

## شناسایی ژن‌های مسیر بیوسنتزی تری‌ترین و سزکویی‌ترین در میوه هندوانه ابوجهل (*Citrullus colocynthis* L.) با استفاده از فناوری توالی‌یابی نسل جدید (NGS)

معصومه درافشان<sup>۱</sup>، مهدی سلطانی حویزه<sup>۲\*</sup> و وحید شریعتی<sup>۳</sup>

۱- دانشجوی کارشناسی ارشد، گروه ژنتیک و به‌نژادی گیاهی، واحد اهواز، دانشگاه آزاد اسلامی، اهواز، ایران

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

پست الکترونیک: soltani.m@iauhvaz.ac.ir

۳- استادیار، پژوهشگاه ملی مهندسی ژنتیک و زیست‌فناوری، تهران، ایران

تاریخ پذیرش: اردیبهشت ۱۳۹۹

تاریخ اصلاح نهایی: اسفند ۱۳۹۸

تاریخ دریافت: آبان ۱۳۹۸

### چکیده

شناسایی ژن‌های بیوسنتز کننده متابولیت‌های اختصاصی گیاهان دارویی امروزه با سرعت و دقت فراوان با استفاده از فناوری‌های جدید مطالعه ترنسکریپتوم مانند توالی‌یابی RNA انجام می‌شود. این مطالعه به‌منظور شناسایی ژن‌های اختصاصی مسیر بیوسنتز تری‌ترین‌ها و سزکویی‌ترین‌های موجود در بافت میوه هندوانه ابوجهل (*Citrullus colocynthis* L.) انجام شد. با استخراج RNA از بافت میوه‌های هندوانه ابوجهل برداشت شده از منطقه اندیمشک واقع در استان خوزستان در سال ۱۳۹۶، تکنیک توالی‌یابی RNA با استفاده از بن‌سازه Illumina HiSeq 2500 اجرا شد. مراحل بیوانفورماتیکی شامل یک پارچه‌سازی نوپدید با استفاده از نرم‌افزار Evidential-gene و تفسیر کارکردی با استفاده از پایگاه اطلاعاتی KAAS انجام گردید. از تعداد ۲۱۹۵۲۸۸۵ توالی دارای کیفیت بالا، تعداد ۵۵۳۱۱ تک‌ژن یک‌پارچه تولید شد و این تک‌ژن‌ها به‌صورت موازی در پایگاه‌های اطلاعاتی مختلف بارگذاری شدند. در پایگاه KAAS تعداد ۱۷۳۵۹ تک‌ژن در ۱۳۴ مسیر گیاهی آنوتیت (Annotate) شدند. از میان مسیرهای متابولیت‌های ثانویه مختلف و مهم در بافت میوه گیاه هندوانه ابوجهل، به مسیر ژنی "تری‌ترین‌ها" و "سزکویی‌ترین‌ها" تعداد ۳۹ تک‌ژن و ۸ ژن اورتولوگ اختصاص داشت. آنالیز ترنسکریپتوم این گیاه دارویی با هدف شناسایی ژن‌های مسیرهای بیوسنتزی متابولیت‌های ثانویه جنبه‌های پژوهشی و اجرایی مختلفی از جمله مهندسی مسیر بیوسنتز متابولیت‌های دارویی گیاهی را زمینه‌سازی می‌کنند.

واژه‌های کلیدی: کوکوربیتاسین، توالی‌یابی RNA، متابولیت‌های ثانویه، یک‌پارچه‌سازی نوپدید، تری‌ترین‌ها.

### مقدمه

پخش شده است. هندوانه ابوجهل به‌طور سنتی برای درمان زردی، سیاتیک، آسم، سرطان و دیابت استفاده می‌شود (Abo et al., 2008; Upadhyay et al., 2007). تحقیقات دارویی نشان می‌دهد که این گیاه دارویی اثرهای گسترده

گیاه دارویی هندوانه ابوجهل از خانواده کدویان (Cucurbitaceae) به‌عنوان سیب تلخ یا کلوسینت شناخته می‌شود که به‌طور گسترده‌ای در بیابان‌های سراسر جهان

فارنزیل دی فسفات (FPP: Farnesyl Diphosphate) با استفاده از آنزیم بیوسنتزی فارنزیل دی فسفات سینتاز (FPS: Farnesyl Diphosphate Synthase) کاتالیز می شوند. ماده تشکیل شده FPP برای مسیرهای بیوسنتزی ترین وابسته به MVA، از جمله تری ترین ها و سزکویی ترین ها است (Soltani Howyzeh *et al.*, 2018).

خانواده های گیاهی مختلف دارای زیرمجموعه های منحصر بفردی از آنزیم های تری ترین و سزکویی ترین هستند. در سال های اخیر، تعداد آنزیم های بیوسنتزی تری ترین ها و سزکویی ترین های شناسایی شده به طور گسترده ای افزایش یافته است، این امر عمدتاً به دلیل افزایش قابل توجهی در اطلاعات جمع آوری شده از توالی یابی ژنوم و ترنسکرپتوم موجودات است. این مطالعات برنامه های گسترده ای را در زمینه ژنتیک و سازمان دهی خوشه ژنوم، همبستگی یا تکامل مولکولی آنزیم ها مطرح کرده و با هدف توضیح تمام مسیرهای بیوسنتزی یا توصیف تمام خانواده های آنزیمی دنبال می شود (Moses *et al.*, 2015; Boutanaev *et al.*, 2015; Seki *et al.*, 2015; Miettinen *et al.*, 2017b). به همین ترتیب، علاقه به ازدیاد تولید تری ترین ها و سزکویی ترین ها که دارای ارزش تجاری بالقوه هستند در گیاهان و میزبان های مختلف هتروولوگوس (غیر خودی) افزایش یافته است (Jiang *et al.*, 2016). موگروسید وی (mogroside V)، تری ترین ساپونین در عصاره میوه مانک (*Siraitia grosvenori*) به عنوان گیاه دارویی و شیرین کننده طبیعی (Tang *et al.*, 2011)، گینوسید (Gypenosides)، تری ترین ساپونین در گیاه مکینو (*Gynostemma pentaphyllum*) (Chen *et al.*, 2016)، کارانتین (Charantin) به عنوان تری ترین کوکوربیتان در گیاه کارلا یا خیار چنبر تلخ (*Momordica charantia*) با عملکرد مفید دارویی مانند فعالیت ضد سرطان، ضد دیابت و ضد باکتری (Cuong *et al.*, 2017)، گوسیپول و آلدییدهای سزکویی ترین در پنبه به عنوان ترکیب های دفاعی (Tian *et al.*, 2018) و لاکتون های سزکویی ترین در خانواده Asteraceae (Padilla-Gonzalez *et al.*, 2016) را می توان به عنوان نمونه هایی از تری ترین ها و سزکویی ترین های مسیر

بیولوژیکی مانند سمیت سلولی، آنتی اکسیدانی، فعالیت های ضد میکروبی، ضد دیابت و کاهنده چربی دارد (Gurudeeban & Ramanathan, 2010). مصرف هندوانه ابو جهل فواید دارویی فراوانی دارد، اما اخیراً نگرانی هایی در رابطه با مسمومیت ها و عوارض بافتی مانند آسیب روده و کلیه ناشی از مصرف آن مطرح شده است (Javadzadeh *et al.*, 2013). بررسی های پیشین، ترکیب های فیتوشیمیایی این گیاه دارویی را به فلاونوئیدها، آلکالوئیدها، استروئیدها، تریپن ها، ساپونین های تری ترینی، سزکویی ترین ها، کوکوربیتاسین ها و کلوسنتوسیدها گروه بندی می کنند (Uma & Sekar, 2014). در گیاهان مختلف متابولیت های اختصاصی ترین ها به عنوان یک گروه بزرگ از محصولات طبیعی دارای کاربردهای متنوع صنعتی می باشند. تری ترین ها و سزکویی ترین ها طبقه بزرگ و مهم از متابولیت های ثانویه ترین ها با ساختارها و عملکردهای متنوع را تشکیل می دهند. این متابولیت ها قابلیت بالایی برای انواع کاربردهای تجاری مانند تولید واکسن، داروهای ضد سرطان، مکمل های غذایی و لوازم آرایشی دارند (Miettinen *et al.*, 2017a). تری ترین ها جزء اصلی عصاره گیاهی بوده که به عنوان ترکیب های سمی به منظور سازوکارهای بیولوژیکی، ضد عفونی کننده و دفاعی در برابر میکروب ها، قارچ ها و آفات فعالیت کرده، در حالی که سزکویی ترین ها ماده اصلی اسانس روغنی شناخته شده که در رایحه، طعم، رنگدانه های فتوسنتزی و فعالیت های آنتی اکسیدانی نقش دارند (Langenheim, 1994; Moses *et al.*, 2014).

گیاهان تری ترین ها و سزکویی ترین ها را در مسیرهای بیوسنتزی بلند و شاخه دار تولید می کنند. در واقع، این ترکیب های اولیه و اختصاصی همه در اطراف اسکلت ترین براساس شش ایزوپنتنیل پیروفسفات (IPP: Isopentenyl Pyrophosphate) در مسیر اصلی سیتوزولی موالونیک اسید (Mevalonic Acid: MVA) ساخته می شوند (Thimmappa *et al.*, 2014). بخشی از مجموعه IPP تولید شده توسط ایزومراز آن به دی متیل آلیل پیروفسفات (DMAPP: Dimethylallyl Pyrophosphate) تبدیل می شود. دو مولکول IPP و DMAPP برای تبدیل به

توالی‌یابی نسل جدید، ترنسکرپتوم بافت میوه گیاه دارویی هندوانه ابوجهل برای شناسایی ژن‌های تولیدکننده آنزیم‌های مسیر بیوسنتزی تری‌ترین‌ها و سزکویی‌ترین‌ها برای اولین بار مورد تجزیه و تحلیل قرار گرفته است.

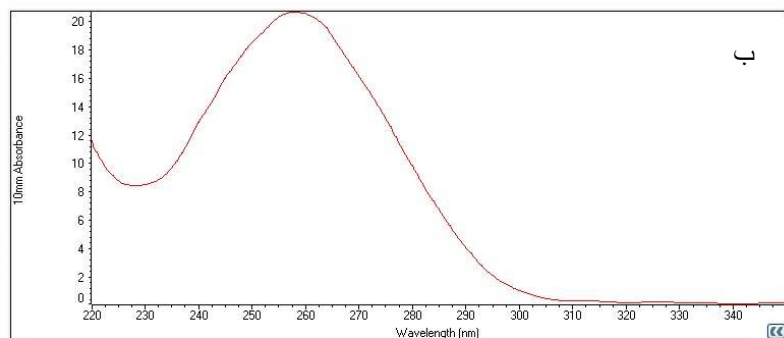
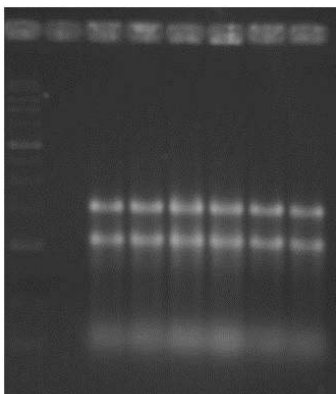
### مواد و روش‌ها

گیاه دارویی هندوانه ابوجهل از منطقه اندیمشک واقع در استان خوزستان در محدوده جغرافیایی  $48/139^{\circ}E$  و  $32/403^{\circ}N$  در خردادماه سال ۱۳۹۶ به صورت تصادفی ۶ بوته انتخاب و میوه رسیده آنها در مرحله فنولوژیکی رسیدگی میوه نمونه‌گیری شد. نمونه‌ها بلافاصله با استفاده از ازت مایع به آزمایشگاه تحقیقاتی دانشکده کشاورزی و منابع طبیعی دانشگاه آزاد اسلامی واحد اهواز منتقل شده و نگهداری نمونه‌ها در فریزر با دمای  $-80^{\circ}C$  درجه سانتی‌گراد تا زمان آماده‌سازی استخراج RNA انجام شد. با استفاده از کیت استاندارد کیاژن مطابق با دستورالعمل شرکت سازنده RNA کل از نمونه‌های بافت گیاهی (میوه کامل) استخراج شده، کمیّت و کیفیت آنها به ترتیب با استفاده از ژل آگارز ۱٪ و اسپکتروفوتومتر نانودراپ تعیین شد (شکل ۱) و ۶ نمونه RNA استخراج شده با هم ادغام و یک نمونه پولد شده برای توالی‌یابی به مؤسسه ژنومی (BGI) واقع در کشور چین ارسال گردید.

بیوسنتزی متابولیت‌های ثانویه اشاره نمود. امروزه توالی‌یابی RNA یک ابزار موفق و پرطرفدار برای کشف فرایند ترنسکرپتوم شناخته می‌شود، زیرا توالی‌یابی RNA دارای مزایای متعددی نسبت به توالی‌یابی DNA مانند هزینه کمتر، بازده بالاتر و ویژگی‌های پیشرفته‌تر می‌باشد (Zheng et al., 2014; Rabbani et al., 2016). این تکنولوژی توالی‌یابی نسل جدید (NGS) برای شناسایی ژن‌ها و آنزیم‌های گوناگون براساس بیوسنتز متابولیت‌های ثانویه مختلف از جمله تری‌ترین‌ها، فلاونوئیدها و آلکالوئیدها در گیاهان مختلف استفاده می‌شود (Soltani Howyze et al., 2018; Dorafshan et al., 2018; Amiripour et al., 2018).

از آن‌جا که در گیاه دارویی هندوانه ابوجهل بسیاری از متابولیت‌های ثانویه شناسایی شده، با این حال اطلاعات ژنتیکی این گیاه دارویی محدود بوده و مسیرهای بیوسنتزی متابولیت‌های ثانویه مشخص نشده است. همچنین در مطالعه پیشین توالی‌یابی ترنسکرپتوم هندوانه ابوجهل توسط Wang و همکاران (۲۰۱۴) دو محدودیت مهم وجود داشت. محدودیت اول نمونه مورد استفاده برای توالی‌یابی شامل یک بافت برگ بوده و محدودیت دوم توالی‌یابی آن در شرایط تنش خشکی بوده، نه اینکه در محیط طبیعی و به صورت خودرو رشد کرده باشد (Wang et al., 2014). در این مطالعه با استفاده از ابزار

الف



ب

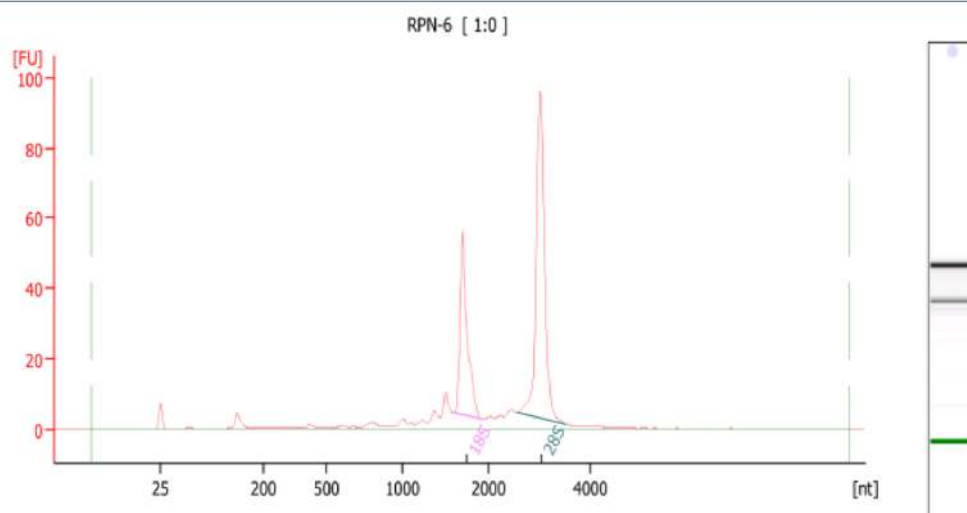
شکل ۱- الف) عکس الکتروفورز RNA استخراج شده روی ژل آگارز ۱٪، ب) اسپکتروم جذب نوری دستگاه نانودراپ

برای تعیین کیفیت RNA در طول موج‌های مختلف

۱ نمایانگر تخریب شده‌ترین و ۱۰ نمایانگر سالم‌ترین RNA است. عدد RIN برای نمونه ارسال شده در این تحقیق برابر ۹/۱ و در سطح A بود (شکل ۲). به این معنی که نمونه دارای شرایط لازم و کافی برای ساخت کتابخانه است. سپس با استفاده از بن‌سازه Illumina HiSeq2500 به‌صورت خوانش‌های دو طرفه ۲×۱۵۰ توالی‌یابی انجام شد.

پیش از توالی‌یابی با استفاده از تکنولوژی آجیلنت بیوآنالیزر ۲۱۰۰ (Bioanalyzer®2100 Agilent) خلوص RNA ارزیابی شد. دستگاه آجیلنت بیوآنالیزر ۲۱۰۰ دارای یک نرم‌افزار الگوریتم برای محاسبه عدد یکپارچگی RNA یا RNA Integrity Number (RIN) با یک نمایش دیجیتال توزیع اندازه مولکول‌های RNA می‌باشد. عدد RIN براساس یک سیستم شماره‌گذاری از ۱ تا ۱۰ می‌باشد که

Sample Name	Concentration (ng/μL)	OD260/280	OD260/230	RIN	28S/18S	Test Result
RPN-6	183	2.11	2.16	9.1	1.8	Level A



#### Overall Results for sample 9 : RPN-6

RNA Area:	305.1	RNA Integrity Number (RIN):	9.1 (B.02.07)
RNA Concentration:	183 ng/μl	Result Flagging Color:	<span style="background-color: #ccccff; border: 1px solid black; display: inline-block; width: 15px; height: 10px;"></span>
rRNA Ratio [28s / 18s]:	1.8	Result Flagging Label:	RIN: 9.10

#### Fragment table for sample 9 : RPN-6

Name	Start Size [nt]	End Size [nt]	Area	% of total Area
18S	1,589	1,940	66.9	21.9
28S	2,558	3,512	119.0	39.0

شکل ۲- نتایج تعیین کیفیت نمونه ارسال شده برای توالی‌یابی توسط دستگاه آجیلنت بیوآنالیزر ۲۱۰۰ و تعیین سطح کیفیت نمونه

ایلومینا توسط نرم‌افزار Trimmomatic (نسخه ۰/۳۲) حذف گردید (شکل ۳). سپس با استفاده از نرم‌افزار FastQC خصوصیات کتابخانه ارزیابی و کارایی پیرایش داده‌ها تأیید شد.

کیفیت اولیه خوانش‌ها با استفاده از نرم‌افزار FASTQC (نسخه ۰،۱۰،۱) مشخص و توالی‌های دارای طول کوچکتر از ۵۰ نوکلئوتید، نوکلئوتیدهای دارای امتیاز کیفیت Phred (Phred quality score) کمتر از ۲۰ و همچنین آدپتورهای



شکل ۳- کیفیت توالی‌بازها در خوانش‌های رفت و برگشت توالی‌یابی شده قبل و بعد از تریم کردن

استفاده از سیستم شناسایی qRT-PCR (دستگاه Corbett Rotor-Gene 6000 instrument, Corbett Life TaKaRa SYBR® Green و کیت Permixon Ex Taq™ II (Science, Australia) انجام شد. ژن کنترل داخلی اکتین گیاه هندوانه ابوجهل برای تخمین مقدار نسبی رونوشت تک‌ژن‌ها مورد استفاده قرار گرفت (Si et al., 2010). برای تحلیل داده‌های qRT-PCR از روش  $\Delta\Delta CT$  استفاده شد. دو تکرار تکنیکی برای همه آزمایش‌های qRT-PCR استفاده شد. اولیگونوکلوئوتیدهای اختصاصی ژن‌های انتخاب شده برای تحلیل qRT-PCR در جدول ۱ نشان داده شده است. روش ساخت cDNA براساس دستورالعمل کیت خریداری شده شرکت ترموساینتیفیک انجام شد. cDNA حاصل در دمای ۲۰- درجه سانتی‌گراد نگهداری شد. واکنش زنجیره‌ای پلیمرز کمی با استفاده از فناوری رنگ سایبرگرین انجام شد. شرایط بهینه برای اجرای واکنش PCR زمان واقعی در حجم ۲۰ میکرولیتر و از هر نمونه دو تکرار فراهم گردید. برای جلوگیری از بروز خطاهای احتمالی در هنگام برداشت مقادیر کم مواد و نیز

با توجه به اینکه توالی ژنوم گیاه دارویی هندوانه ابوجهل وجود نداشته، روش یک‌پارچه‌سازی نوپدید ترنسکرپتوم (De novo Assembly Transcriptome) مورد استفاده قرار گرفت. به این منظور خوانش‌های دارای کیفیت بالا با استفاده از نرم‌افزار Evidential-gene یک‌پارچه‌سازی شدند. پس از بهینه‌سازی خوانش‌ها، K-mer استاندارد اجرا شده و برای سایر پارامترهای نرم‌افزار از تنظیمات پیش‌فرض استفاده گردید. توالی تک‌ژن‌های یک‌پارچه شده در پایگاه KAAS (<http://www.Genome.Jp/kegg/kaas>) بارگذاری شد و با استفاده از شناسه‌های اختصاصی KO (Orthology KEGG) تفسیر شدند. شناسه‌های اختصاصی KO با استفاده از روش Single-directional Best Hit (SBH) اجرا گردید. هر شناسه KO مجموعه‌ای از ارتولوگ‌های ژنی است که به صورت مستقیم به یک مسیر KEGG پیوست شده و برای ارتباط دادن ژنوم با مسیرهای زیستی بکار می‌رود.

برای تأیید نتایج بررسی بیان کمی برخی تک‌ژن‌های انتخاب شده در مسیر بیوسنتز تری‌ترین و سزکویی‌ترین با

اصلی یک تیوب حاوی تمام مواد لازم برای انجام qRT-PCR بجز cDNA نیز در دستگاه ترموسایکلر قرار داده شد. در صورت آلودگی در تیوب فاقد cDNA، سنتز DNA انجام می‌شود. چرخه حرارتی qRT-PCR در جدول ۲ نشان داده شده است.

برای راحتی و سرعت عمل بیشتر، اقدام به تهیه محلول‌های پایه گردید. تمام مراحل تهیه محلول qRT-PCR و نیز افزودن cDNA بر روی یخ انجام شد. برای تشخیص آلودگی در واکنش PCR زمان واقعی از کنترل منفی استفاده گردید. برای این کار به همراه نمونه‌های

جدول ۱- لیست آغازگرهای اولیگونکلئوتیدی مورد استفاده برای ژن‌های انتخابی در qRT-PCR

ردیف	ژن	رفت (۳-۵)	برگشت (۳-۵)
۱	Curbitadienol synthase (CPQ)	GTATTTCCATCGGGTTTTGACTG	CCATATCGTCAAAGAAAGACAACCC
۲	Squalene monooxygenase (SQLE)	GAAGAAGAGGCATTACAAGGAAG	TCTACTCTATTAAGCGTCCCAAG
۳	Actin	CAACATACATAGCAGGCACA	TGACTGAGGCTCCACTCAAC

جدول ۲- چرخه حرارتی واکنش زنجیره‌ای پلیمرز

تعداد چرخه	مرحله	زمان	دما (°C)
یک چرخه	واسرشت‌سازی	۱۰ دقیقه	۹۵
	واسرشت‌سازی	۱۵ ثانیه	۹۵
۴۰ چرخه	اتصال	۱ دقیقه	۶۰
	تکنیر	۲ دقیقه	TO-GO
یک چرخه	تکنیر نهایی	۵ ثانیه	۶۰

## نتایج

اسیدچرب " (۱۳۸ تک‌ژن، ko01212) و "بیوسنتز اسیدآمینه" (۵۰۳ تک‌ژن، ko01230) طبقه‌بندی شدند (جدول ۳). هر یک از پنج مسیر اصلی دارای مسیر فرعی و زیر مجموعه‌هایی هستند. مسیر اصلی متابولیت ثانویه در کل گیاه دارای تعدادی از مسیرهای فرعی مهم از جمله متابولیسم ترپنوئیدها و پلی‌کتیدها (۳۴۹ تک‌ژن) که شامل زیر مجموعه‌های ترپنوئیدی "بیوسنتز اسکلت ترپنوئید" (۱۰۱ تک‌ژن، ko00900)، "بیوسنتز کاروتنوئید" (۶۲ تک‌ژن، ko00906)، "بیوسنتز تری‌ترپنوئید و سزکویی‌ترپنوئید" (۵۵ تک‌ژن، ko00909)، "بیوسنتز دی‌ترپنوئید" (۴۱ تک‌ژن، ko00904)، "بیوسنتز مونوترپنوئید" (۲۳ تک‌ژن، ko00902) می‌باشد.

در این مطالعه از پایگاه اتوماتیک KAAS به منظور نقشه‌یابی مسیرهای KEGG برای ژن‌های ارتولوگ اختصاصی گیاه مربوطه استفاده گردید. از تعداد کل ۵۵۳۱۱ تک‌رونوشت یک‌پارچه شده تعداد ۱۷۳۵۹ تک‌ژن از میان آنها در پایگاه KAAS شناسایی شدند که این تعداد تک‌ژن در ۱۳۴ مسیر زیستی قرار گرفتند. تک‌ژن‌های تفسیر شده در گیاه هندوانه ابوجهل به پنج مسیر اصلی متابولیسمی در پایگاه داده KEGG شامل "مسیرهای متابولیکی" (۳۹۳۵ تک‌ژن، ko01100)، "بیوسنتز متابولیت ثانویه" (۲۰۹۸ تک‌ژن، ko01110)، "متابولیسم کربن" (۵۴۱ تک‌ژن، ko01200)، "متابولیسم اسید اوکسوکربوکسیلیک-۲" (۱۲۶ تک‌ژن، ko01210)، "متابولیسم

جدول ۳- بررسی کلی نقشه مسیرهای اصلی بیوسنتز متابولیت‌ها و مسیر بیوسنتز تری‌ترین و سزکویی‌ترین شناسایی شده در پایگاه اطلاعاتی KEGG برای رونوشت‌های شناسایی شده هندوانه اوجهل

ردیف مسیر	مسیرهای KEGG	شماره Ko مسیرهای KEGG	تعداد کل اورتولوگ‌های* در مسیر KEGG	تعداد اورتولوگ‌های شناسایی شده در مسیر KEGG	درصد اورتولوگ شناسایی شده در مسیر KEGG	تعداد تک‌ژن‌های شناسایی شده گیاه در مسیر KEGG	درصد تک‌ژن‌های شناسایی شده گیاه در نقشه
مسیرهای اصلی بیوسنتز	مسیرهای متابولیسم	۱۱۰۰	۲۹۲۴	۸۸۶	۳۰/۳	۳۹۳۵	۴/۴
	بیوسنتز متابولیت‌های ثانویه	۱۱۱۰	۱۱۴۶	۳۹۶	۳۴/۵	۲۰۹۸	۵/۲
	متابولیسم کرین	۱۲۰۰	۳۴۹	۹۵	۲۷/۲	۵۴۱	۵/۷
	متابولیسم اسید اوسوکربوکسیلیک-۲	۱۲۱۰	۸۲	۲۸	۳۴/۱	۱۲۶	۴/۵
	متابولیسم اسیدجرب	۱۲۱۲	۷۰	۲۵	۳۵/۷	۱۳۸	۵/۵
	بیوسنتز اسید آمینه‌ها	۱۲۳۰	۲۳۰	۹۸	۴۲/۶	۵۰۳	۵/۱
مسیر فرعی		۰۰۹۰۹	۷۳	۸	۱۰/۹	۳۹	۴/۸

\*: اورتولوگ ژن‌های همولوگی هستند که به دلیل گونه‌زایی از هم جدا شده‌اند و با یک ژن اجدادی در ژنوم جدشان مرتبط هستند. انتظار می‌رود اورتولوگ‌ها دارای ساختار و عملکرد مشابه باشد.

میوه گیاه هندوانه ابوجهل تنها ۸ ژن اورتولوگ (۱۰/۹٪) وجود داشت. برای این ۸ ژن اورتولوگ ۳۹ تک ژن شناسایی گردید، یعنی به طور متوسط به ازای هر ژن اورتولوگ مسیر ۴/۸ تک ژن وجود دارد (جدول ۳).

تری ترین ها و سزکویی ترین ها به عنوان اجزاء اصلی تشکیل دهنده اسانس و عصاره گیاه دارویی هندوانه ابوجهل از پیش ماده IPP حاصل از مسیر MVA سنتز می شوند. در مسیر بیوسنتز تری ترین ها و سزکویی ترین ها با شناسه ko00909 تعداد ۷۳ ژن اورتولوگ وجود دارد که در بافت

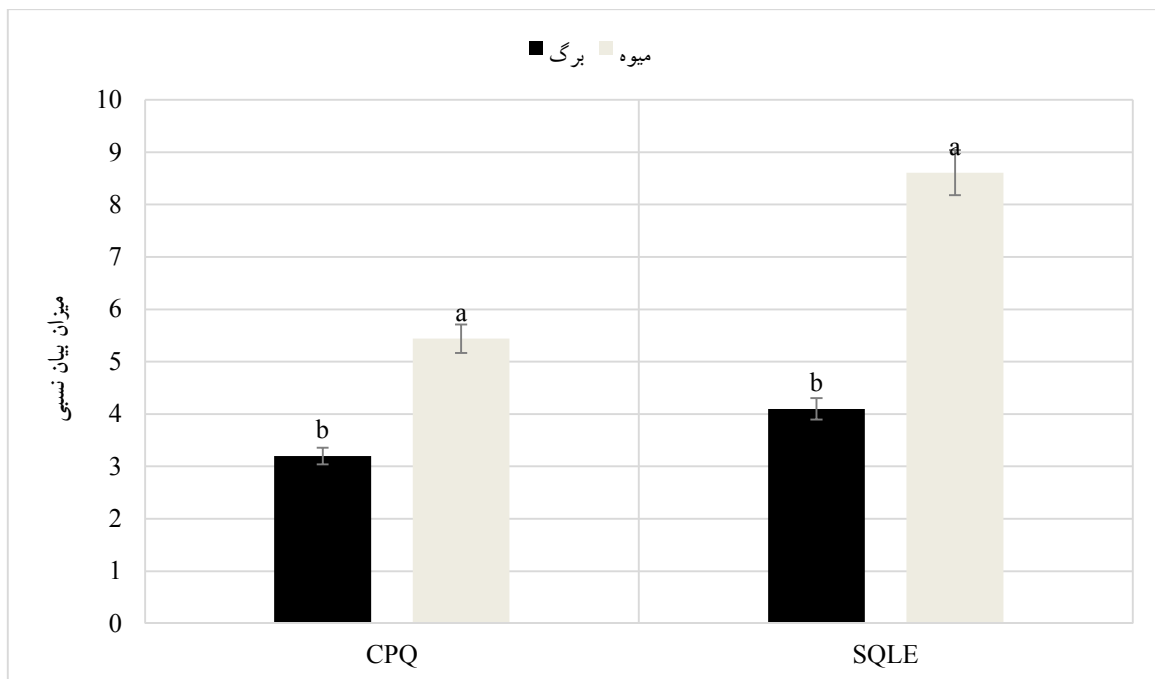
جدول ۴- آنزیم های اختصاصی مسیر بیوسنتز تری ترین و سزکویی ترین ترنسکریپتوم میوه هندوانه ابوجهل براساس نتایج پایگاه KEEG

مسیر	شناسه K	شناسه تک ژن هندوانه ابوجهل	توصیف آنزیم	عدد* EC
مسیر بیوسنتز تری ترین	K00801	505110	FDFT1; farnesyl-diphosphate farnesyltransferase	[EC:2.5.1.21]
	K00511	7168, 113835, 113836, 145266, 489173, 492517, 506297, 622126, 1007950, 1019764, 1138399, 1192324, 1228914, 1228916, 1303925	SQLE, ERG1; squalene monoxygenase	[EC:1.14.14.17]
	K15812	78428, 1309093	CPQ; cucurbitadienol synthase	[EC:5.4.99.33]
	K15813	8761, 88998, 89000, 89001, 89002, 89003, 212620, 1192633, 1192639, 1286333, 1286336, 1286339, 1286344, 1286347	LUP4; beta-amyrin synthase	[EC:5.4.99.39]
	K15820	187921	IMS1; isomultiflorenol synthase	[EC:5.4.99.36]
مسیر بیوسنتز سزکویی ترین	K15891	503459, 1028192	FLDH; NAD <sup>+</sup> -dependent farnesol dehydrogenase	[EC:1.1.1.216]
	K14175	69217, 499377, 794740, 1123239, 138917	NES1; (3S,6E)-nerolidol synthase	[EC:4.2.3.48]
	K15803	24743, 25842, 20901, 20902, 68357, 120935, 122099, 125832, 226880, 580837, 1166565, 1199757, 1332460, 1332463, 1332468	GERD; (-)-germacrene D synthase	[EC:4.2.3.75]

این تحقیق بیان ژن های کوربیتادینول سینتاز (CPQ) و اسکوالن سینتاز (SQLE) در بافت میوه و برگ مورد ارزیابی قرار گرفت. نتایج نشان داد که میزان بیان نسبی هر دو ژن در بافت میوه بیش از بافت برگ گیاه دارویی هندوانه ابوجهل بود و اختلاف بیان ژن ها در سطح آماری ۵٪ معنی دار بود. به طوری که میزان بیان نسبی ژن های کوربیتادینول سینتاز (CPQ) و اسکوالن سینتاز (SQLE) در بافت میوه به ترتیب ۱/۷ و ۲/۱ برابر بیشتر از میزان بیان نسبی این ژن ها در برگ بود (شکل ۴).

در مسیر بیوسنتز تری ترین آنزیم های اختصاصی فارنزیل-دی فسفات فارنزیل ترانسفراز (FDFT1)، اسکوالن مونواکسیژناز (SQLE, ERG1)، کوکوروبیتادینول سینتاز (CPQ)، بتا-آمیرین سینتاز (LUP4) و ایزومولتی فلورنول سینتاز (IMS1) شناسایی شدند (جدول ۴). همچنین در مسیر بیوسنتز سزکویی ترین آنزیم های اختصاصی NAD<sup>+</sup>-دی پندنت فارنزول دی هیدروژناز (FLDH)، (3S,6E)-نرولیدول سینتاز (NES1) و (-) -جرماکرن دی سینتاز (GERD) مشخص شدند (جدول ۴). برای تأیید داده های حاصل از توالی یابی بدست آمده در





شکل ۴- نمودار بررسی بیان ژن‌های کوربیتادینول سینتاز (CPQ) و اسکوالن سینتاز (SQLE) در بافت میوه و برگ گیاه دارویی هندوانه ابوجهل

## بحث

و همکاران (۲۰۱۴) در بافت برگ گیاه هندوانه ابوجهل تحت شرایط تنش خشکی برای شناسایی ژن‌های مقاوم به خشکی می‌باشد (Wang et al., 2014). Wang و همکاران (۲۰۱۴) تعداد ۲۰۵۸۱ کانتینگ را بدست آوردند که از میان آنها تعداد ۲۵۴۵ دارای تفاوت بیان معنی‌دار در طی تنش خشکی بودند. به دلیل اینکه هدف مورد نظر مطالعه قبلی متفاوت با هدف این مطالعه است، از این رو هیچ داده‌ای از شناسایی ژن‌های مسیر بیوسنتز متابولیت‌های ثانویه به‌ویژه ژن‌های تولیدکننده آنزیم‌های مسیر بیوسنتزی تری‌ترین‌ها و سزکویی‌ترین‌ها تاکنون گزارش نشده است. به طوری که در مورد شناسایی ژن‌های مسیر متابولیت‌های ثانویه سزکویی‌ترین و تری‌ترین در سایر مطالعات می‌توان به موارد ذیل اشاره نمود.

بابونه آلمانی (*Matricaria chamomilla*) یکی از محبوب‌ترین گیاهان علفی از نظر فعالیت دارویی متابولیت‌های سزکویی‌ترین، مانند (-)آلفا-بیسابولول

گیاهان انواع مختلفی از تری‌ترین‌های طبیعی، از جمله رنگدانه‌های فتوسنتزی، مولکول‌های سیگنالیک و ترکیب‌های دفاعی تولید می‌کنند. اینها متابولیت‌های پیچیده‌ای هستند که به صورت انتخابی در مراحل مختلف رشد و نمو شکل می‌گیرند، به نحوی که بعضی از آنها خواص دارویی و کاربردهای صنعتی پیدا کرده‌اند. روش‌های جدید توالی‌یابی ترنسکریپتومی و ژنومی می‌تواند منجر به کشف ژن‌های مسیرهای مرتبط با بیوسنتز متابولیت‌های ثانویه شود، همچنین به درک نقش فیزیولوژیکی و بیولوژیکی این محصولات طبیعی در گیاهان مختلف کمک کند. این مطالعه اولین مطالعه بافت میوه گیاه دارویی هندوانه ابوجهل با استفاده از ابزار توالی‌یابی نسل جدید برای شناسایی ژن‌های تولیدکننده آنزیم‌های مسیر بیوسنتزی تری‌ترین‌ها و سزکویی‌ترین‌ها می‌باشد. بررسی منابع نشان داد که تنها مطالعه تحلیل ترانسکریپتوم قبلی انجام شده در این گیاه توسط Wang

Roch 454-GS-FLX، آنزیم‌های اسکلت‌ترین و مسیر بیوسنتزی تری‌ترین از جمله اعضای خاصی از خانواده سیتوکروم P450 (CYP450s) را مشخص کرد (Bhambhani et al., 2017). عشقه یا پیچک (*Hedera helix*) مهمترین گیاه دارویی سنتی در اروپاست. از آنجا که اجزاء اصلی فعال، تری‌ترین‌ها هستند، تجزیه و تحلیل ترنسکریپتوم با استفاده از بن‌سازه Illumina HiSeq 2500 برای برگ و ریشه آن انجام شد. پس از تفسیر کارکردی و طبقه‌بندی دو مسیر MVA و MEP تعداد ۴۱۰ تک ژن مربوط به بیوسنتز تری‌ترین سایونین از جمله هدراکوسید سی (Hederacosid C) و آلفا-هدرین ( $\alpha$ -hederin) گزارش گردید (Sun et al., 2017). براساس تجزیه و تحلیل بیوانفورماتیکی خوانش‌های تکنولوژی Illumina HiSeq 2000، از بافت‌های برگ و ریزوم گیاه شقایق نعمانی (*Anemone flaccida*) خانواده Ranunculaceae تمام آنزیم‌های احتمالی درگیر در بیوسنتز تری‌ترین‌ها در دو مسیر MVA و MEP شناسایی شدند. علاوه بر این، در مجموع ۱۲۶ مورد CYP450s و ۳۲ مورد UDP-glucosyltransferase (UDPGs) به‌عنوان نامزدهای سنتزکننده تری‌ترین سایونین انتخاب شدند (Zhan et al., 2016). جینسینگ (*Panax notoginseng*) گیاه دارویی مهم از خانواده Araliaceae دارای اجزاء فعال تری‌ترین سایونین است. با توجه به محدودیت اطلاعات ژنتیکی این گیاه با استفاده از تکنولوژی Roch 454-GS-FLX مجموعه گسترده‌ای از EST برای ریشه آن بدست آمد. نتایج این تحقیق نشان داد که رونوشت کد شده دامارن‌دیول سینتاز (DS: dammarenediol synthase) که اولین آنزیم کلیدی در مسیر بیوسنتز تری‌ترین‌های سایونین است، در ریشه این گیاه بسیار بیان می‌شود. همچنین ژن‌های CYP450s (pn00158, pn02132) و UDPGs (pn00082) شناسایی شدند که به احتمال زیاد در هیدروکسیله کردن یا گلیکوزیله کردن آگیلکون‌ها در مسیر بیوسنتز تری‌ترین سایونین درگیر هستند (Luo et al., 2011).

( $\alpha$ -bisabolol) و بیسابولول اکسید A (bisabolol oxide A) می‌باشد. ژن‌های مختلف بیان شده در گل، ساقه، ریشه و برگ مورد بررسی قرار گرفتند و ۱۱۱۵ ژن با استفاده از دانشنامه ژنوم و ژن‌های کیوتو (KEGG) به مسیر بیوسنتز متابولیت‌های متفاوتی اختصاص یافتند. ژن‌های مسیر سنتز سزکویی‌ترین‌ها در بافت‌های مختلف گیاه بابونه اختلاف معنی‌داری را نشان دادند (Zhang et al., 2018). چوب گیاه صندل بالغ گرمسیری (*Santalum album*) تولیدکننده یکی از پرترفدارترین رایحه‌های جهان به‌شمار می‌رود. براساس الگوی بیان ژن‌ها در این گیاه، آنزیم‌های سیتوکروم P450 (SaCYP736A167) به‌عنوان آنزیم‌های احتمالی درگیر در بیوسنتز سزکویی‌ترین انتخاب شدند. این آنزیم‌ها عمدتاً سنتالول‌ها (*Santalols*) یا برگاموتول (*bergamotol*) تولید می‌کنند، از این رو قابلیت تولید صنعتی این ترکیب‌ها برای جایگزینی روش استخراج از عطر روغن چوب صندل وجود دارد (Celedon et al., 2016). اسانس‌های معطر استخراج شده از میوه‌های تازه گیاه لیتسه آ (*Litsea cubeba*) متعلق به خانواده Lauraceae دارای ارزش‌های پزشکی و اقتصادی متنوع هستند. اجزاء غالب در این روغن‌های ضروری مونوترپن‌ها و سزکویی‌ترین‌ها می‌باشند. با جستجو در برابر KEGG، ۱۶۱۳۰ تک ژن در ۲۹۷ مسیر KEGG اختصاص داده شد و ۶۱ تک ژن که شامل مسیرهای مولونات (mevalonate) و ۲-سی-متیل-دی-اریتریټول-۴-فسفات (2-C-methyl-D-erythritol-4-phosphate) مربوط به بیوسنتز اسکلت‌ترین می‌شود. از ۱۲۹۶۳ تک ژن، ۲۸۵ مورد با استفاده از پایگاه داده PlantCyc به مسیرهای تری‌ترین اشاره شد. علاوه بر این، ۱۴ ژن تری‌ترین سینستاز در ترنسکریپتوم شناسایی گردید (Han et al., 2013).

درخت چریش (*Azadirachta indica* A.) که معمولاً به نام Neem شناخته می‌شود، منبع تری‌ترین‌ترین بوده که اهمیت اقتصادی دارد. تجزیه و تحلیل ترنسکریپتومی از مراحل مختلف رشد برگ و میوه با استفاده از تکنولوژی

در اسانس و عصاره گیاهان دارویی و تنظیم آنها برای تولید محصولات و داروهای گیاهی شود.

### منابع مورد استفاده

- Abo, K., Fred-Jaiyesimi, A. and Jaiyesimi, A., 2008. Ethnobotanical studies of medicinal plants used in the management of diabetes mellitus in South Western Nigeria. *Journal of Ethnopharmacology*, 115(1): 67-71.
- Amiripour, M., Sadat Nouri, S.A., Shariati, V. and Soltani Howyzeh, M., 2018. Identification of terpenoid backbone biosynthetic pathway genes in Ajowan (*Trachyspermum ammi* L.) by RNA-Seq. *Modern Genetic Journal*, 13(1): 133-141.
- Bains, S., Thakur, V., Kaur, J., Singh, K. and Kaur, R., 2018. Elucidating genes involved in sesquiterpenoid and flavonoid biosynthetic pathways in *Saussurea lappa* by de novo leaf transcriptome analysis. *Genomics*, 116(6): 1474-1482.
- Bhambhani, S., Lakhwani, D., Gupta, P., Pandey, A., Dhar, Y.V., Bag, S.K., Asif, M.H. and Trivedi, P.K., 2017. Transcriptome and metabolite analyses in *Azadirachta indica*: identification of genes involved in biosynthesis of bioactive triterpenoids. *Scientific Reports*, 7(5043): 1-12.
- Boutanaev, A.M., Moses, T., Zi, J., Nelson, D.R., Mugford, S.T., Peters, R.J. and Osbourn, A., 2015. Investigation of terpene diversification across multiple sequenced plant genomes. *Proceedings of the National Academy of Sciences*, 112(1): E81-E88.
- Celedon, J.M., Chiang, A., Yuen, M.M., Diaz-Chavez, M.L., Madilao, L.L., Finnegan, P.M., Barbour, E.L. and Bohlmann, J., 2016. Heartwood-specific transcriptome and metabolite signatures of tropical sandalwood (*Santalum album*) reveal the final step of (Z)-santalol fragrance biosynthesis. *The Plant Journal*, 86(4): 289-299.
- Chen, Q., Ma, C., Qian, J., Lan, X., Chao, N., Sun, J. and Wu, Y., 2016. Transcriptome sequencing of *Gynostemma pentaphyllum* to identify genes and enzymes involved in Triterpenoid biosynthesis. *International Journal of Genomics*, 2016: 7840914.
- Cuong, D.M., Jeon, J., Morgan, A.M., Kim, C., Kim, J.K., Lee, S.Y. and Park, S.U., 2017. Accumulation of charantin and expression of triterpenoid biosynthesis genes in bitter melon (*Momordica charantia*). *Journal of agricultural and food chemistry*, 65(33): 7240-7249.
- Dorafshan, M., Soltani Howyzeh, M. and Shariati, V., 2019. Identification of terpenoid backbone biosynthetic pathway genes in fruit of *Citrullus colocynthis* medical

تکنولوژی توالی‌یابی نسل جدید نیز در مطالعات پروفایل ترنسکریپتومی سایر گیاهان برای بیوسنتز سزکویی‌ترین‌ها از جمله گیاه بابونه آلمانی (*Matricaria chamomilla*) (Tao et al., 2016)، گیاه فلفل بحری (*Capsicum spp.*) (Lee et al., 2017)، گیاه قسط (*Astraceae Saussurea lappa*) از خانواده (Bains et al., 2018) و برای بیوسنتز تری‌ترین‌ها در گیاه جینسینگ آمریکای شمالی (*Panax quinquefolius*) (Wu et al., 2013) و جینسینگ آمریکایی (*Panax japonicas*) (Rai et al., 2016) مورد استفاده قرار گرفته است.

اسانس‌های معطر و تلخ مزه و عصاره‌های سمی استخراج شده از میوه‌های تازه گیاه دارویی هندوانه ابوجهل اهمیت پزشکی و اقتصادی متنوع دارند. اجزاء غالب در این روغن‌های اسانسی تری‌ترین‌ها و سزکویی‌ترین‌ها هستند. البته شناخت سازوکارهای مولکولی از بیوسنتز تری‌ترین‌ها برای بهبود عملکرد و کیفیت آنها ضروریست. به این ترتیب، توالی‌یابی ترنسکریپتوم با توان بالای گیاه دارویی هندوانه ابوجهل برای تولید مقادیر زیادی از توالی‌های رونویسی با هدف کشف ژن، به‌ویژه ژن‌های مرتبط با بیوسنتز تری‌ترین‌ها و سزکویی‌ترین‌ها ضرورت دارد. در این پژوهش استفاده از تکنولوژی RNA-Seq براساس فناوری‌های NGS منجر به شناسایی ژن‌های مرتبط با مسیر بیوسنتزی تری‌ترین‌ها و سزکویی‌ترین‌ها در گیاه دارویی هندوانه ابوجهل شد. این مطالعه اولین تجزیه و تحلیل نوپدید ترنسکریپتوم را در بافت میوه هندوانه ابوجهل نشان می‌دهد. همچنین اطلاعاتی در مورد شناسایی و بیان ژن‌های مربوط به سایر مسیرهای متابولیت‌های ثانویه فعال توزیع شده در بافت‌های متفاوت این گیاه دارویی را به ما ارائه می‌دهد. بنابراین تحقیقات مسیرهای بیوسنتزی متابولیت‌های ثانویه و عملکردهای مرتبط می‌تواند مبنای طرح‌های پژوهشی و اجرایی در زمینه به‌نژادی مولکولی، مهندسی متابولیک و مهندسی ژنتیک با هدف افزایش متابولیت‌های اختصاصی

- Padilla-Gonzalez, G.F., dos Santos, F.A. and Da Costa, F.B., 2016. Sesquiterpene lactones: more than protective plant compounds with high toxicity. *Critical Reviews in Plant Sciences*, 35(1): 18-37.
- Rabbani, B., Nakaoka, H., Akhondzadeh, S., Tekin, M. and Mahdieh, N., 2016. Next generation sequencing: implications in personalized medicine and pharmacogenomics. *Molecular BioSystems*, 12(6): 1818-1830.
- Rai, A., Yamazaki, M., Takahashi, H., Nakamura, M., Kojoma, M., Suzuki, H. and Saito, K., 2016. RNA-seq transcriptome analysis of *Panax japonicus*, and its comparison with other *Panax species* to identify potential genes involved in the saponins biosynthesis. *Frontiers in Plant Science*, 7: 481.
- Seki, H., Tamura, K. and Muranaka, T., 2015. P450s and UGTs: key players in the structural diversity of triterpenoid saponins. *Plant and Cell Physiology*, 56(8): 1463-1471.
- Si, Y., Dane, F., Rashotte, A., Kang, K. and Singh, N.K., 2010. Cloning and expression analysis of the Ccrboh gene encoding respiratory burst oxidase in *Citrullus colocynthis* and grafting onto *Citrullus lanatus* (watermelon). *Journal of Experimental Botany*, 61(6): 1635-1642.
- Soltani Howyzeh, M., Sadat Nouri, S.A., Shariati, V. and Amiripour, M., 2018. Comparative transcriptome analysis to identify putative genes involved in *thymol biosynthesis* pathway in medicinal plant *Trachyspermum ammi* L. *Scientific Reports*, 8(1): 13405.
- Sun, H., Li, F., Xu, Z., Sun, M., Cong, H., Qiao, F. and Zhong, X., 2017. De novo leaf and root transcriptome analysis to identify putative genes involved in triterpenoid saponins biosynthesis in *Hedera helix* L. *PLOS ONE*, 12(8): e0182243.
- Tang, Q., Ma, X., Mo, C., Wilson, I.W., Song, C., Zhao, H., Yang, Y., Fu, W. and Qiu, D., 2011. An efficient approach to finding *Siraitia grosvenorii* triterpene biosynthetic genes by RNA-seq and digital gene expression analysis. *BMC Genomics*, 12(1): 343.
- Tao, T., Liu, X., Chang, J., Xu, F. and Yin, Y., 2016. Cloning and characterisation of the gene encoding acetyl-coa c-acetyltransferase in *Matricaria chamomilla*. *Journal of Pharmaceutical, Chemical and Biological*, 4(3): 386-393.
- Thimmappa, R., Geisler, K., Louveau, T., O'Maille, P. and Osbourn, A., 2014. Triterpene biosynthesis in plants. *Annual Review of Plant Biology*, 65: 225-257.
- Tian, X., Ruan, J.X., Huang, J.Q., Yang, C.Q., Fang, X., Chen, Z.W., Hong, H., Wang, L.J., Mao, Y.B. and Lu, S., 2018. Characterization of gossypol biosynthetic plant by RNA sequencing. *Iranian Journal of Medicinal and Aromatic Plants Research*, 35(4): 691-702.
- Gurudeeban, S. and Ramanathan, T., 2010. Antidiabetic effect of *Citrullus colocynthis* in alloxan-induced diabetic rats. *Inventi Rapid: Ethnopharmacology*, 1: 12.
- Han, X.J., Wang, Y.D., Chen, Y.C., Lin, L.Y. and Wu, Q.K., 2013. Transcriptome sequencing and expression analysis of terpenoid biosynthesis genes in *Litsea cubeba*. *PLOS ONE*, 8(10): e76890.
- Javadzadeh, H.R., Davoudi, A., Davoudi, F., Valizadegan, G., Goodarzi, H., Mahmoodi, S., Ghane, M.R. and Faraji, M., 2013. *Citrullus colocynthis* as the cause of acute rectorrhagia. *Case Reports in Emergency Medicine*, 5p.
- Jiang, Z., Kempinski, C., Bush, C.J., Nybo, S.E. and Chappell, J., 2016. Engineering triterpene and methylated triterpene production in plants provides biochemical and physiological insights into terpene metabolism. *Plant Physiology*, 170(2): 702-716.
- Langenheim, J.H., 1994. Higher plant terpenoids: a phytocentric overview of their ecological roles. *Journal of Chemical Ecology*, 20(60): 1223-1280.
- Lee, H.A., Kim, S., Kim, S. and Choi, D., 2017. Expansion of sesquiterpene biosynthetic gene clusters in pepper confers nonhost resistance to the Irish potato famine pathogen. *New Phytologist*, 215(3): 1132-1143.
- 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(5): S5.
- Miettinen, K., Inigo, S., Kreft, L., Pollier, J., De, Bo. C., Botzki, A., Coppens, F., Bak, S. and Goossens, A., 2017a. The TriForC database: a comprehensive up-to-date resource of plant triterpene biosynthesis. *Nucleic Acids Research*, 46(1): D586-D594.
- Miettinen, K., Pollier, J., Buyst, D., Arendt, P., Csuk, R., Sommerwerk, S., Moses, T., Mertens, J., Sonawane, P.D. and Pauwels, L., 2017b. The ancient CYP716 family is a major contributor to the diversification of eudicot triterpenoid biosynthesis. *Nature Communications*, 8: 14153.
- Moses, T., Papadopoulou, K.K. and Osbourn, A., 2014. Metabolic and functional diversity of saponins, biosynthetic intermediates and semi-synthetic derivatives. *Critical Reviews in Biochemistry and Molecular Biology*, 49(6): 439-462.
- Moses, T., Pollier, J., Shen, Q., Soetaert, S., Reed, J., Erffelinck, M.L., Van, Nieuwerburgh, F.C., Bossche, R.V., Osbourn, A. and Thevelein, J.M., 2015. OSC2 and CYP716A14v2 catalyze the biosynthesis of triterpenoids for the cuticle of aerial organs of *Artemisia annua*. *The Plant Cell*, 27(1): 286-301.

- assembled and profiled across seasonal development. *BMC Genomics*, 14(1): 564.
- Zhan, C., Li, X., Zhao, Z., Yang, T., Wang, X., Luo, B., Zhang, Q., Hu, Y. and Hu, X., 2016. Comprehensive analysis of the triterpenoid saponins biosynthetic pathway in *Anemone flaccida* by transcriptome and proteome profiling. *Frontiers in Plant Science*, 7: 1094.
  - Zhang, W., Tao, T., Liu, X., Xu, F., Chang, J. and Liao, Y., 2018. De novo assembly and comparative transcriptome analysis: novel insights into sesquiterpenoid biosynthesis in *Matricaria chamomilla* L. *Acta Physiologiae Plantarum*, 40(7): 129.
  - Zheng, X., Xu, H., Ma, X., Zhan, R. and Chen, W., 2014. Triterpenoid saponin biosynthetic pathway profiling and candidate gene mining of the *Ilex asprella* root using RNA-Seq. *International Journal of Molecular Sciences*, 15(4): 5970-5987.
  - pathway. *Proceedings of the National Academy of Sciences*, 115(23): E5410-E5418.
  - Uma, C. and Sekar, K., 2014. Phytochemical analysis of a folklore medicinal plant *Citrullus colocynthis* L. (bitter apple). *Journal of Pharmacognosy and Phytochemistry*, 2(6): 195-202.
  - Upadhyay, B., Roy, S. and Kumar, A., 2007. Traditional uses of medicinal plants among the rural communities of Churu district in the Thar Desert, India. *Journal of Ethnopharmacology*, 113(3): 387-399.
  - Wang, Z., Hu, H., Goertzen, L.R., McElroy, J.S. and Dane, F., 2014. Analysis of the *Citrullus colocynthis* transcriptome during water deficit stress. *PLoS One*, 9(8): e104657.
  - Wu, D., Austin, R.S., Zhou, S. and Brown, D., 2013. The root transcriptome for North American ginseng

## Identification of triterpene and sesquiterpene biosynthetic pathway genes in *Citrullus colocynthis* L. fruit using Next Generation Sequencing (NGS) technology

M. Dorafshan<sup>1</sup>, M. Soltani Howyzeh<sup>2\*</sup> and V. Shariati<sup>3</sup>

1- M.Sc. student, Department of Genetics and Plant Breeding, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran

2\*- Corresponding author, Faculty member, Department of Genetics and Plant Breeding, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran, E-mail: soltani.m@iauahvaz.ac.ir

3- Faculty Member, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran

Received: October 2019

Revised: March 2020

Accepted: May 2020

### Abstract

Identification of the genes biosynthesizing medicinal plants-specific metabolites is now performed with great speed and accuracy using new transcriptome study technologies such as RNA sequencing. The present study was carried out to find the specific genes in the biosynthetic pathway of triterpenes and sesquiterpenes in the fruit tissue of colocynth (*Citrullus colocynthis* L.). After RNA extraction from the tissue of colocynth fruits harvested from Andimeshk region in Khuzestan province in 2017, RNA sequencing technique was performed using the Illumina HiSeq2500 platform. The bioinformatics steps including de novo assembly, using the Evidential-gene software, and functional annotation, using the KAAS database, were performed. In the KAAS database, 17359 unigenes were annotated in 134 plant pathways. Among the different and important secondary metabolites pathways in the fruit tissue of colocynth, 39 unigenes and 8 orthologous genes were assigned to the triterpenes and sesquiterpenes gene pathways. Transcriptome analysis of this medicinal plant with the aim of identifying the genes of secondary metabolites biosynthetic pathways underlies various research and practical aspects such as biosynthetic pathway engineering of herbal medicines.

**Keywords:** Cucurbitacine, RNA sequencing, secondary metabolites, De novo assembly, terpenes.