genome. This showed that *T. ammi* was an Apiaceae family member and had a high similarity to the carrot genome. A total of 3437 unigenes (43%) were categorized functionally, which among them 2219 unigenes (64.6%) belonged to the "metabolic process" category and 71 unigenes (1.2%) were assigned to the "secondary metabolic process". In this study, 12 genes were detected in the terpenoid backbone biosynthesis pathway, that their transcripts were containing a microsatellite. These SSRs probably contribute to the genes expression and the biosynthesis of metabolites, especially secondary metabolites. The development of these markers can be used for future studies of marker-assisted selection, genetic diversity and construct genetic maps in this medicinal plant.

Keywords: Apiaceae, Terpenoids, Transcriptome, Secondary metabolic

* Corresponding Author, E-mail: noori@ut.ac.ir