

شناسایی و تعیین خصوصیات میکرو RNAهای محافظت شده *Coriandrum sativum* L. با استفاده از داده‌های توالی‌یابی نسل جدید

رضا میردریکوند^{۱*}، سیدسجاد سهرابی^۲، سیدمحسن سهرابی^۲، کامران سمیعی^۳

۱. استادیار، گروه زراعت و اصلاح نباتات، واحد خرم‌آباد، دانشگاه آزاد اسلامی، خرم‌آباد، ایران

۲. محقق، باشگاه پژوهشگران جوان و نخبگان، دانشگاه آزاد اسلامی، خرم‌آباد، ایران

۳. استادیار، گروه کشاورزی، واحد کنگاور، دانشگاه آزاد اسلامی، کنگاور، ایران

(تاریخ دریافت: ۹۷/۶/۱۴ - تاریخ تصویب: ۱۳۹۸/۲/۱)

Identification and characterization of conserved miRNAs of *Coriandrum sativum* L. using next-generation sequencing data

Reza Mir Drikvand^{1*}, Seyed Sajad Sohrabi², Seyed Mohsen Sohrabi², Kamran Samiei³

1. Assistant Professor, Department of Agronomy and Plant Breeding, Khorramabad Branch, Islamic Azad University, Khorramabad, Iran.

2. Researcher, Young Researchers and Elite Club, Khorramabad branch, Islamic Azad University, Khorramabad, Iran.

3. Assistant Professor, Department of Agronomy and Plant Breeding, Kangavar Branch, Islamic Azad University, Iran.

(Received: Sep. 5, 2018 - Accepted: Apr. 21, 2019)

Abstract

MicroRNAs (miRNAs) are a class of small and noncoding RNAs with length of 18-24 nucleotides that control the expression of target genes at the transcriptional and post-transcriptional levels in plants. The miRNAs play an important role in different processes such as growth and development, cell proliferation and response to stresses in plants. Coriander or *Coriandrum sativum* L. is a plant of Apiaceae family with different nutritional and pharmaceutical applications. Up to now, the genome of this plant has not been sequenced and there is no report of miRNAs identification has been recorded for it. The present study was performed to identify the conserved miRNAs and their target genes in transcriptome of coriander plant. Firstly, Transcriptome of seed and leaf tissues was assembled and non-coding transcripts were identified and considered as miRNA precursor. Finally, among candidate sequences, three miRNAs named csa-miR162, csa-miR169 and csa-miR399 belong to three conserved families were identified after strict filtering. Identified miRNAs showed differential expression between seed and leaf tissues and also role of their target genes in different biological processes was confirmed. In general, given the regulatory roles of identified miRNAs on broad spectrum of gene networks and biological processes of coriander plant in the present study, these miRNAs can be used as candidate genes to improve qualitative and quantitative yield and resistance to different stresses in this plant.

Keywords: Coriander plant, miRBase, miRNA, Secondary Structure, Transcriptome

چکیده

میکرو RNAها (miRNAs) گروهی از مولکول‌های RNA کوچک و غیر کدکننده با طولی حدود ۱۸-۲۴ نوکلئوتید هستند که بیان ژن‌های هدف خود را در سطوح مختلف رونویسی و پس از رونویسی در گیاهان کنترل می‌کنند. میکرو RNAها در فرآیندهای مختلفی مانند رشد و نمو، فرآیندهای زیستی، تکثیر سلولی و پاسخ به تنش‌ها در گیاهان نقش مهمی بازی می‌کنند. گشنیز زراعی با نام علمی (*Coriandrum sativum* L.) گیاهی از خانواده چتریان (Apiaceae) است که دارای کاربردهای غذایی و دارویی مختلفی است. ژنوم این گیاه تاکنون توالی‌یابی نشده است و هیچ‌گونه گزارشی از شناسایی میکرو RNAها برای آن ثبت نشده است. مطالعه حاضر به منظور شناسایی میکرو RNAهای محافظت شده بالقوه و ژن‌های هدف آنها در ترنسکرپتوم گیاه گشنیز صورت گرفت. ابتدا ترنسکرپتوم بافت‌های بذر و برگ این گیاه سرهم‌بندی شد و رونوشت‌های غیر کدکننده به پروتئین شناسایی و به‌عنوان توالی‌های کاندید پیش‌ساز میکرو RNA در نظر گرفته شدند. در نهایت از بین توالی‌های کاندید سه میکرو RNA با نام‌های csa-miR162، csa-miR169 و miR169 و csa-miR399 متعلق به سه خانواده محافظت شده میکرو RNA پس از اعمال فیلترهای سخت‌گیرانه شناسایی شد. میکرو RNAهای شناسایی شده دارای بیان متفاوتی در بافت‌های بذر و برگ بودند و نقش ژن‌های هدفشان در فرآیندهای مختلف زیستی نیز مورد تأیید قرار گرفت. در مجموع با توجه به اینکه میکرو RNAهای شناسایی شده در مطالعه حاضر دارای نقش تنظیمی بر طیف وسیعی از شبکه‌های ژنی و فرآیندهای زیستی مختلف در گیاه گشنیز هستند، می‌توان از آنها به‌عنوان ژن‌های کاندید در بهبود عملکرد کمی و کیفی و همچنین مقاومت به تنش‌های مختلف در این گیاه بهره برد.

واژه‌های کلیدی: ترنسکرپتوم، ساختار ثانویه، گیاه گشنیز، میکرو RNA، miRBase

al., 2011; Darughe et al., 2012; de Almeida Freires et al., 2014; Laribi et al., 2015; Mandal and Mandal, 2015).
 با توجه به این‌که تولید متابولیت‌های ثانویه در گیاهان و سایر موجودات زنده، از طریق شبکه پیچیده‌ای از ژن‌ها کنترل می‌شود، استفاده از روش‌های اصلاحی سنتی به‌تنهایی برای افزایش یا کاهش این متابولیت‌ها کارساز نبوده و انتخاب مناسبی نخواهند بود. به همین علت در سال‌های اخیر، با پیشرفت‌های قابل توجهی که در ژنومیکس گیاهی و حجم اطلاعات موجود بوجود آمده است، بررسی و شناسایی ژن‌های دخیل در بیوستنز متابولیت‌های ثانویه و همچنین شبکه‌های تنظیمی آن‌ها مورد توجه قرار گرفته است و با سرعت بالایی در حال انجام است. توالی‌یابی با کارایی بالا (High-Throughput Sequencing)، یکی از جدیدترین روش‌های ژنومیکس است که با تولید داده‌های ژنومیکس در مقیاس بالا راه را برای شناسایی ژن‌ها و همچنین بازسازی شبکه‌های ژنی هموار کرده است (Hao et al., 2012; Goossens et al., 2003).

میکرو RNAها (miRNAs) عناصر زیستی کلیدی و مؤثری در تنظیم شبکه‌های ژنی در موجودات مختلف هستند (Li et al., 2010; Tian et al., 2015). این عناصر زیستی گروهی از RNAهای غیرکدکننده هستند که از طریق فرآیند تداخل RNA (RNA interference)، بیان بسیاری از ژن‌ها را در سطوح مختلف رونویسی، پردازش، پایداری RNA و ترجمه تنظیم می‌کنند. مکانیسم عمل میکرو RNAها از طریق جفت‌شدگی با نواحی مکمل خود در توالی هدف است. به‌طور کلی، یک ژن میکرو RNA اولیه (pri-miRNA) گیاهی، معمولاً دارای یک ساختار ساقه-حلقه می‌باشد که در اولین مرحله از پردازش درون هسته‌ای، توسط آنزیم DCL1 (که یک RNase III است) برش خورده و ساختار دیگری به نام pre-miRNA را ایجاد

مقدمه

گشنیز زراعی با نام علمی (*Coriandrum sativum* L.) گیاهی دیپلوئید ($2n=2x=22$)، علفی، یک‌ساله و دگرگرده افشان از خانواده چتریان (Apiaceae) است که دارای دو گونه وحشی (*Corandrium tardylium*) و زراعی (*Coriandrum sativum* L.) می‌باشد (Diederichsen, 1996). هر چند کشورهای جنوب غرب آسیا و شمال آفریقا خاستگاه اصلی گشنیز به‌شمار می‌روند (Hedge and Lamond, 1972)، با این حال کشت تجاری این گیاه در کشورهای اتحاد جماهیر شوروی سابق، مجارستان، هلند، رومانی، جمهوری چک، اسلواکی، مراکش، کانادا، هند، پاکستان، ایران، ترکیه، مکزیک و آرژانتین رایج است (López et al., 2008). در کشور ایران نیز، گیاه گشنیز در استان‌های همدان، قزوین، آذربایجان غربی و شرقی، کرمان، کرمانشاه، بوشهر، سیستان و بلوچستان و یزد و دیگر نقاط کشت و کار می‌شود.

گیاه گشنیز به صورت سبزی تازه و همچنین به عنوان ادویه‌ی عطر و طعم دهنده در صنایع غذایی مورد استفاده قرار می‌گیرد. از میوه‌ی این گیاه برای استخراج روغن در صنایع عطرسازی استفاده می‌شود. روغن استخراج شده از میوه‌ی گشنیز در صنایع داروسازی برای تولید داروهای خلط‌آور، ضدنفخ، ضداسپاسم، ضد استرس، ضد فراموشی و آنتی‌اکسیدان استفاده می‌شود. تفاله‌ی حاصل از گیاه گشنیز برای تامین فیبر و پروتئین دام‌ها کاربرد دارد (Weiss, 2002; Marangoni and Moura, 2011; Mandal and Mandal, 2015). اسانس گیاه گشنیز دارای ویژگی‌های بیولوژیکی مختلفی است که از آن جمله می‌توان به ویژگی ضد باکتریایی، ضد قارچی، ضد سرطانی، ضد دیابت، ضد جهش و آنتی‌اکسیدانی اشاره کرد (Lo Cantore et al., 2004; Gomez-Flores et al., 2010; Rajeshwari and Andallu, 2011; Silva et

می‌کند. در دومین مرحله از پردازش، حلقه انتهایی برش خورده و توالی دو رشته‌ای با طول متوسط ۲۴ - ۱۸ نوکلئوتید ایجاد می‌شود. در آخر با ورود ساختار دو رشته به سیتوپلاسم و اتصال و تشکیل کمپلکس خاموشی القا شده توسط RNA (RNA-induced silencing complex یا RISC)، ساختار تک رشته و بالغ میکرو RNA (Mature miRNA) ایجاد می‌شود (Castel and Martienssen, 2013). اولین میکرو RNA گیاهی در گیاه *Arabidopsis thaliana* شناسایی و تعیین خصوصیت شده است (Park et al., 2002). از آن زمان تاکنون تعداد زیادی میکرو RNA و ژن‌های هدفشان در گونه‌های مختلف گیاهی مانند *Malus domestica*، *Gossypium arboreum*، *Nicotiana tabacum* و *Panicum virgatum* غیره شناسایی شده است (Gleave et al., 2008; Frazier et al., 2010; Xie et al., 2010; Wang et al., 2012). به طور کلی، برای شناسایی میکرو RNA ها از سه روش همسانه‌سازی مستقیم، ژنتیک مستقیم و روش‌های بیوانفورماتیکی استفاده می‌شود؛ در حالی که برای شناسایی ژن‌های هدف آن‌ها سه روش بیوانفورماتیکی، بیوشیمیایی و روش‌های مبتنی بر امیکس مورد استفاده قرار می‌گیرد (Li et al., 2010). در این بین، روش‌های بیوانفورماتیکی سریع‌تر و بسیار کم هزینه‌تر از سایر روش‌ها بوده و از ویژگی محافظت‌شدگی بالای میکرو RNA ها برای شناسایی آن‌ها استفاده می‌کنند. روش‌های بیوانفورماتیکی، با استفاده از روش‌های مبتنی بر همولوژی و با مقایسه پایگاه داده‌های میکرو RNA های شناسایی شده (<http://www.mirbase.org>) با پایگاه‌های داده ارگانسیم موردنظر (مانند پایگاه داده‌های ترنسکرپتومی) به شناسایی میکرو RNA های محافظت شده بالقوه، می‌پردازند (Li et al., 2010; Gleave et al., 2008; Tian et al., 2015). با توجه به در دسترس نبودن ژنوم گیاه گشنیز و نبود هیچ گونه اطلاعاتی در مورد میکرو RNA های این گیاه، هدف از این تحقیق شناسایی میکرو RNA های محافظت شده گیاه گشنیز با استفاده از داده‌های توالی‌یابی نسل بعدی (Next Generation Sequencing, NGS) است. در این تحقیق برای اولین بار، میکرو RNA های محافظت شده از کتابخانه‌های ترنسکرپتومی گیاه گشنیز شناسایی، بررسی و تعیین خصوصیت شدند.

مواد و روش‌ها

جمع‌آوری و کیفیت‌سنجی داده‌ها

در این پژوهش، از پنج مجموعه داده (Dataset) که حاوی خوانش‌های کوتاه توالی‌یابی شده (Short Reads یا خوانش‌ها) از بافت‌های برگ و مراحل مختلف نموی بذر گشنیز بودند، استفاده شد. مجموعه داده‌های مورد نظر با کدهای دسترسی SRR1700630، SRR1700819، SRR1700873، SRR5410771 و SRR5410814 از پایگاه داده SRA سایت NCBI با استفاده از ابزار ASPERA دریافت شدند.

با توجه به تأثیر کیفیت داده‌های ورودی بر نحوه سرهم‌بندی (Assembly) ترنسکرپتوم، کیفیت اولیه توالی‌های خام (Raw reads) برای همه نمونه‌ها با استفاده از نرم‌افزار FastQC (Version 0.11.6) سنجیده شد. خطاهای حاصل از توالی‌یابی ناقص و همچنین توالی‌های آداپتور از خوانش‌های کوتاه به ترتیب با استفاده از نرم‌افزارهای RCorrector (Song and Florea, 2015) و Cutadapt (Martin, 2011) حذف شدند. در نهایت به منظور حذف توالی‌های با کیفیت پایین (با میانگین Phred score کمتر از ۳۰)، توالی‌های با میزان بالای نوکلئوتیدهای خوانش نشده (N) و توالی‌های با اندازه

RNA بالغ بود، از پایگاه mirbase موجود در سایت <http://www.mirbase.org> دریافت شدند. شناسایی رونوشت‌های با بیشترین شباهت با توالی‌های میکرو RNA بالغ با استفاده از ابزار BLASTn و پارامترهای $E\text{-value} \leq 10$ و $mismatch < 4$ صورت گرفت. میکرو RNAهای بالغ دریافت شده به‌عنوان ورودی (Query) و رونوشت‌های ایجاد شده به‌عنوان پایگاه مورد جستجو (Subject) مورد استفاده قرار گرفتند. آستانه‌ی طول میکرو RNAهای بالقوه بین ۱۸ تا ۲۴ نوکلئوتید تعیین شد. توالی‌های حاصل از هم‌ردیفی با ابزار BLASTn، علیه پایگاه داده پروتئین‌های غیرتکراری (Non-Redundant protein database, NR) با استفاده از ابزار BLASTx با $E\text{-value} \leq 0.001$ مورد هم‌ردیفی قرار گرفته و توالی‌های دارای نتیجه در این پایگاه حذف شدند.

پیش‌بینی ساختار دوم میکرو RNAها

شباهت بالای یک رونوشت با توالی میکرو RNA بالغ به‌تنهایی برای تعیین صحت یک میکرو RNA شناسایی شده کافی نیست. به همین علت، نوع تاخوردگی و میزان انرژی آزاد ساختار میکرو RNA اولیه با استفاده از ابزار mfold موجود در سایت <http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form> در رونوشت‌های حاصل، بررسی شد (Zuker, 2003). به‌منظور ایجاد اطمینان از قرارگیری توالی میکرو RNA بالغ در محل و جهت صحیح احتمالی خود در ساختار پیش‌ساز میکرو RNA، یک ناحیه‌ی ۱۰۰ نوکلئوتیدی در بالادست و پایین‌دست میکرو RNA شناسایی شده از توالی رونوشت‌های انتخابی با استفاده از ابزار Bedtools استخراج شد (Quinlan, 2014). توالی‌هایی که دارای ساختار ساقه-حلقه مناسب بوده و همچنین در یکی از بازوها شامل توالی میکرو RNA بالغ نیز بودند به‌عنوان میکرو RNAهای احتمالی مورد

کمتر از ۵۰ جفت باز از نرم‌افزار (version 0.27) Trimmomatic استفاده شد (Bolger et al., 2014).

سرهم‌بندی نوپدید (De novo assembly) خوانش‌ها

پس از فرآیند تریمنینگ (Trimming)، از خوانش‌های با کیفیت بالا (Clean reads) به‌عنوان توالی‌های ورودی برای سرهم‌بندی ترنسکرپتوم گشنیز، استفاده شد. سرهم‌بندی نوپدید با استفاده از نرم‌افزار CLC Genomics Workbench 7.5.0 طبق پارامترهای پیش‌فرض نرم‌افزار انجام شد (Matvienko, 2015). این نرم‌افزار از الگوریتم De Bruijn Graph برای سرهم‌بندی خوانش‌ها استفاده می‌کند. در این روش خوانش‌ها به قطعات کوچک‌تر یا k-mer تبدیل شده سپس از هم‌پوشانی k-merها برای سرهم‌بندی و تشکیل قطعات به‌هم‌پیوسته بزرگ‌تری به نام کانتیگ استفاده می‌شود (Grabherr et al., 2011). نرم‌افزار سرهم‌بندی قطعات را تا رسیدن به کانتیگ‌ها ادامه می‌دهد. همچنین به‌منظور حذف کانتیگ‌های اضافی از بسته نرم‌افزاری EvidentialGene (V2013.07.27) (<http://arthropods.eugenec.org/EvidentialGene>) استفاده شد (Gilbert, 2016).

کمیت سرهم‌بندی نوپدید با تعیین آماره‌های تعداد کانتیگ، طول بزرگ‌ترین کانتیگ، N50، طول متوسط کانتیگ‌ها توسط نرم‌افزار Transrate (v1.0.3) صورت گرفت (Smith-Unna et al., 2016).

شناسایی میکرو RNAهای محافظت شده

پس از سرهم‌بندی نوپدید و حذف کانتیگ‌های تکراری، از رونوشت‌های حاصل برای شناسایی میکرو RNAهای محافظت‌شده استفاده شد. به این منظور، کلیه‌ی توالی‌های میکرو RNA ثبت‌شده برای ۸۸ گونه گیاهی که شامل ۱۰۴۱۱ توالی میکرو

ابزار psRNATarget موجود در آدرس <http://plantgrn.noble.org/psRNATarget> و تحت پارامترهای پیش فرض این ابزار صورت گرفت (Dai and Zhao, 2011). مستندسازی و تفسیر عملکردی (Functional annotation) ژن های هدف میکرو RNA ها، با استفاده از ابزار BLASTx موجود در پایگاه NCBI انجام شد (Altschul *et al.*, 1990). رونوشت های هدف میکرو RNA ها علیه پایگاه NR با در نظر گرفتن حد آستانه $E-5$ value ≤ 1.0 مورد هم ردیفی محلی (Local Gene ontology, GO) ژن های هدف میکرو RNA ها با استفاده از نرم افزار Blast2GO صورت گرفت.

نتایج و بحث

کنترل کیفیت و سرهم بندی نوپدید خوانش ها امروزه با تلفیق روش های اصلاح کلاسیک و روش های مبتنی بر زیست فناوری و مهندسی ژنتیک، تشخیص جامع همه ژن ها و عوامل درگیر در بروز یک صفت، یکی از گام های مهم در طراحی برنامه های اصلاحی به شمار می رود. در سال های اخیر همزمان با ظهور فناوری های مبتنی بر NGS، تولید داده های امیکس در گیاهان به طور چشم گیری افزایش یافته است (Singh *et al.*, 2015)، به طوری که برای اغلب گیاهان مدل، ژنوم مرجع به صورت کامل مستندسازی شده یا در حال انجام است. در گیاهان غیر مدل اعم از گیاهان دارویی یا گیاهانی که دارای اندازه ژنوم بزرگی هستند (Unamba *et al.*, 2015)، سرهم بندی ترنسکرپتوم با استفاده از داده های حاصل از توالی یابی RNA (RNA sequencing, RNA-Seq)، امکان شناسایی اغلب ژن ها و عوامل دخیل در شبکه های زیستی از جمله میکرو RNA ها را، بدون نیاز به یک توالی ژنوم مرجع، ایجاد می کند (Duan *et al.*, 2012; Mathiyalagan *et al.*, 2013).

بررسی قرار گرفتند. به منظور تعیین صحت ساختار میکرو RNA های احتمالی از پارامترهای درصد GC/AU، حداقل انرژی آزاد تاخوردگی (Minimal free energy, MFE Adjusted minimal free energy, AMFE) و شاخص حداقل انرژی آزاد تاخوردگی (Minimal free energy index, MFEI) استفاده شد (روابط ۱، ۲ و ۳) (Xie *et al.*, 2005; Padmashree and Ramachandraswamy, 2016).

رابطه ۱) $-\Delta G$ = حداقل انرژی آزاد تاخوردگی (kcal/mol)

رابطه ۲) = حداقل انرژی آزاد تاخوردگی تصحیح شده $100 \times (\text{طول میرنای پیش ساز} / \text{MFE})$

رابطه ۳) = حداقل انرژی آزاد تاخوردگی (GC% تصحیح شده) / حداقل انرژی آزاد تاخوردگی

در نهایت، میکرو RNA های احتمالی با بررسی مواردی همچون: تاخوردگی مناسب و وجود ساختار ساقه حلقه، وجود توالی میکرو RNA ی بالغ در یک بازو در ناحیه ساقه، عدم شکستگی در حلقه میکرو RNA، حداکثر ۶ جفت نوکلئوتید غیر منطبق در ساختار میکرو RNA، محتوای ۳۰ تا ۷۰ درصدی AU، حداکثر اندازه فاصله در ساختار میکرو RNA ۳ نوکلئوتید و حداقل انرژی آزاد تاخوردگی، حداقل انرژی آزاد تاخوردگی تصحیح شده و میزان شاخص حداقل انرژی آزاد تاخوردگی بیشتر از ۸۵ کیلوکالری بر مول (Bonnet *et al.*, 2004; Zhang *et al.*, 2006b; Li *et al.*, 2014; Mehta *et al.*, 2016).

شناسایی و مستندسازی ژن های هدف میکرو RNA

شناسایی ژن های هدف میکرو RNA با استفاده از

در مطالعه حاضر، پس از ادغام تمام مجموعه داده‌های حاصل از RNA-Seq بافت‌های برگ و دانه گشنیز و به دنبال آن تعیین کیفیت اولیه و حذف خوانش‌های با کیفیت پایین، ۱۲۸۳۲۰۵۶۷ خوانش با کیفیت بالا حاصل شد (جدول ۱). درصد توالی‌های با کیفیت بالاتر از Q30، در تمامی نمونه‌ها بیشتر از ۹۵ درصد بود، که نشان دهنده کیفیت مناسب داده‌های ورودی برای سرهم‌بندی ترنسکریپتوم بود. از سرهم‌بندی خوانش‌های تمیز شده، ۵۱۹۹۶ رونوشت حاصل شد، که این تعداد، پس از حذف توالی‌های تکراری با استفاده از بسته نرم‌افزاری EvidentialGene، به ۳۲۱۴۲ رونوشت منحصر به فرد تقلیل یافت. مقایسه پارامترهای کمی و کیفی سرهم‌بندی نهایی و اولیه نشان داد که EvidentialGene موجب افزایش محسوس پارامتر N50 از ۱۴۸۶ به ۱۷۱۹ شده است (جدول ۲). تفاوت کارکرد حذف توالی‌های تکراری توسط EvidentialGene نسبت به سایر نرم‌افزارهای رایج مانند CD-HIT-EST در این مورد خلاصه می‌گردد که EvidentialGene نه تنها از شباهت بین توالی‌های خام استفاده می‌کند، بلکه قاب خوانش باز و توالی پروتئینی احتمالی را برای ایجاد خوشه‌های ژنی و در نهایت کاهش افزونگی کانتیگ‌ها (Redundancy) در نظر می‌گیرد (Gilbert, 2018). وجود کانتیگ‌های اضافی، فرآیند تجزیه و تحلیل‌های مقایسه‌ای با هدف تشخیص ژن‌های با بیان افتراقی (Differentially Expressed Genes, DEGs) را با خطا همراه می‌کند، از این رو انتخاب نرم‌افزار با کارایی مناسب در تشخیص کانتیگ‌های اضافی نقش موثری در افزایش اعتبار نتایج تحلیل داده‌های حاصل از توالی‌یابی RNA خواهد داشت (Ono et al., 2015). حجم نهایی ترنسکریپتوم و محتوای GC برای ترنسکریپتوم گشنیز به ترتیب ۳۹/۴۷ Mb و ۴۰ درصد تخمین زده شد. مقایسه پارامترهای

سرهم‌بندی ایجاد شده در مطالعه حاضر نسبت به مطالعه‌ای که تنها با استفاده از داده‌های حاصل از توالی‌یابی بافت بذر ایجاد شده بود (Galata et al., 2014)، نشان داد که هر چند تعداد رونوشت‌های سرهم‌بندی شده در مطالعه حاضر کمتر می‌باشد، با این حال میانگین طول رونوشت‌های ایجاد شده در این مطالعه (۱۲۲۸ bp) به طور چشم‌گیری بیشتر از میانگین طول رونوشت‌های موجود در مطالعه مذکور (۵۱۹ bp) است. از این رو می‌توان گفت که سرهم‌بندی ایجاد شده در مطالعه حاضر با بهره‌گیری از تمامی اطلاعات ترنسکریپتومی گشنیز ظرفیت بالایی جهت شناسایی ژن‌ها و عوامل ژنتیکی دخیل در بیوسنتز متابولیت‌های ثانویه و همچنین شبکه‌های تنظیمی آن‌ها در گشنیز دارد.

شناسایی میکرو RNAهای محافظت شده

هم‌ردیفی محلی میکرو RNAهای گیاهی با رونوشت‌های ایجاد شده در خلال سرهم‌بندی ترنسکریپتوم گشنیز، منجر به شناسایی ۴۴۴۷ رونوشت شد که با توالی‌های میکرو RNA بالغ گیاهی مطابقت داشتند. پس از حذف رونوشت‌های کدکننده پروتئین، (رونوشت‌هایی که دارای رکورد مشابهی با پایگاه پروتئینی NR بودند)، تعداد رونوشت میکرو RNAهای احتمالی به ۵۴۹ توالی رسید.

در مطالعه‌ای مشابه، که با هدف شناسایی میکرو RNAهای محافظت شده در گیاه شیرین برگ (*Stevia rebaudiana*) صورت گرفته بود، کارایی بالای استفاده از داده‌های RNA-Seq و سرهم‌بندی نوپدید آن‌ها برای شناسایی میکرو RNAهای محافظت شده نشان داده شده بود (Mehta et al., 2016). نتایج مطالعه‌ی مذکور نشان داد که از ۱۴۱۸۵۸ رونوشت ایجاد شده توسط فرآیند سرهم‌بندی نوپدید، ۴۶۱۶ رونوشت با میکرو RNAهای بالغ تطابق یافته که از این بین ۳۸۱ مورد به توالی‌های غیر کدکننده مربوط بودند. این در حالی

مطالعات مشابه نیز مورد تأیید قرار گرفته است (Mathiyalagan *et al.*, 2013; Gleave *et al.*, 2008; Mehta *et al.*, 2016; Xie *et al.*, 2012; Wang *et al.*, 2010). استفاده بهینه از روش های محاسباتی مستلزم افزایش کمیت و کیفیت داده های ورودی است؛ از این رو می توان گفت رهیافت مبتنی بر توالی یابی با کارایی بالا، با تولید حجم وسیعی از داده ها، مسیر را برای استفاده از روش های بیوانفورماتیکی هموارتر می سازد.

بود که کل توالی های (EST Expressed Sequence Tag) ثبت شده برای گیاه شیرین برگ تنها شامل ۳/۹ درصد از توالی های حاصل از RNA-Seq بود (Mehta *et al.*, 2016). شایان ذکر است که تاکنون تنها ۴ توالی EST برای گونه گشنیز انتشار یافته است، که نسبت به سایر گیاهان زراعی و برخی از گیاهان دارویی در این زمینه دارای فقر مطلق می باشد. کارایی استفاده از روش های محاسباتی برای شناسایی میکرو RNA های محافظت شده در سایر

جدول ۱. اطلاعات آماری مجموعه داده استفاده شده در مطالعه حاضر

شماره	نوع	مرحله ی	تعداد خوانش های	تعداد خوانش های تمیز	درصد خوانش های با کیفیت بیشتر
دسترسی	بافت	رشدی	خام	شده	از Q30
SRR1700630	بذر	اولیه	۱۶۹۲۲۳۴۰	۱۴۰۱۷۴۷۶	۹۶/۶۸
SRR1700819	بذر	میانی	۲۰۰۹۵۳۴۰	۱۶۳۶۳۸۳۰	۹۶/۷۹
SRR1700873	بذر	انتهایی	۱۵۹۱۲۹۲۰	۱۳۲۵۰۷۸۹	۹۷/۰۲
SRR5410771	برگ	-	۴۳۷۴۶۱۲۰	۳۹۸۹۸۰۳۰	۹۵/۳۶
SRR5410814	برگ	-	۴۹۱۶۳۱۰۸	۴۴۷۹۰۴۴۲	۹۵/۳۴

جدول ۲. مشخصات کمی و کیفی سرهم بندی ایجاد شده قبل و بعد از حذف کانتیگ های اضافی

سرهم بندی		ویژگی
قبل از حذف	بعد از حذف	
کانتیگ های اضافی	کانتیگ های اضافی	
۵۱۹۹۶	۳۲۱۴۲	تعداد رونوشت
۱۵۵	۱۹۲	کوتاه ترین اندازه رونوشت (جفت باز)
۱۵۴۳۲	۱۵۴۳۲	بلندترین اندازه رونوشت (جفت باز)
۴۷۱۱۲۰۳۰	۳۹۴۷۲۸۸۲	تعداد جفت باز
۹۰۵/۹۵	۱۲۲۸/۰۵	میانگین طول (جفت باز)
۳۰	۴	تعداد رونوشت با طول کمتر از ۲۰۰ جفت باز
۱۶۱۶۶	۱۵۶۸۱	تعداد رونوشت با طول بیشتر از ۱ کیلو جفت باز
۸	۸	تعداد رونوشت با طول بیشتر از ۱۰ کیلو جفت باز
۲۰۲۶۳	۲۰۰۵۱	تعداد رونوشت های دارای قاب خوانش باز
۶۸/۰۳	۷۱/۹۴	میانگین درصدی از طول رونوشت ها که با قاب خوانش باز پوشش می خورد
۳۶۶	۵۹۷	N90 (bp)
۸۸۵	۱۲۰۹	N70 (bp)
۱۴۸۶	۱۷۱۹	N50 (bp)
۲۱۲۷	۲۳۳۴	N30 (bp)
۳۴۸۶	۳۷۱۱	N10 (bp)
۴۰/۴۴	۴۰/۹۷	%GC

۰/۶۴، ۰/۶۲، ۰/۶۶ و ۰/۸۵ کیلوکالری بر مول گزارش شده است (Zhang et al., 2006b). میانگین پارامتر MFEI برای سه میکرو RNA شناسایی شده ۰/۹۳- کیلوکالری بر مول محاسبه شد که این میزان نیز صحت ساختار ثانویه توالی شناسایی شده را تأیید می‌نماید. تعداد میکرو RNA شناسایی شده در مطالعه حاضر با نتایج مطالعه ژانگ و همکاران (۲۰۰۶) تطابق دارد. در مطالعه ژانگ و همکاران (۲۰۰۶) احتمال حضور میکرو RNA در توالی‌های EST، ۰/۰۱ درصد گزارش شده است (Zhang et al., 2006a; Zhang et al., 2006b). از این رو می‌توان گفت که در مطالعه حاضر قرابت بالایی بین تعداد میکرو RNA شناسایی شده (۳ میکرو RNA) با تعداد مورد انتظار (۳ ~ ۳/۲۱ = ۳/۲۱ × ۰/۰۰۰۱ × ۳۲۱۴۲) وجود دارد. بررسی میزان بیان میکرو RNA شناسایی شده بر حسب تعداد خوانش‌ها در هر کیلوباز میلیون (Reads Per Kilobase Million, RPKM) در دو بافت برگ و بذر (شکل ۲) نشان داد که بیان سه میکرو RNA شناسایی شده در بافت برگ نسبت به بافت بذر بیشتر بوده و این دو بافت از نظر بیان سه میکرو RNA شناسایی شده در دو گروه کاملاً جدا قرار می‌گیرند. از سوی دیگر گروه بندی میکرو RNA های شناسایی شده نشان داد که از سه گزینه مورد بررسی csa-miR162 بیشترین میزان بیان را در هر سه دوره نمو بذر به خود اختصاص داده است، این در حالی بود که در بافت برگ csa-miR169 بیشترین میزان بیان را نشان داد.

miR169 یکی از محافظت شده ترین توالی‌های میکرو RNA در گیاهان به شمار می‌رود که در رشد و نمو گیاهان و پاسخ به تنش‌های محیطی نقش دارد (Li et al., 2017). این خانواده از میکرو RNA به‌عنوان بزرگترین گروه میکرو RNA در آرکیدوپسیس شناخته شده‌اند، با این حال تاکنون تنها برای برخی از اعضای این گروه عملکرد شناخته شده‌ای تصور می‌شود (Xu

با توجه به این نکته که انتخاب میکرو RNA های محافظت شده تنها بر اساس معیار همولوژی نتایج مطلوبی به‌همراه نخواهد داشت، ساختار ثانویه رونوشت‌های شامل میکرو RNA های احتمالی (۵۴۹ رونوشت) توسط وبسایت mfold مورد بررسی قرار گرفت (Hammond, 2015; Xie et al., 2010). ساختار ثانویه توالی‌های کاندید از نظر تناسب ساختار ساقه - حلقه و محل قرارگیری توالی میکرو RNA بالغ در ساختار ثانویه پیش‌بینی شده، در مرحله نخست کنترل شد. در گام دوم، توالی‌هایی که از نظر مقدار حداکثر فاصله و تعداد نوکلئوتید غیر منطبق در ساختار، نامطلوب شناسایی شده بودند، حذف شدند. از ۵۴۹ توالی مورد بررسی تنها ۱۲ رونوشت حاوی میکرو RNA بالغ دارای پارامترهای لازم بوده و از مابقی توالی‌ها در مراحل بعد چشم‌پوشی شد. در نهایت بر اساس پارامتر درصد AU و MFEI (جدول ۳)، سه میکرو RNA محافظت شده با پتانسیل بسیار بالا متعلق به خانواده‌های miR169، miR162 و miR399 شناسایی شد (شکل ۱).

در هر سه میکرو RNA شناسایی شده درصد A/U بیشتر از ۵۰ می‌باشد که این موضوع ثبات و صحت ساختار ثانویه پیش‌ساز میکرو RNA های شناسایی شده را تأیید می‌نماید. ساختار پیش ساز میکرو RNA یی که درصد A+U بالایی دارند نسبت به دیگر فرم‌های RNA غنی از باندهای هیدروژنی می‌باشند. باندهای هیدروژنی و جفت شدن بازها اغلب باعث پایداری ساختار سنجاق سری میکرو RNA ها می‌شوند (Zhang et al., 2006a; Zhang et al., 2006b). همچنین در مطالعات متعدد پارامتر MFEI به عنوان مناسب‌ترین معیار برای شناسایی میکرو RNA گزارش شده است (Gleave et al., 2008; Mandhan et al., 2006; Nakano et al., 2012). مقدار MFEI برای تشخیص و تمایز بین tRNA، rRNA، mRNA و microRNA، به ترتیب بیشتر از ۰/۵۹،

شده در مطالعه حاضر که با استفاده از ابزار psRNATarget صورت گرفت، نشان داد که در مجموع ۱۰ رونوشت توسط این میکرو RNAها هدف قرار داده می‌شوند. تعداد ژن‌های هدف برای میکرو RNAهای csa-miR169، csa-miR162 و csa-miR399 به ترتیب ۴، ۳ و ۳ رونوشت بود (شکل ۲). تخریب mRNA یا عدم ترجمه به پروتئین نتیجه انطباق توالی میکرو RNAی بالغ موجود در کمپلکس خاموشی با توالی mRNA ژن هدف می‌باشد. از ۱۰ ژن هدف ۸ ژن با تخریب شدن و ۲ ژن با عدم ترجمه شدن توسط ۳ میکرو RNA شناسایی شده تنظیم می‌شوند. بررسی ارتباط بین بیان میکرو RNAهای شناسایی شده با ژن‌های هدف در ترنسکریپتوم گشیز نشان داد که در اغلب موارد با افزایش بیان میکرو RNA، بیان ژن‌های هدف در بافت مورد نظر کاهش یافته و در مقابل هنگامی که بیان یک میکرو RNA در یک بافت نسبت به سایر میکرو RNAها کمتر بوده، به تناسب ژن‌های هدف آن نیز در همان بافت دارای بیان بیشتری بودند (شکل ۳).

گروه‌بندی کارکردی ژن‌های هدف با استفاده نرم‌افزار Blast2GO نشان داد که در زمینه عملکرد مولکولی (Molecular Function, MF) (شکل ۴-الف) گروه‌های "اتصال"، "انتقال دهنده‌ها" و "فعالیت فروکافتی" بیشترین فراوانی را به خود اختصاص داده‌اند. این در حالی بود که در زمینه فرآیند زیستی (Biological Processes, BP) (شکل ۴-ب) دسته "فرآیند سلولی" و "فرآیند متابولیکی" به ترتیب با فراوانی ۲۳ و ۱۵ درصد بیشتر از سایر گروه‌ها تحت تأثیر قرار می‌گیرند. "بخش‌های سلولی" و "سلول" به طور مشترک با ۳۲ درصد بیشترین اهداف میکرو RNAهای شناسایی شده در زمینه اجزای سلولی (Cellular Components, CC) (شکل ۴-ج) را به خود اختصاص دادند. با توجه به نتایج هستی‌شناسی

(*et al.*, 2014). نتایج بررسی منابع نشان داد که برخی از اعضای خانواده miR169 در تنش‌های خشکی (Zhang *et al.*, 2011)، سرما (Lee *et al.*, 2010)، شوری (Zhao *et al.*, 2009) و کمبودازت (Zhao *et al.*, 2011) افزایش بیان نشان داده و احتمالاً نقش مؤثری در تنظیم زمان گل‌دهی طی تنش‌های اکسیداتیو دارند (Xu *et al.*, 2014). بررسی مقادیر بیان میکرو RNAهای شناسایی شده نیز نشان داد که تقریباً در دو بافت مورد نظر تغییر بیان محسوسی برای csa-miR169 ثبت نشده است.

نتایج مطالعه حاضر تغییر بیان محسوسی برای csa-miR399 در بافت برگ نسبت به بافت بذر نشان نداد. در مطالعات متعدد نقش miR399 در تنظیم هومئوستازی فسفات به خوبی شرح داده شده است (Bari *et al.*, 2006; Pant *et al.*, 2008; Chiou *et al.*, 2006; Kim *et al.*, 2011) هر چند در برخی از مطالعات تغییر میزان بیان این خانواده از میکرو RNAها تحت تیمارهای دمایی نیز به اثبات رسیده است (Lee *et al.*, 2010).

بر خلاف دو میکرو RNAی csa-miR169 و csa-miR399 برای میکرو RNAی csa-miR162 در بافت بذر بیان بالاتری نسبت به بافت برگ گشیز مشاهده شد که این موضوع احتمالاً به حوزه وسیع فعالیت و میزان بالای بیان این میکرو RNA اشاره دارد. میکرو RNAی miR162 هر چند در بسیاری از گونه‌های گیاهی تاکنون شناسایی نشده است با این حال در گونه‌های گیاهی به‌عنوان یک عامل پاسخ به تنش‌های غیر زیستی مانند تنش خشکی و شوری شناخته می‌شود (Barciszewska-Pacak *et al.*, 2016).

شناسایی ژن‌های هدف میکرو RNAهای شناسایی شده

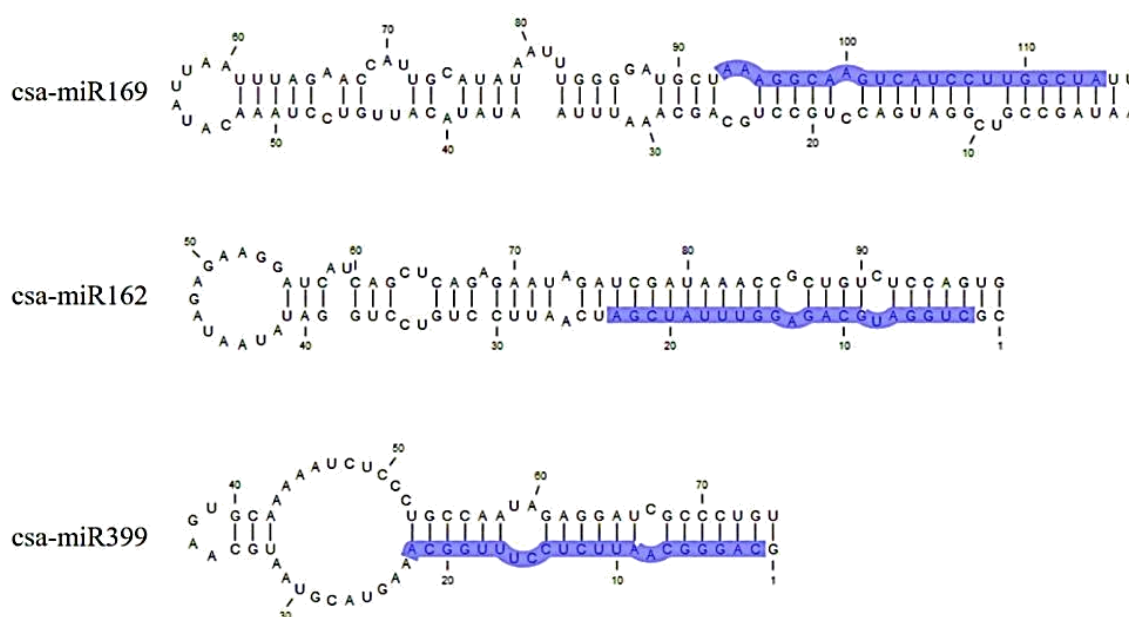
نتایج بررسی ژن‌های هدف میکرو RNAهای شناسایی

رویشی به زایشی، انتقال پیام و پاسخ به تنش‌های زیستی و غیرزیستی، نقش مهمی ایفا می‌کنند (Baldrich *et al.*, 2018; Chen, 2005; Rhoades *et al.*, 2002; Jung *et al.*, 2009; Jung *et al.*, 2013).

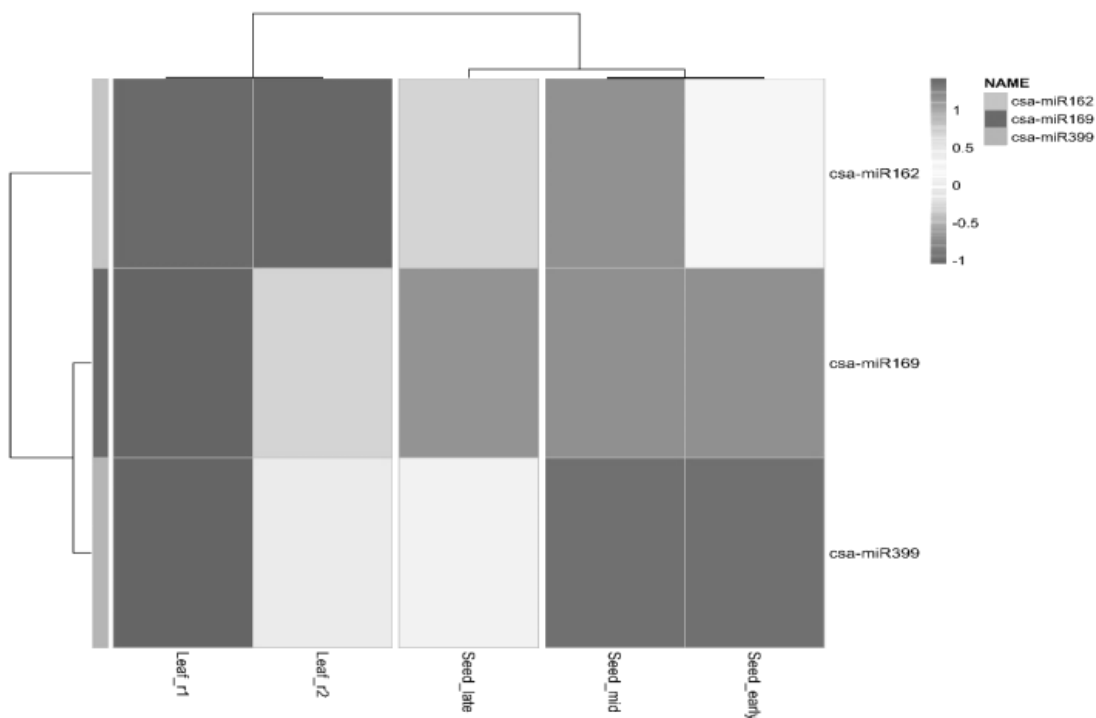
ژن‌های هدف می‌توان گفت، میکرو RNAهای شناسایی شده در مطالعه حاضر نیز مشابه با نتایج مطالعاتی که بر روی میکرو RNAهایی که تاکنون شناسایی شده‌اند، در اغلب فرآیندهای زیستی و متابولیکی همچون تمایز، رشد و نمو، انتقال از مرحله

جدول ۳. ویژگی‌های میکرو RNAهای شناسایی شده در گیاه گشنیز

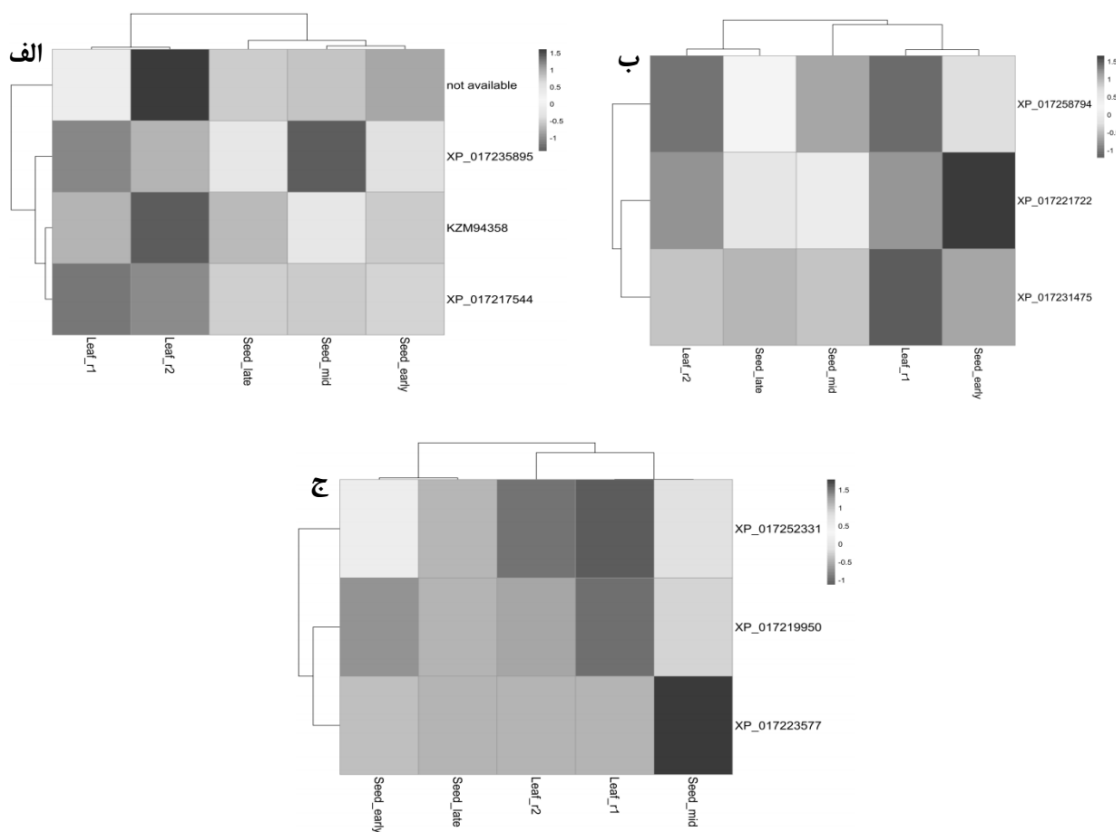
نام میکرو RNAی بالغ			ویژگی‌ها
csa-miR399	csa-miR162	csa-miR169	
evgLocus_26625	evgLocus_20661	evgLocus_16743	نام رونوشت
۲۴۱	۱۵۷۱	۷۸۹	طول رونوشت
۰/۶۳۰	۰/۶۱۸	۴/۸۸۶	RPKM رونوشت در بافت برگ
۰/۱۱۷	۱۶/۰۷۹	۰/۰۳۵	RPKM رونوشت در بافت بذر
۱۵۶..۱۳۶	۹۹۰..۹۷۰	۵۲۴..۵۰۳	مختصات توالی میکرو RNAی بالغ روی رونوشت
.	.	.	تعداد نوکلئوتیدهای غیر منطبق
۰/۰۰۰۰۸	۰/۰۰۰۰۸	۰/۰۰۰۰۳	ارزش مورد انتظار
۴۱/۱۰	۴۱/۱۰	۴۳/۰۰	bit score
۷۴	۹۸	۱۱۶	طول توالی پیش‌ساز میکرو RNA (جفت باز)
۵'-UGCCAAAGGA	۵'-UCGAUAAACCUC	۵'-UAGCCAAGGAUG	توالی میکرو RNA شناسایی شده (جفت باز)
GAAUUGCCCUG-۳'	UGCAUCCAG-۳'	ACUUGCCUUU-۳'	طول میکرو RNA شناسایی شده (جفت باز)
۲۱	۲۱	۲۲	درصد GC
۴۸/۶۴	۴۴/۸۹	۳۷/۰۶	درصد AU
۵۱/۳۵	۵۵/۱۰	۶۲/۹۳	MFE (kcal/mol)
۳۲/۰۰	۴۴/۸۰	۳۸/۶۰	AMFE
۴۳/۲۴	۴۵/۷۱	۳۲/۲۵	MFEI
۰/۸۸	۱/۰۱	۰/۸۹	



شکل ۱. ساختار پیش‌ساز برای سه میکرو RNA شناسایی شده در گشنیز. توالی میکرو RNAی بالغ با رنگ آبی مشخص شده است.

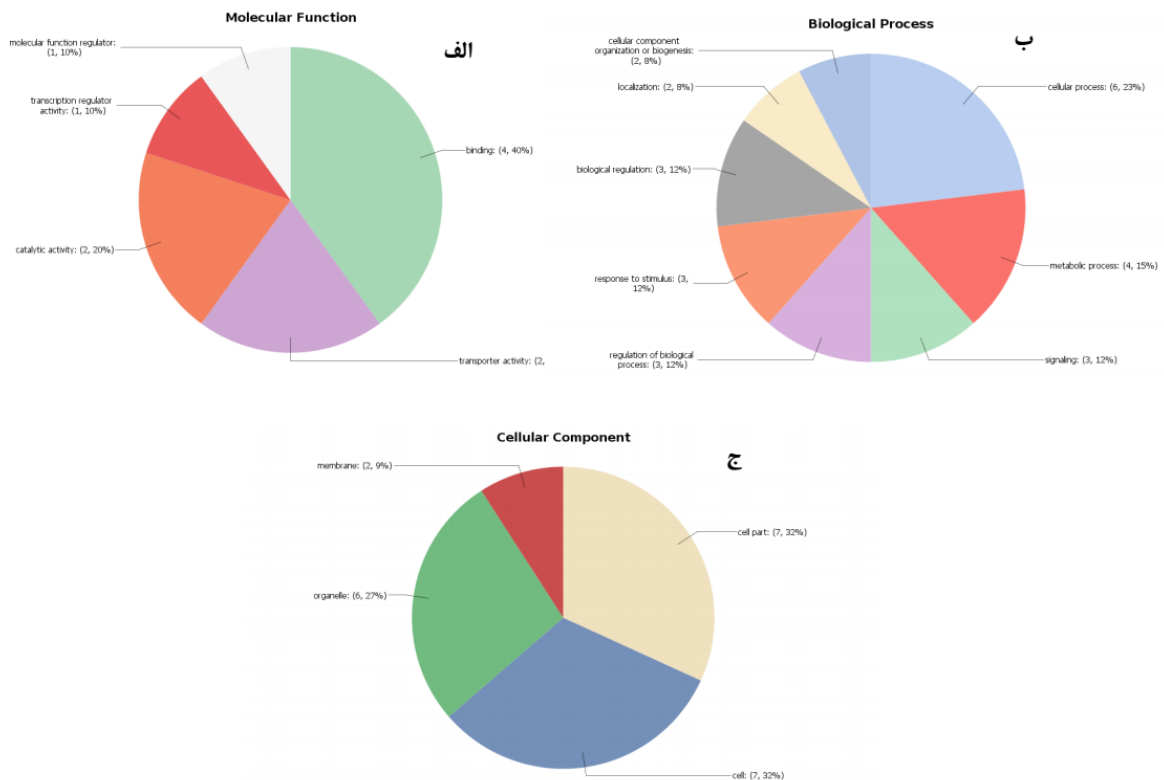


شکل ۲. نقشه حرارتی (Heat map) بیان میکرو RNAهای شناسایی شده در بافت های بذر و برگ گشنیز



شکل ۳. نقشه حرارتی بیان ژن های هدف میکرو RNAهای شناسایی شده در گشنیز.

الف، ب و ج: به ترتیب نقشه حرارتی بیان ژن های هدف csa-miR169، csa-miR399 و csa-miR162 را نشان می دهد.



شکل ۴. تحلیل هستی‌شناسی ژن‌های هدف پیش‌بینی‌شده برای میکرو RNAهای شناسایی‌شده در گیاه گشنیز. الف: عملکرد مولکولی؛ ب: فرآیندهای زیستی و ج: اجزای سلولی.

نتیجه‌گیری کلی

تولید متابولیت‌های ثانویه در گیاهان دارویی، از طریق شبکه پیچیده‌ای از ژن‌ها و عوامل کلیدی مانند میکرو RNAها کنترل می‌شود، از این رو استفاده از روش‌های اصلاح سنتی به‌تنهایی برای افزایش یا کاهش این متابولیت‌ها کارساز نبوده و انتخاب مناسبی نخواهند بود. گیاه گشنیز نیز از جمله گیاهان دارویی است که برنامه‌های اصلاحی آن به علت فقدان اطلاعات ژنتیکی با سرعت کم در حال پیشروی است. در مطالعه حاضر به منظور افزایش محتوای ژنتیکی گشنیز، تمامی اطلاعات ترنسکرپتومی موجود این گیاه به منظور دسترسی به یک رفرنس ترنسکرپتوم جامع سرهم‌بندی شد. نتایج نشان داد که ترنسکرپتوم ایجادشده در این پژوهش از نظر ویژگی‌های کمی و کیفی برتری قابل ملاحظه‌ای نسبت به مطالعات پیشین دارد، هر چند تا رسیدن به یک رفرنس ترنسکرپتوم جامع

مستندسازی شده فاصله دارد. رفرنس ترنسکرپتوم ایجاد شده در مطالعه حاضر منبع بسیار غنی برای بررسی و تعیین خصوصیات میکرو RNAهای محافظت شده گیاه گشنیز تشخیص داده شد، به طوری که برای اولین بار توالی سه میکرو RNAی محافظت‌شده در گیاه گشنیز به ثبت رسید. بررسی بیان میکرو RNAهای شناسایی شده و ژن‌های هدفشان در دو بافت بذر و برگ گشنیز ارتباط منطقی این عوامل کنترلی بر شبکه‌های زیستی گشنیز را اثبات نمود. در نهایت با توجه به نقش تنظیمی میکرو RNAهای شناسایی شده بر گروه وسیعی از عوامل ژنتیکی و تأثیر بالقوه آن‌ها در تغییر طیف وسیعی از ژن‌ها پایین‌دستی شبکه‌های ژنی، می‌توان از میکرو RNAهای شناسایی‌شده به‌عنوان ژن‌های کاندید در مبحث مهندسی ژنتیک با هدف افزایش تحمل به تنش‌های زیستی و غیر زیستی و همچنین ارتقای صفات کمی و کیفی گیاه گشنیز استفاده نمود.

اسلامی واحد خرم‌آباد انجام شده است، که بدین وسیله قدردانی می‌شود.

سپاسگزاری

مقاله حاضر با حمایت معاونت پژوهشی دانشگاه آزاد

REFERENCES

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*. 215 (3): 403-410.
- Baldrich P, Beric A, Meyers BC (2018) Despacito: the slow evolutionary changes in plant microRNAs. *Current Opinion in Plant Biology*. 42: 16-22.
- Barciszewska-Pacak M, Knop K, Jarmołowski A, Szweykowska-Kulińska Z (2016) Arabidopsis thaliana microRNA162 level is posttranscriptionally regulated via splicing and polyadenylation site selection. *Acta Biochimica Polonica*. 63 (4): 811-816.
- Bari R, Pant BD, Stitt M, Scheible WR (2006) PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. *Plant Physiology*. 141 (3): 988-999.
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30 (15): 2114-2120.
- Bonnet E, Wuyts J, Rouzé P, Van de Peer Y (2004) Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. *Bioinformatics*. 20 (17): 2911-2917.
- Castel SE, Martienssen RA (2013) RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nature Reviews Genetics*. 14 (2): 100-112.
- Chen X (2005) microRNA biogenesis and function in plants. *FEBS Letters*. 579 (26): 5923-5931.
- Chiou TJ, Aung K, Lin SI, Wu CC, Chiang SF, Su CL (2006) Regulation of phosphate homeostasis by microRNA in Arabidopsis. *Plant Cell*. 18(2): 412-421.
- Dai X, Zhao PX (2011) psRNATarget: a plant small RNA target analysis server. *Nucleic Acids Research*. 39 (s2): 155-159.
- Darughe F, Barzegar M, Sahari M (2012) Antioxidant and antifungal activity of Coriander (*Coriandrum sativum* L.) essential oil in cake. *International Food Research Journal*. 19 (3): 1253-1260.
- de Almeida Freires I, Murata RM, Furletti VF, Sartoratto A, de Alencar SM, Figueira GM, de Oliveira Rodrigues JA, Duarte MCT, Rosalen PL (2014) *Coriandrum sativum* L.(coriander) essential oil: antifungal activity and mode of action on *Candida* spp., and molecular targets affected in human whole-genome expression. *PLoS One*. 9 (6): e99086.
- Diederichsen A (1996) Coriander (*Coriandrum sativum* L.). Promoting the conservation and use of underutilized and neglected crops. International Potato Center, Lima, Peru/International Plant Genetic Resources Institute, Rome, Italy.
- Duan J, Xia C, Zhao G, Jia J, Kong X (2012) Optimizing de novo common wheat transcriptome assembly using short-read RNA-Seq data. *BMC Genomics*. 13 (1): 392.
- Frazier TP, Xie F, Freistaedter A, Burklew CE, Zhang B (2010) Identification and characterization of microRNAs and their target genes in tobacco (*Nicotiana tabacum*). *Planta*. 232 (6): 1289-1308.
- Galata M, Sarker LS, Mahmoud SS (2014) Transcriptome profiling, and cloning and characterization of the main monoterpene synthases of *Coriandrum sativum* L. *Phytochemistry*. 102: 64-73.
- Gilbert D (2016) Accurate and complete

- gene construction with EvidentialGene. F1000Res. 5.
- Gilbert DG (2018) Genes of the Pig, *Sus scrofa*, reconstructed with Evidential Gene. doi.org/10.1101/412130
- Gleave AP, Ampomah-Dwamena C, Berthold S, Dejnopratt S, Karunairetnam S, Nain B, Wang YY, Crowhurst RN, MacDiarmid RM (2008) Identification and characterisation of primary microRNAs from apple (*Malus domestica* cv. *Royal Gala*) expressed sequence tags. *Tree Genetics & Genomes*. 4 (2): 343-358.
- Gomez-Flores R, Hernández-Martínez H, Tamez-Guerra P, Tamez-Guerra R, Quintanilla-Licea R, Monreal-Cuevas E, Rodríguez-Padilla C (2010) Antitumor and immunomodulating potential of *Coriandrum sativum*. *Journal of Natural Products*. 3:54-63.
- Goossens A, Häkkinen ST, Laakso I, Seppänen-Laakso T, Biondi S, De Sutter V, Lammertyn F, Nuutila AM, Jung HJ, Park SJ, Kang H (2013) Regulation of RNA metabolism in plant development and stress responses. *Journal of Plant Biology*. 56: 123-129.
- Jung JH, Seo PJ, Park CM (2009) MicroRNA biogenesis and function in higher plants. *Plant Biotechnology Reports*. 3 (2): 111-126.
- Kim W, Ahn HJ, Chiou TJ, Ahn JH (2011) The role of the miR399-PHO2 module in the regulation of flowering time in response to different ambient temperatures in *Arabidopsis thaliana*. *Molecules and Cells*. 32 (1): 83-88.
- Laribi B, Kouki K, M'Hamdi M, Bettaieb T (2015) Coriander (*Coriandrum sativum* L.) and its bioactive constituents. *Fitoterapia*. 103: 9-26.
- Lee H, Yoo SJ, Lee JH, Kim W, Yoo SK, Fitzgerald H, Carrington JC, Ahn JH (2010) Genetic framework for flowering-time regulation by ambient temperature-responsive miRNAs in *Söderlund H, Zabeau M* (2003) A functional genomics approach toward the understanding of secondary metabolism in plant cells. *Proceedings of the National Academy of Sciences*. 100 (14): 8595-8600.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*. 29: 644-652.
- Hammond SM (2015) An overview of microRNAs. *Advanced Drug Delivery Reviews*. 87: 3-14.
- Hao DC, Chen SL, Xiao PG, Liu M (2012) Application of high-throughput sequencing in medicinal plant transcriptome studies. *Drug Development Research*. 73 (8): 487-498.
- Hedge I, Lamond J (1972) *Coriandrum* L. *Flora of Turkey*. 4: 330-331.
- Arabidopsis*. *Nucleic Acids Research*. 38 (9): 3081-3093.
- Li L, Xu J, Yang D, Tan X, Wang H (2010) Computational approaches for microRNA studies: a review. *Mammalian Genome*. 21 (1-2): 1-12.
- Li X, Hou Y, Zhang L, Zhang W, Quan C, Cui Y, Bian S (2014) Computational identification of conserved microRNAs and their targets from expression sequence tags of blueberry (*Vaccinium corybosum*). *Plant Signaling & Behavior*. 9 (9): e29462.
- Li Y, Zhao SL, Li JL, Hu XH, Wang H, Cao XL, Xu YJ, Zhao ZX, Xiao ZY, Yang N, Fan J, Huang F, Wang WM (2017) Osa-miR169 negatively regulates rice immunity against the blast fungus *magnaporthe oryzae*. *Frontiers in Plant Science*. 8: 2.
- Lo Cantore P, Iacobellis NS, De Marco A, Capasso F, Senatore F (2004) Antibacterial activity of *Coriandrum*

- sativum* L. and *Foeniculum vulgare* Miller var. *vulgare* (Miller) essential oils. *Journal of Agricultural and Food Chemistry*. 52 (26): 7862-7866.
- López PA, Widrlechner MP, Simon PW, Rai S, Boylston TD, Isbell TA, Bailey TB, Gardner CA, Wilson LA (2008) Assessing phenotypic, biochemical, and molecular diversity in coriander (*Coriandrum sativum* L.) germplasm. *Genetic Resources and Crop Evolution*. 55 (2): 247-275.
- Mandal S, Mandal M (2015) Coriander (*Coriandrum sativum* L.) essential oil: Chemistry and biological activity. *Asian Pacific Journal of Tropical Biomedicine*. 5 (6): 421-428.
- Mandhan V, Kaur J, Singh K (2012) smRNAome profiling to identify conserved and novel microRNAs in *Stevia rebaudiana* Bertoni. *BMC Plant Biology*. 12 (1): 197.
- Marangoni C, Moura NFd (2011) Sensory profile of Italian salami with coriander (*Coriandrum sativum* L.) essential oil. *Food Science and Technology*. 31 (1): 119-123.
- Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal*. 17(1): 10-12.
- Mathiyalagan R, Subramaniam S, Natarajan S, Kim YJ, Sun MS, Kim SY, Kim YJ, Yang DC (2013) In silico profiling of microRNAs in Korean ginseng (*Panax ginseng* Meyer). *Journal of Ginseng Research*. 37(2): 227-247.
- Matvienko M (2015) CLC Genomics Workbench. *Plant and Animal Genome Conference*. Qiagen Bioinformatics Workshop at PAG 2015.
- Mehta A, Gupta H, Rawal R, Mankad A, Tiwari T, Patel M, Ghosh A (2016) In Silico MicroRNA Identification from *Stevia rebaudiana* Transcriptome Assembly. *European Journal of Medicinal Plants*. 15 (2): 1-14.
- Nakano T, Suzuki K, Fujimura T, Shinshi H (2006) Genome-Wide Analysis of the ERF Gene Family in Arabidopsis and Rice. *Plant Physiology*. 140 (2): 411-432.
- Ono H, Ishii K, Kozaki T, Ogiwara I, Kanekatsu M, Yamada T (2015) Removal of redundant contigs from de novo RNA-Seq assemblies via homology search improves accurate detection of differentially expressed genes. *BMC Genomics*. 16: 1031.
- Padmashree D, Ramachandraswamy N (2016) Identification and characterization of conserved miRNAs with its targets mRNA in *Trichinella Spiralis*. *Bioinformatics*. 12 (5): 279.
- Pant BD, Buhtz A., Kehr J, Scheible, WR (2008) MicroRNA399 is a long-distance signal for the regulation of plant phosphate homeostasis. *The Plant Journal*, 53: 731-738.
- Park W, Li J, Song R, Messing J, Chen X (2002) CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in *Arabidopsis thaliana*. *Current Biology*. 12 (17): 1484-1495.
- Purswlove J, Brown E, Green C, Robbins S (1981) *Spices*. CABI, Wallingford, Oxfordshire.
- Quinlan AR (2014) BEDTools: the Swiss-army tool for genome feature analysis. *Current Protocols in Bioinformatics*. 47 (1): 11.12. 11-11.12. 34.
- Rajeshwari U, Andallu B (2011) Medicinal benefits of coriander (*Coriandrum sativum* L.). *Spatula DD*. 1 (1): 51-58.
- Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of Plant MicroRNA Targets. *Cell*. 110 (4): 513-520.
- Silva F, Ferreira S, Queiroz JA, Domingues FC (2011) Coriander (*Coriandrum sativum* L.) essential oil:

- its antibacterial activity and mode of action evaluated by flow cytometry. *Journal of Medical Microbiology*. 60: 1479-1486.
- Simon J (1990) Essential oils and culinary herbs. Indianapolis, Indiana, USA.
- Singh VK, Singh AK, Singh S, Singh BD (2015) Next-Generation Sequencing (NGS) Tools and Impact in Plant Breeding. In: Al-Khayri JM, Jain SM, Johnson DV (eds) *Advances in Plant Breeding Strategies: Breeding, Biotechnology and Molecular Tools*. Springer International Publishing, Cham, pp 563-612.
- Smith-Unna R, Bournnell C, Patro R, Hibberd JM, Kelly S (2016) TransRate: reference-free quality assessment of de novo transcriptome assemblies. *Genome Research*. 26 (8): 1134-1144.
- Song L, Florea L (2015) Rcorrector: efficient and accurate error correction for Illumina RNA-seq reads. *GigaScience*. 4 (1): 48.
- Tian T, Wang J, Zhou X (2015) A review: microRNA detection methods. *Organic & Biomolecular Chemistry*. 13 (8): 2226-2238.
- Unamba CIN, Nag A, Sharma RK (2015) Next Generation Sequencing Technologies: The Doorway to the Unexplored Genomics of Non-Model Plants. *Frontiers in Plant Science*. 6: 1074.
- Wang M, Wang Q, Wang B (2012) Identification and characterization of microRNAs in Asiatic cotton (*Gossypium arboreum* L.). *PLoS One*. 7 (4): e33696.
- Weiss EA (2002) *Spice crops*. CABI, Wallingford, Oxfordshire.
- Xie F, Frazier TP, Zhang B (2010) Identification and characterization of microRNAs and their targets in the bioenergy plant switchgrass (*Panicum virgatum*). *Planta*. 232 (2): 417-434.
- Xie Z, Allen E, Fahlgren N, Calamar A, Givan SA, Carrington JC (2005) Expression of Arabidopsis miRNA genes. *Plant Physiology*. 138 (4): 2145-2154.
- Xu MY, Zhang L, Li WW, Hu XL, Wang MB, Fan YL, Zhang CY, Wang L (2014) Stress-induced early flowering is mediated by miR169 in Arabidopsis thaliana. *Journal of Experimental Botany*. 65 (1): 89-101.
- Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA (2006a) Conservation and divergence of plant microRNA genes. *The Plant Journal*. 46 (2): 243-259.
- Zhang B, Pan X, Cox S, Cobb GP, Anderson TA (2006b) Evidence that miRNAs are different from other RNAs. *Cellular and Molecular Life Sciences CMLS*. 63 (2): 246-254.
- Zhang X, Zou Z, Gong P, Zhang J, Ziaf K, Li H, Xiao F, Ye Z (2011) Over-expression of microRNA169 confers enhanced drought tolerance to tomato. *Biotechnology Letters*. 33 (2): 403-409.
- Zhao B, Ge L, Liang R, Li W, Ruan K, Lin H, Jin Y (2009) Members of miR-169 family are induced by high salinity and transiently inhibit the NF-YA transcription factor. *BMC Molecular Biology*. 10 (1): 10-29.
- Zhao M, Ding H, Zhu JK, Zhang F, Li WX (2011) Involvement of miR169 in the nitrogen- starvation responses in Arabidopsis. *New Phytologist*. 190 (4): 906-915.
- Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Research*. 31 (13): 3406-3415.