

ORIGINAL ARTICLE

Identifying genes related to terpenoids biosynthesis pathway and investigating phylogenetic relationship of different *Thymus* species using chloroplast data

Aboozar Soorni^{*}, Parnian Karimzadeh, Samira Dehghani

Department of Biotechnology,
College of Agriculture, Isfahan
University of Technology, Isfahan,
Iran

Correspondence

Aboozar Soorni
Email: soorni@iut.ac.ir

ABSTRACT

Thyme species are very important due to the production of secondary metabolites such as terpenoids. Since the identification of key genes such as genes related to terpenoids biosynthesis pathway can play an effective role in plant breeding programs, especially thyme species, the present study was aimed to investigate the transcriptomes of *T. daenensis*, *T. vulgaris*, *T. lancifolius*, *T. persicus*, *T. pubescens* to identify key genes in the biosynthesis of monoterpenoids, chloroplast genes sequence and evaluation of similarities and differences among these species. For this purpose, total RNAs extracted from vegetative growth were sent to Macrogen of Korea for sequencing with the Illumina HiSeq 2500 platform. After assembling the sequences using various tools, the best results was selected and transcripts were documented in different databases. Then, according to the documented results, key genes responsible in the synthesis of terpenoids and chloroplast gene sequence were identified, and then phylogenetic relationships among species was investigated. According to the evaluation indicators, the best assembly was a product of Binpacker tools. Based on the results, the sequence of 10 genes involved in the synthesis of terpenoids was obtained. Interestingly, among the identified TPSs, most of the contigs were classified into the TPSb and TPSa classes of terpenoids. The sequence of 73 chloroplast genes was extracted from the transcriptome data and finally the phylogenetic relationship was evaluated according to 400, 70 bp of cpDNA. The study of phylogenetic relationships showed a close genetic relationship between *T. daenensi* and *T. vulgaris* which can introduce *T. daenensis* as an appropriate replacement for *T. vulgaris* in different purpose, especially in pharmacological applications. The results show that *Z. multiflora* can most probably be as one of the ancestors of *Thymus*, which is significantly different from *Thymus* species in terms of its genetic structure, especially the key genes of the terpene biosynthesis pathway.

KEYWORDS

Thyme, Transcriptome, Phylogenetic relationships, Terpenoids, Chloroplast.

How to cite

Soorni, A., Karimzadeh, P., & Dehghani, S. (2023). Identification of genes related to the terpenes pathways and evaluation of phylogenetic relationship of different *Thymus* species using chloroplast data. *Crop Biotechnology*, 12(41), 23-36.

نشر به علمی

زیست فناوری گیاهان زراعی

«مقاله پژوهشی»

شناسایی ژن‌های مرتبط با مسیر بیوسنتز ترپنوئیدها و بررسی روابط فیلوژنتیکی گونه‌های مختلف آویشن با استفاده از اطلاعات کلروپلاستی

ابوذر سورنی*، پرنیان کریمزاده، سمیرا دهقانی

گروه زیست‌فناوری، دانشکده کشاورزی، دانشگاه صنعتی اصفهان، اصفهان، ایران

چکیده

گونه‌های آویشن به دلیل تولید متابولیت‌های ثانویه‌ای مانند ترپنوئیدها از اهمیت ویژه‌ای برخوردار هستند. از آنجایی که شناسایی ژن‌های کلیدی مانند ژن‌های مسیر بیوسنتز ترپنوئیدها می‌تواند نقش موثری در برنامه‌های اصلاحی گیاهان به ویژه گونه‌های آویشن داشته باشد، این مطالعه با هدف بررسی ترنسکریپتوم گونه‌های *T. daenensis* و *T. vulgaris*، *T. pubescens*، *T. persicus*، *T. lancifolius* به منظور شناسایی ژن‌های کلیدی بیوسنتز مونوترپن‌ها، توالی ژن‌های کلروپلاستی و بررسی تفاوت و تشابه میان گونه‌ها انجام گرفت. بدین منظور RNAهای کل استخراج شده از مرحله رویشی جهت توالی‌یابی با پلتفرم Illumina HiSeqTM 2500 به شرکت ماکروژن کره ارسال گردیدند. پس از سرهم‌بندی توالی‌ها با ابزارهای مختلف، بهترین نتیجه انتخاب و مستندسازی ترنسکریپت‌ها در پایگاه‌های اطلاعاتی انجام شد. در نهایت بر اساس نتایج مستندسازی، ژن‌های کلیدی مسیر سنتز ترپنوئیدها، و توالی‌های کلروپلاستی شناسایی شده و روابط فیلوژنتیکی میان گونه‌ها مورد بررسی قرار گرفت. بر اساس شاخص‌های ارزیابی بهترین سرهم‌بندی محصول ابزار Binpacker بود. بر اساس نتایج حاصل توالی ۱۰ ژن درگیر در مسیر سنتز ترپنوئیدها بدست آمد. از میان *Terpene synthase (TPSs)* شناسایی شده، بیشترین کانتینگ‌ها به ترتیب مربوط به ترپن‌های دسته *TPSa* و *TPSb* بود. از اطلاعات ترنسکریپتومی، توالی ۷۳ ژن کلروپلاستی استخراج شدند. از اطلاعات کلروپلاستی بدست آمده، در مجموع ۴۰۰۷۰bp در بررسی روابط فیلوژنتیکی مورد استفاده قرار گرفت. نتایج حاصل از بررسی روابط فیلوژنتیکی، ارتباط ژنتیکی نزدیک بین *T. daenensis* و *T. vulgaris* را نشان داد که می‌تواند معرف گونه *T. daenensis* به‌عنوان جایگزینی مناسب برای گونه رسمی و اروپایی *T. vulgaris* جهت مقاصد مختلف بال‌اخص کاربردهای دارویی باشد. نتایج این مطالعه نشان می‌دهد به احتمال زیاد می‌توان *Z. multiflora* را به‌عنوان یکی از اجداد *Thymus* در نظر گرفت که به لحاظ ساختار ژنتیکی به خصوص ژن‌های کلیدی مسیر بیوسنتز ترپن‌ها تفاوت قابل توجهی با گونه‌های جنس آویشن دارد.

واژه‌های کلیدی

آویشن، ترنسکریپتوم، روابط فیلوژنتیکی، ترپنوئیدها، کلروپلاست.

نویسنده مسئول:

ابوذر سورنی

رایانامه: soorni@iut.ac.ir

استاد به این مقاله:

سورنی، ابوذر، پرنیان کریمزاده، سمیرا دهقانی (۱۴۰۲). شناسایی ژن‌های مرتبط با مسیر بیوسنتز ترپنوئیدها و بررسی روابط فیلوژنتیکی گونه‌های مختلف آویشن با استفاده از اطلاعات کلروپلاستی. فصلنامه علمی زیست فناوری گیاهان زراعی، ۱۳(۴۱)، ۲۳-۳۶.

مقدمه

گونه‌های آویشن می‌باشد که این امر موجب ایجاد کموتایپ‌های تیمول^۲، کارواکرول^۳، آلفا ترپینئول^۴، آلفا توجون^۵، ژرانیول^۶، لینالول در این جنس شده است (Izadiyan et al., 2018). اما امروزه یکی از چالش‌های محققین، ابداع تکنیک‌های نوین، جهت شناسایی مسیر تولید ترکیبات ثانویه می‌باشد، زیرا بدون درک درست از مسیر سنتز متابولیت‌ها، رسیدن به اهدافی از جمله افزایش میزان متابولیت‌ها، تولید متابولیت‌های جدید و گسترش مارکرهای مولکولی بسیار دشوار، زمان‌بر و پرهزینه است.

در سال‌های اخیر فناوری زیستی با استفاده از روش‌هایی نظیر تجزیه و تحلیل ترنسکریپتوم به کمک تکنیک RNA-Seq کارایی گیاهان دارویی را در جهت تولید دارو افزایش داده است. علی‌رغم پیشرفت‌های اخیر، دانش ما از اطلاعات ژنومی، ترنسکریپتومی و ژن‌های کلیدی دخیل در سنتز متابولیت‌های ثانویه کلیدی آویشن بسیار محدود است. از طرف دیگر اهمیت مونوترپن‌ها باعث شده که تعریف‌های زیادی در مورد مسیر بیوسنتزی آن‌ها در گیاهان ارائه شود و همچنین ژن‌های دخیل در بیوسنتز این محصولات همسانه‌سازی شود و تکنیک‌های دست‌ورزی ژنتیکی برای افزایش معنی‌دار تولید مونوترپن‌ها در گیاهان مورد استفاده قرار گیرد.

با توجه به اهمیت اقتصادی این ترکیب‌های ترپنی که خزانه غنی را برای ترکیبات دارویی فراهم می‌سازد، با آشکارسازی تغییرات پویای این گونه متابولیت‌های ثانویه به‌طور هم‌زمان با داده‌های ترنسکریپتومی باعث فهم بهتر شبکه‌های تنظیمی بیوسنتز این مواد خواهد شد که درنهایت می‌تواند مهندسی متابولیک این ترکیبات را تسهیل کند. در واقع اطلاعات موجود مبین این موضوع است که آویشن یک مدل تجربی مناسب برای بررسی وقایع تنظیمی در بیان ترپن سینتاز و تجمع ترپن فرار در طول رشد و تکامل گیاه می‌باشد. با این حال، مطالعه بر روی جنبه‌های مولکولی بیوسنتز ترپن‌ها و همچنین شناسایی و تعیین توالی ژن‌های مهم در تولید متابولیت‌های بالارزش در این گیاه انجام نگرفته است. از طرف دیگر علیرغم اهمیت زیاد گونه‌های آویشن از نظر پراکنش و متابولیت‌های ثانویه، اطلاعات محدودی از روابط فیلوژنتیک برخی از گونه‌ها در دسترس است. امروزه پیشرفت‌های جدید در زمینه‌ی مطالعات مولکولی، امکانات

آویشن با نام عمومی Thyme گیاهی دولپه متعلق به خانواده نعناعیان (Lamiaceae) و زیر خانواده Nepetoideae است که از نظر فیلوژنی با جنس‌های *Zataria*، *Origanum* و *Micromeria* قرابت و خویشاوندی دارد (Jamzad et al., 2009). کشور ایران به دلیل وسعت و تنوع شرایط اکولوژی و همچنین تنوع و انعطاف‌پذیری این گیاه، تعداد قابل‌توجهی از گونه‌های آویشن را در خود جا داده است (Omidbeigi et al., 2014)، به طوری که تعداد ۱۸ گونه آویشن در فلور ایران یافت می‌شود. از این تعداد ۱۴ گونه و زیرگونه توسط رشینگر^۱ در فلور ایرانیکا گزارش شده که بیش از هشت گونه آن اندمیک ایران می‌باشد. بیش‌ترین پراکنش این گیاه در شمال و غرب کشور است که هرچه به طرف جنوب و شرق کشور پیش رفته تعداد آن‌ها کاهش می‌یابد (Jamzad et al., 1994). در میان گونه‌های مختلف، آویشن دناهی گونه‌ی انحصاری ایران می‌باشد که ضمن کمیت و کیفیت اسانس بالا (درصد تیمول و کارواکرول بیشتر در مقایسه با سایر گونه‌های دیگر) (Rustaiee et al., 2011)، پراکنش نسبتاً وسیعی در کشور دارد که بیانگر سازگاری و متحمل بودن این گیاه با شرایط محیطی متفاوت است (Jamzad et al., 2009). با توجه به درصد بالای تولید اسانس در زیرگونه دناهی و نیز درصد بسیار بالای ترکیبات فنلی در آن (به‌ویژه تیمول)، می‌توان این زیرگونه را به‌عنوان جایگزین مناسبی برای گونه رسمی و اروپایی *T. vulgaris* جهت مقاصد مختلف بال‌اخص کاربردهای دارویی در نظر گرفت. همچنین با توجه به سازگاری مناسب آن به مناطق اقلیمی مختلف ایران یکی از گونه‌های کاندید جهت اهلی‌سازی و کشت صنعتی به شمار می‌رود (Rustaiee et al., 2011).

در خصوص گونه‌های آویشن، مبحث شناسایی و تعیین موقعیت تاکسونومیک گونه‌های مختلف به دلیل هیبریداسیون بین‌گونه‌ای در این جنس دشوار است. علاوه بر هیبریداسیون، تنوع ریخت‌شناسی موجود در گونه‌ها نیز مسئله شناسایی را دشوار می‌کند (Haji Akhondi & Jamzad et al., 1994; Farahani, 2007). به دلیل دگرگرفته‌افشان بودن و هیبریداسیون‌های فراوان بین‌گونه‌ای، تنوع ژنتیکی و در نتیجه آن تنوع مورفولوژیکی، تنوع در ترکیبات شیمیایی و بازده اسانس در این جنس به‌وفور دیده می‌شود. چندانکی شیمیایی از خصوصیات

2. Thymol
3. Carvacrol
4. α -Terpineol
5. α -Thujone
6. Geraniol

1. Rechingher

کاماترین سنتاز، CYP71D178 و CYP71D180 در بافت‌های گل و برگ آویشن باغی نشان می‌دهد سطح بیان این ژن‌ها در گل بالاتر از برگ می‌باشد که احتمالاً به دلیل تراکم بالای کرک‌ها در این بافت است (MashhadyMalekzadeh *et al.*, 2017).

بررسی میزان نسبی ژن‌های مسیر سنتز تیمول (MVA) (MEP) در آویشن باغی رقم واریکو ۳ تحت تنش سرمایی نشان داد که تنش سرمایی بر میزان بیان ژن‌ها اثر معنی‌داری دارد و میزان بیان ژن TPS5 در بازه‌های متفاوت زمانی تنش افزایش می‌یابد و در بازه زمانی ۴۸ ساعت به بیشترین بیان می‌رسد و در ژن TPS2 بیشترین میزان بیان در بازه زمانی ۳ ساعت ۳/۴۶ برابر نسبت به شاهد مشاهده شد (Habibi *et al.*, 2020). مطالعه بیان ژن‌های کلیدی در بیوسنتز مونوترپن‌ها در بافت‌های مختلف و در پاسخ به الیسیتورهای غیرزیستی در گیاه دارویی مرزه تابستانه (*Hortensis Satureja*) نشان داد که هر دو ژن DXR و GTS در بافت‌های مختلف (ریشه، ساقه، برگ و گل آذین) دارای بیان‌های متفاوتی هستند و در بافت‌های هوایی به‌ویژه گل آذین و برگ دارای بیان بیشتری می‌باشند. همچنین بیان این ژن‌ها تحت تاثیر الیسیتورهای غیر زیستی شامل اسید سالیسیلیک، متیل جاسمونات و اشعه UV-B تغییرات چشم‌گیری نشان می‌دهد (Ghobadi *et al.*, 2016).

در ارتباط با ژن‌های کلروپلاستی جنس *Thymus* در تحقیقی که توسط زونیکا و همکاران انجام شد، قسمت‌هایی از توالی کلروپلاستی ژن‌های *matK*, *rbcL*, *psbA* در گونه‌ای از این گیاه با نام *T. serpyllum* گزارش شده است. علاوه بر این اطلاعات مربوط به بخشی‌هایی از توالی ژن‌های کلروپلاستی *matK* *rpl16* و *rbcL* از گونه‌ی *T. vulgaris* در بانک اطلاعات موجود است (Bunsawat *et al.*, 2004; Federici *et al.*, 2013) که در مجموع نشان از عدم وجود اطلاعات کافی در این زمینه و مباحث پیرامون آن، برای گیاه آویشن است. در نتیجه بررسی ژن‌های کلروپلاستی گونه‌های جنس *Thymus* در جهت تحلیل و بررسی روابط فیلوژنتیکی و تکاملی ضروری به نظر می‌رسد.

روش شناسی پژوهش

جمع‌آوری نمونه‌ی گیاهی، استخراج RNA کل از برگ‌ها و توالی‌یابی

بذور گونه‌های *T. pubescens* *T. persicus* *T. lancifolius* *T. vulgaris* و *T. daenensis* از موسسه جنگل‌ها و مراتع

مناسبی در جهت پاسخگویی به بسیاری از پرسش‌ها در زمینه‌ی رده‌بندی گونه‌ها، روابط فیلوژنتیک و تکاملی گونه‌ها فراهم کرده‌است. لذا به دلیل اهمیت مطالعات فیلوژنتیک در درک چگونگی تکامل ژن‌ها، ژنوم‌ها و گونه‌ها، مطالعه‌ی حاضر با هدف شناسایی ژن‌های کلیدی مسیر بیوسنتز ترپنوئیدها، ژن‌های کلروپلاستی و بررسی روابط فیلوژنتیکی و تخمین زمان انشقاق گونه‌های آویشن بر مبنای اطلاعات کلروپلاستی به روش بی‌زین انجام گرفت.

پیشینه پژوهش

از آنجایی در خصوص شناسایی تمامی ژن‌های مسیر بیوسنتز ترپنوئیدها و تعیین توالی آن‌ها در آویشن مطالعه جامعی صورت نگرفته است، اما در خصوص بررسی روابط فیلوژنتیکی گونه‌ها و ارزیابی تنوع ژنتیکی آن‌ها مطالعات زیادی صورت گرفته است، هر چند این ارزیابی‌ها بر مبنای تعداد محدودی ژن بوده است. در بحث تنوع گیاه آویشن، نتایج بررسی تنوع ژنتیکی ۱۵ توده آویشن شیرازی با استفاده از نشانگرهای RAPD و ISSR، نشان داد که تنوع ژنتیکی نسبتاً بالایی در میان این توده‌ها وجود دارد. در مطالعه‌ی دیگری که مجیری و همکاران در سال ۹۲ بر روی ۷۰ توده‌ی آویشن انجام دادند، این توده‌ها به شش گروه تقسیم شدند که بیشترین شباهت بین دو توده‌ی *T. kotschyanus* و *T. transcaucasicus* از قزوین و گیلان و کمترین شباهت مربوط به توده *T. lancifolius* فارس و *T. fedtschenkoi* از آذربایجان غربی بود. نتایج تحقیقی که در سال ۲۰۰۹ بر روی تنوع ژنتیکی، مورفولوژیکی و فیلوژنتیکی هفت توده آویشن بر اساس نشانگر ISSR انجام شد، نشان داد که از میان ۲۰ آغازگر، ۷ آغازگر باندهای چندشکل تولید کردند. در درخت فیلوژنتیکی ترسیم شده بر اساس این نشانگر دو کلاد اصلی A و B شکل گرفت که گونه‌های *T. serpyllum*، *T. vilosus* و *T. praecox* از دیگر گونه‌ها تفکیک شدند. در این مطالعه گونه‌ی *T. vulgaris* خارج از هر دو کلاد قرار گرفت (Smolik *et al.*, 2009).

در خصوص بررسی سطح بیان ژن‌های درگیر در مسیر سنتز ترکیباتی همچون تیمول و کارواکرول در گیاه آویشن و برخی گونه‌های گیاهی نزدیک به آن نیز مطالعاتی صورت گرفته است، که محدود به یک یا چند ژن می‌باشد. به عنوان مثال الگوی بیان ژن‌های دخیل در بیوسنتز تیمول و کارواکرول شامل DXR،

نمایند به دست آمدند. در این خصوص آستانه کیفیت بازها ۲۰ از میانگین ۴ باز و حداقل طول ۵۰bp در نظر گرفته شد.

سرهم‌بندی قرائت‌ها

خوانش‌های با کیفیت بالا متعلق به هر گونه به صورت جداگانه با پارامترهای بهینه شده با استفاده از نرم‌افزارهای Trinity v2.4.0 (Liu *et al.*, 2011) و Binpacker v1.0 (Grabherr *et al.*, 2011) بر اساس تک K-mer انتخابی ۲۵ به صورت *De novo* Transcriptome Assembly (سرهم‌بندی بدون استفاده از ژنوم مرجع و به کمک قطعات هم‌پوشان) و با استفاده از سیستم کامپیوتری با مشخصات CPU Core 64, RAM 256G و سیستم‌عامل (Linux ubuntu v16) سرهم‌بندی شدند. کانتیگ‌های به دست آمده از سرهم‌بندی خوانش‌ها با پارامترهایی همچون شمار کانتیگ‌ها، اندازه‌ی کل، میانگین طول کانتیگ‌ها و N50 بدست آمده از اطلاعات آماری هر ابزار ارزیابی شدند.

ترسیم مسیر سنتز ترپنوئیدها (تیمول - کارواکرول)

رونوشت‌های به دست آمده از بهترین نتیجه سرهم‌بندی توسط ابزار Trinotate v2.0.2 (<http://trinotate.github.io/>) مستندسازی شدند. بر اساس نتایج حاصل از مستندسازی رونوشت‌ها، ژن‌های دخیل در مسیر سنتز تیمول و کارواکرول شناسایی شده و مسیر کلی سنتز این ترکیبات بر اساس ژن‌های کد کننده ترسیم شد. در این خصوص به منظور افزایش سطح اعتماد به توالی‌های شناسایی شده از بانک اطلاعاتی Terzyme (Priya *et al.*, 2018) استفاده شد. بدین منظور نتایج سرهم‌بندی برای تمامی گونه‌ها در بانک اطلاعاتی Terzyme مورد کاوش و تجزیه و تحلیل قرار گرفتند تا با استفاده از این بانک اطلاعاتی توالی‌های مربوط به ترپنوئیدسنتازها شناسایی و طبقه‌بندی شوند.

شناسایی ژن‌های کلروپلاستی و استخراج قطعات مشترک

برای شناسایی ژن‌های کلروپلاستی، از جنس‌های نزدیک به جنس *Thymus* (به عنوان رفرنس استفاده شد، به این ترتیب که فایل Fasta توالی کامل کدشونده^۳ گونه‌ی *Mentha logifolia* با شماره دستیابی NC-032054 و گونه‌ی *Salvia officinalis* با شماره دستیابی NC-038165 از پایگاه داده‌ی NCBI (www.ncbi.nlm.nih.gov) به دست آمد. هر دو رفرنس (*Salvia*,

کشور تهیه و به مدت ۴۵ روز در گلدان‌های پلاستیکی تحت شرایط گلخانه‌ای کشت داده شدند. سپس برگ‌های جوان در مرحله رشد رویشی جهت استخراج RNA با استفاده از فویل آلومینیومی جمع‌آوری و بلافاصله در ازت مایع قرار داده شدند. پس از انتقال به آزمایشگاه تا زمان استخراج RNA در فریزر در دمای °C ۸۰- نگهداری گردیدند. استخراج RNA کل از برگ‌های جمع‌آوری شده با استفاده از کیت RNeasy mini kit (Qiagen) انجام شد. جهت بررسی کمیت و کیفیت RNA، ۲ میکرولیتر از RNA به همراه ۰/۵ میکرولیتر GelRed (مدل Biotium)، بر روی ژل آگارز ۱ درصد با استفاده از بافر یک درصد (Tris 40mM, Acetic acid 20m M, EDTA 1mM) با ولتاژ ۸۵ ولت بر سانتی‌متر به مدت یک ساعت الکتروفورز شد و سپس باندهای RNA ریوزومی ۲۸s و ۱۸s آن توسط دستگاه Viber Lourmat gel document تحت نور فرابنفش با طول موج ۳۱۲ نانومتر عکس‌برداری و مورد بررسی قرار گرفت. همچنین برای تعیین کمیت RNA از دستگاه نانودراپ (مدل Implen)، استفاده شد. پس از تایید کمیت و کیفیت RNAهای استخراج شده، ساخت کتابخانه‌ی cDNA (بعد از حذف حضور احتمالی DNA از طریق استفاده از آنزیم DNaseI) و توالی‌یابی با خوانش دوسویه (رفت^۱ و برگشت^۲) به طول ۱۵۰ pb با استفاده از پلتفرم Illumina HiSeqTM 2500 در شرکت ماکروژن کره انجام شد.

ارزیابی کیفیت و ویرایش توالی‌ها

نتایج حاصل از دستگاه توالی‌یابی خوانش‌های خام نامیده می‌شوند که در قالب فایل Fastq ذخیره شده‌اند. داده‌های خام گاهی حاوی خوانش‌های نامطلوبی مانند آداپتورها و خوانش‌های با کیفیت پایین می‌باشند که باید حذف شوند. به همین دلیل بعد از دریافت اطلاعات، می‌بایست توالی‌ها توسط ابزارهای کنترل کیفیت مورد بررسی قرار گیرند. در این تحقیق کنترل کیفیت توالی‌ها توسط نرم‌افزارهای FastQC v0.11 (Andrews *et al.*, 2015) صورت گرفت. با کمک برنامه‌ی Trimmomatic v0.32 (Bolger *et al.*, 2014) تصحیح و ویرایش توالی‌ها انجام شد و طی آن خوانش‌های تمیز با حذف آداپتورها، آغازگرها و توالی‌های بی‌کیفیت که می‌توانند در آنالیزهای بعدی مشکلات اساسی ایجاد

می‌پردازد که برای این پارامتر از منوی کشویی مقدار تجربی^۸ انتخاب شد، که فراوانی‌ها را به نسبت مشاهدات در داده‌ها تنظیم می‌کند. در انتها، گزینه نرخ متوسط جهش^۹ نیز انتخاب شد که در واقع به هر پارتیشن اجازه می‌دهد نرخ نسبی جهش تخمین زده-شده‌ی مرتبط با خود را داشته‌باشد. در نهایت، تمامی تنظیمات این بخش به کلیه توالی‌ها ارجاع داده شد. ساعت مولکولی آرام در تنظیمات مدل ساعت^{۱۰} به صورت ساعت مولکولی سخت‌گیرانه^{۱۱} انتخاب شد. درخت پریور با استفاده از مدل Calibrated yule که مدلی ساده از گونه‌زایی است انتخاب و بر اساس اطلاعات به دست آمده از سایت <http://www.timetree.org/> (Kumar et al., 2017) برای جنس *Salvia* و *Mentha* در مقابل جنس *Thymus* توزیع پریور بر روی گره‌های کالیبره شده به صورت ۳۱/۶ میلیون سال قبل برای *Salvia-Thymus* و ۵۰ میلیون سال قبل برای *Mentha-Thymus* تعریف شد. از مجموعه‌ی درختان، بهترین درخت توافقی همراه با میانگین سنی تمامی گره‌ها بوسیله‌ی نرم‌افزار TreeAnnotator v1.6.1 (موجود در پکیج beast)، حاصل شد. فایل به دست آمده از TreeAnnotator، به برنامه FigTree v1.4.0 (Rambaut & Drummond, 2018) وارد و پس از نمودن درخت و علامت‌دار کردن گزینه‌ی برجسب‌گذاری گره‌ها^{۱۲}، درخت فیلوژنتیکی با زمان انشقاق تخمین زده شده برای هر گره، نمایش داده شد.

یافته‌های پژوهش

کنترل کیفیت داده‌های توالی‌یابی

در این مطالعه، با توجه به اطلاعات حاصل FastQC بازهای بی کیفیت از توالی‌ها حذف شدند، همچنین برای بهبود کیفیت سرهم‌بندی‌ها، بخشی از خوانش‌ها (۲۲/۵۹٪ تا ۳۲/۴۱٪)، با حذف توالی‌های آداپتوری، بازها و خوانش‌های بی کیفیت و کوتاه (کوتاه‌تر از ۵۰bp) نیز حذف شدند. خروجی این مرحله آماره‌های پایه نظیر تعداد خوانش‌ها و اطلاعات کیفی که تصمیمات پیش‌پردازش در مرحله‌ی بعد را هدایت می‌کنند، است. تعداد خوانش برای پنج گونه مذکور قبل و بعد از ویرایش مطابق جدول ۱ می‌باشد.

Mentha) که محتوی ۸۷ ژن کلروپلاستی بودند، در مقابل فایل‌های سرهم‌بندی شده از تمامی گونه‌های آویشن بلاست نوکلئوتید شدند. به این ترتیب که ابتدا یک پایگاه داده‌ی^۱ سفارشی از فایل‌های سرهم‌بندی شده‌ی گونه‌های آویشن ایجاد شد و در نهایت بلاست نوکلئوتید برای فایل‌های رفرنس (به عنوان پرسش^۲) و فایل ترنسکریپتوم گونه‌ها (به عنوان پایگاه داده) اجرا شد. اطلاعات حاصل از بلاست به صورت جدولی^۳ گزارش شد (6-outfmt-). بررسی نتیجه‌ی بلاست هر ژن با فراخوانی نام ژن مورد نظر از فایل خروجی بلاست انجام گرفت. غربالگری نتایج بلاست به صورت انتخاب بیشترین نقاط هدف با بیشترین شباهت^۴ به ازای هر ژن در گونه‌های رفرنس و گونه‌های مورد بررسی، انجام شد. پس از تعیین بهترین شناسه‌ها، از نقاط شروع و پایان تعیین شده در تمامی گونه‌ها برای توالی هر ژن، فایل bed جهت استخراج توالی‌ها تهیه شد. جهت استخراج قطعات هدف مشترک میان گونه‌ها و مرجع (حاصل از بلاست) از داده‌های سرهم‌بندی شده (فایل ورودی) بر اساس فواصل تعیین شده در فایل bed، از نرم‌افزار (Quinlan) bedtools v2.28.0، از نرم‌افزار (Hall, 2010) استفاده شد.

همردیفی توالی‌ها و بررسی روابط فیلوژنتیکی

در این مرحله از نرم‌افزار Muscle v3.8.31 (Edgar, 2004) برای همردیف کردن توالی‌های چندگانه و trimAL v1.4 (Capella-Gutierrez et al., 2009) جهت حذف مناطق ضعیف همردیفی‌های چندگانه، استفاده شد. به منظور بررسی روابط فیلوژنتیکی و تخمین زمان انشقاق میان گونه‌ها ابزار Beast (Rambaut & Drummond, 2014) 2.5.2 استفاده شد. در تنظیمات مربوط به تعریف مدل سایت^۵ در قسمت مدل تکاملی (جایگزینی)، برای شمار دسته‌ی گام^۶، عدد ۴ انتخاب گردید. این تنظیمات امکان مدل‌سازی نرخ تغییرات بین سایت‌ها در هر پارتیشن را فراهم می‌کند. در قسمت مدل جایگزینی (تکامل)، مدل متناسب با داده‌های ما در هر پارتیشن HKY انتخاب شد همچنین برای تخمین جایگزینی، پارامتر تخمین نیز علامت زده شد. در این بخش پارامتر فراوانی^۷ به بررسی فراوانی نوکلئوتیدها

1. Database
2. Query
3. tabular
4. Max similarity
5. Site model
6. Gamma Category Count
7. Frequency

8. Empirical
9. Fix mean mutation rate
10. Clock model
11. Strict molecular clock
12. Node label

جدول ۱. تعداد خوانش‌ها قبل و بعد از فرآیند ویرایش

<i>T. lancifolius</i>	<i>T. persicus</i>	<i>T. pubescens</i>	<i>T. vulgaris</i>	<i>T. daenensis</i>	
۶۵۸۵۹۴۵۶	۶۱۸۱۰۳۴۴	۶۲۲۰۹۱۱۸	۶۴۹۹۷۲۳۲	۶۰۹۸۲۲۱۲	تعداد خوانش قبل از ویرایش
۴۷۱۰۵۲۱۰	۴۷۶۲۴۳۰۲	۴۳۳۸۷۰۷۴	۴۳۹۲۷۰۲۸	۴۳۳۸۴۱۵۲	تعداد خوانش بعد از ویرایش

خطاهای توالی‌یابی و عدم هم‌پوشانی خوانش‌ها اشاره کرد همچنین زمان استخراج RNA و سطح بیان ژن‌ها تأثیر بسزایی در تعداد کانتیگ‌های حاصله دارد، در واقع از اطلاعات ترنسکرپت‌ها با بیان بالا، سرهم‌بندی‌های پیوسته‌تری ایجاد می‌شود. در واقع شمار بالای کانتیگ‌ها، نشان‌دهنده تکه‌تکه بودن ترنسکرپت مرجع و شباهت بالای بسیاری از کانتیگ‌ها به یکدیگر هست. بنابراین کاهش شمار کانتیگ‌ها و افزایش اندازه‌ی کل، سرهم‌بندی‌های پیوسته‌تر با پوشش بالاتر را ایجاد می‌کند (Rana et al, 2016; Smith-Unna et al, 2016). از طرف دیگر بدیهی است که هرچه طول ترنسکرپت‌ها و تعداد آن‌ها بیشتر باشد در جستجوی بلاست تعداد رکوردهای بیشتری در پایگاه‌های اطلاعاتی برای آن‌ها به دست می‌آید و در نتیجه تفسیر عملکردی دقیق‌تری از ترنسکرپتوم و ژن‌های بیان‌شده حاصل می‌گردد.

بر اساس جدول ۲ در پنج گونه‌ی آویشن بیش‌ترین تعداد کانتیگ توسط ابزار Trinity تولید شد درحالی‌که در این گونه‌ها بیشترین اندازه کل سرهم‌بندی در ابزار Binner مشاهده شد. علاوه بر بررسی توزیع اندازه کانتیگ‌ها، شاخص N50 به‌عنوان معیار کیفیت سرهم‌بندی برای هر نمونه محاسبه گردید. از نظر تئوری، سرهم‌بندی‌های بهتر دارای کانتیگ‌های بزرگ‌تر با N50 بالاتر هستند، البته اهمیت این پارامتر همچنان مورد بحث محققین می‌باشد (Salzberg et al, 2012). چراکه N50 بزرگ‌تر لزوماً به معنای تولید سرهم‌بندی‌های بهتر نیست و ممکن است نشان‌دهنده‌ی چایمریسم^۲ (ادغام دو یا چند خوانش در یک کانتیگ در طول سرهم‌بندی) باشد، هرچند مطالعات قبلی روی گیاهان مختلف نشان داده است که N50 بالاتر می‌تواند به‌طور مصنوعی منعکس‌کننده‌ی کیفیت بالای سرهم‌بندی باشد (Schliesky et al., 2012; Liu et al., 2016). در مطالعه حاضر مقدار این شاخص از ۷۳۲bp نوکلئوتید در گونه *T. vulgaris* تا ۹۹۹bp در گونه *T. persicus* برای سرهم‌بندی‌های بدست آمده از ابزار Trinity متغیر بود. از دیدگاه مقایسه، ابزار Binner توانست سرهم‌بندی‌هایی با مقادیر بالاتر N50

سرهم‌بندی ترنسکرپتوم به‌صورت *De novo* و بررسی کیفیت از آنجا که برای آویشن ژنوم مرجعی وجود ندارد، استفاده از مسیر سرهم‌بندی از نو جهت بازسازی ترنسکرپتوم و آنالیزهای بعدی اجتناب‌ناپذیر است. برای سرهم‌بندی خوانش‌ها نرم‌افزارهای زیادی توسعه پیدا کرده‌اند که در پارامترهایی همچون k-mer، N50، تعداد و طول ترنسکرپت‌ها تفاوت نشان می‌دهند. با مقایسه عملکرد نرم‌افزارهای مختلف می‌توان بر اساس هدف و گونه یا جنس مورد مطالعه مناسب‌ترین نرم‌افزار و ترنسکرپتوم مرجع را انتخاب کرد.

به‌منظور ساخت رونوشت‌های بزرگ‌تر، سرهم‌بندی برای هر یک از نمونه‌ها به‌طور جداگانه توسط پکیج Trinity و Binner با k-mer ۲۵ صورت گرفت. بر اساس داده‌های حاصل از پکیج Trinity تعداد کانتیگ‌های کل ترنسکرپت‌های موجود از ۸۵۰۸۴ برای گونه *T. persicus* تا ۱۲۳۶۹۸ برای گونه *T. lancifolius* متغیر بود (جدول ۲). نتایج حاصل از Binner نشان داد تعداد کانتیگ‌های کل ترنسکرپت‌های موجود در گونه‌های *T. daenensis* و *T. lancifolius* به ترتیب کمترین (۶۶۵۹۱) و بیشترین (۹۴۹۳۹) تعداد بودند. نتایج مطالعات نشان می‌دهد که اندازه اسمبلی و تعداد کل کانتیگ‌ها، به طول k-mer بستگی دارد. از نظر تئوری، اندازه بزرگ k-mer باعث کاهش تعداد ترنسکرپت‌ها می‌شود، چرا که k-mers با طول بلندتر سرهم‌بندی‌های پیوسته‌تری از ترنسکرپت‌ها با پوشش بالاتر را ایجاد می‌کنند؛ که به دو دلیل حضور خوانش‌های با فراوانی بالا و وجود ترنسکرپت‌هایی با سطح بیان بالا می‌باشد. در مقابل، k-mers کوتاه به خاطر وجود خطاهای توالی‌یابی و عدم وجود همپوشانی میان قطعات منجر به تولید سرهم‌بندی‌هایی پاره‌پاره می‌شوند. k-mers کوچک می‌توانند به تولید چندین مسیر در گراف اسمبلی منجر شوند؛ بنابراین، ساختار گراف ممکن است ناواضح و مبهم باشد و تجزیه و تحلیل ایزوفرم‌های نمایش داده شده می‌تواند سخت و چالش برانگیز باشد. در واقع به خاطر وجود چندین مسیر مبهم، یافتن یک مسیر واحد و منحصر به فرد در سرهم‌بندی تقریباً امری غیرممکن است (Grabherr et al, 2011; Rana et al, 2016). از دیگر علل افزایش شمار کانتیگ‌ها می‌توان به

1. Fragmentation
2. chimerism

شناسایی ترپن سینتازها و بررسی ساختار موتیف‌ها

در این مطالعه، شناسایی و جداسازی ترپن سینتازها با استفاده از بررسی شباهت آن‌ها در برابر ژن‌های شناسایی شده و مرتبط در پایگاه داده‌ها انجام گرفت. همچنین برای شناسایی ترپن سینتازها از بانک اطلاعاتی Terzyme استفاده شد. بدین منظور نتایج سرهم‌بندی برای تمامی گونه‌ها مورد کاوش و تجزیه و تحلیل قرار گرفتند. بر اساس نتایج این شناسایی، ترپن‌ها در سه خانواده مونو، دی و سزکوئی ترپن‌ها و سه کلاس اصلی و پنج زیرکلاس TPSa، TPSb، TPSc، TPSf و TPSg تقسیم شدند. بر اساس نتایج حاصله زیر خانواده TPSg در گونه‌های *T. lancifolius* مشاهده نشد (جدول ۵). از نظر تعداد، بیشترین و کمترین تعداد مونوترپن‌ها در *T. pubescens* و *T. daenensis* یافت شد. *T. pubescens* کمترین تعداد دی و سزکوئی ترپن‌ها را نشان داد. در همه گونه‌ها، بیشترین تعداد TPS در ارتباط با زیرخانواده TPSb که عمدتاً حاوی مونوترپن و چندین سنتاز سزکوئی ترپن است، می‌باشد. زیرخانواده TPSa فقط شامل سنتازهای سزکوئی ترپن بود، درحالی‌که TPSc و TPSf با سنتازهای دی ترپن مرتبط هستند. زیرخانواده TPSa به‌عنوان بزرگ‌ترین زیرخانواده TPSها فقط TPS-sesqui را در گیاهان رمزگذاری می‌کند اما از دست دادن این زیرخانواده یا عملکرد اعضای آن در برخی از گونه‌ها طی یک دوره تکامل طولانی مشاهده شده است (Jiang et al., 2019) که به نظر می‌رسد مشابه نتایج به‌دست آمده در این مطالعه است.

نسبت به Trinity تولید کند به‌طوری که از ۹۸۵ نوکلئوتید در گونه *T. vulgaris* تا ۱۶۶۵ نوکلئوتید در گونه *T. persicus* متغیر بود.

جداسازی توالی ژن‌های کلیدی مسیر سنتز ترپنوئیدها

بر اساس نتایج حاصل از جستجوی همولوژی، بیشترین تعداد کانتیگ در ژن‌های *HDR*، *HDS*، *DXS* مشاهده شد (جدول ۳). ژن‌های *IDI* و *DXR* تعداد متوسطی از رونوشت‌ها را داشتند و *MDS*، *CMK*، *MCT* و *GDS* کمترین تعداد کانتیگ را داشتند. درواقع می‌توان نتیجه گرفت که تفاوت در تعداد کانتیگ‌ها می‌تواند به نوعی منعکس‌کننده‌ی تفاوت در میزان بیان و فراوانی رونوشت‌ها برای هر ژن باشد.

با توجه به رفرنس و کدون‌های آغاز و پایان می‌توان گفت در اکثر ژن‌ها در اکثر گونه‌ها توالی کامل به‌دست آمده است. در برخی ژن‌ها (مانند *DXS1*) مشاهده شد که برخی گونه‌ها (*T. lancifolius*) در توالی کامل چند باز نسبت به دیگر گونه‌ها کمتر (یا بیشتر) دارد که در طی تکامل و گونه‌زایی ایجاد شده است. بیشترین تفاوت در طول در بین گونه‌ها در ژن *IDI* در آیزوفرم ۲ مشاهده شد (جدول ۴). در ژن *DXS* در آیزوفرم‌های ۱ و ۲ نیز تفاوت در طول توالی‌ها دیده می‌شد که پیش‌بینی می‌شود در این ژن تفاوت به دلیل عدم دستیابی به طول کامل بوده است.

جدول ۲. اطلاعات ترنسکرپت‌های سرهم‌بندی شده با دو ابزار Trinity و Binpacker برای تمامی گونه‌ها

گونه	شمار کانتیگ		اندازه کل (bp)		طول میانگین (bp)		N50(bp)	
	Trinity	Binpacker	Trinity	Binpacker	Trinity	Binpacker	Trinity	Binpacker
<i>T. daenensis</i>	۹۱۱۲۷	۶۶۵۹۱	۵۵/۸۳	۶۸/۴۵	۹۱۴/۳۵	۱۰۰۲۸/۰۳	۸۱۹	۱۵۵۹
<i>T. vulgaris</i>	۱۱۷۸۳۵	۸۸۹۵۳	۶۸/۳۵	۹۴/۵۲	۵۸۰/۱	۱۰۰۶۲/۶۸	۷۳۲	۱۶۰۳
<i>T. pubescens</i>	۹۶۹۷۶	۶۷۸۰۹	۵۹/۹۹	۷۳/۷۲	۶۱۸/۶۴	۱۰۰۸۷/۳	۸۱۸	۱۵۷۴
<i>T. persicus</i>	۸۵۰۸۴	۶۷۱۷۹	۵۸/۴۷	۷۵/۲۴	۶۸۷/۲۳	۱۱۲۰/۰۱	۹۹۹	۱۶۶۵
<i>T. lancifolius</i>	۱۲۳۶۹۸	۹۴۹۳۹	۷۳/۲۴	۹۸/۷۹	۵۹۲/۰۹	۱۰۰۴۰/۶۳	۷۶۷	۱۶۳۴

جدول ۳. تعداد رونوشت‌های جداسازی شده هر ژن از ترنسکرپتوم هر گونه

نام ژن/گونه	<i>T. daenensis</i>	<i>T. vulgaris</i>	<i>T. pubescens</i>	<i>T. persicus</i>	<i>T. lancifolius</i>
DXR	۹	۴	۷	۶	۸
DXS	۹	۱۴	۱۳	۵	۱۳
MCT	۱	۱	۱	۱	۴
CMK	۳	۲	۲	۴	۱
MDS	۱	۲	۴	۲	۱
HDS	۷	۸	۱۶	۱۲	۱۱
HDR	۱۰	۲۱	۱۳	۸	۱
IDI	۳	۳	۵	۶	۴
GGPS	۳	۲	۷	۲	۲
GDS	۱	۷	۲	۲	۲
total	۴۶	۶۴	۷۰	۹۷	۴۷

جدول ۴. طول کانتیگ هر ژن قبل و بعد ویرایش نسبت به رفرنس *Salvia miltiorrhiza* و *Mentha piperita*

نام ژن	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ	طول کانتیگ		
	<i>T. daenensis</i>	<i>T. vulgaris</i>	<i>T. pubescence</i>	<i>T. persicus</i>	<i>T. lancifolius</i>	<i>T. armeniacus</i>	<i>Z. multiflora</i>	Refrence															
<i>DXR</i>	۲۸۲۹	۱۴۰۷	۶۲۲۵	۱۴۰۷	۱۹۲۲	۱۴۰۷	۱۹۰۳	۱۴۰۷	۱۷۵۱	۱۴۰۷	۲۱۶۳	۱۴۰۷	۴۲۱۴	۱۴۰۷	۲۱۶۳	۱۴۰۷	۲۱۶۳	۱۴۰۷	۲۱۶۳	۱۴۰۷	۲۱۶۳	۱۴۰۷	
<i>DXS</i>	۲۵۰۷	۲۱۴۵	۳۳۸۰	۲۱۴۵	۲۶۱۵	۲۱۴۵	۲۷۴۰	۲۱۴۵	۲۳۹۵	۲۱۲۲	۲۴۷۹	۲۱۴۵	۲۰۶۷	۲۱۴۵	۲۱۴۵	۲۱۴۵	۲۱۴۵	۲۱۴۵	۲۱۴۵	۲۱۴۵	۲۱۴۵	۲۱۴۵	۲۱۴۵
<i>DXS2</i>	۳۳۷۳	۲۱۷۴	۳۶۵۹	۲۱۷۴	۳۲۴۲	۲۱۷۴	۱۳۵۹	۱۲۵۸	۱۹۹۸	۱۳۵۶	۲۱۹۱	۱۸۸۳	۲۶۶۹	۲۱۷۴	۲۱۷۴	۲۱۷۴	۲۱۷۴	۲۱۷۴	۲۱۷۴	۲۱۷۴	۲۱۷۴	۲۱۷۴	۲۱۷۴
<i>DXS3</i>	۱۱۷۳	۱۱۶۴	۲۳۴۴	۲۱۲۶	۱۲۷۱	۱۲۶۳	۱۱۹۲	۱۱۸۳	۱۸۵۹	۱۶۳۳	۲۶۶۱	۲۱۴۲	۲۴۸۶	۲۱۴۲	۲۱۴۲	۲۱۴۲	۲۱۴۲	۲۱۴۲	۲۱۴۲	۲۱۴۲	۲۱۴۲	۲۱۴۲	۲۱۴۲
<i>MCT</i>	۱۲۶۳	۹۱۵	۱۶۵۹	۹۱۵	۱۳۹۶	۹۱۵	۱۳۴۵	۹۱۵	۲۴۹۱	۹۱۵	۱۳۸۷	۹۱۵	۱۲۷۸	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵
<i>CMK</i>	۱۳۶۲	۱۲۱۴	۱۴۸۹	۱۲۱۴	۱۵۴۶	۱۲۱۴	۱۳۰۰	۱۲۱۴	۱۵۷۷	۱۲۱۴	۱۶۰۹	۱۲۱۴	۱۶۶۹	۱۲۱۴	۱۲۱۴	۱۲۱۴	۱۲۱۴	۱۲۱۴	۱۲۱۴	۱۲۱۴	۱۲۱۴	۱۲۱۴	۱۲۱۴
<i>MDS</i>	۹۷۲	۷۰۵	۹۶۵	۷۰۵	۱۰۲۳	۷۰۵	۹۹۷	۷۰۵	۹۴۴	۷۰۵	۱۱۷۲	۷۰۵	۱۱۹۵	۷۰۵	۷۰۵	۷۰۵	۷۰۵	۷۰۵	۷۰۵	۷۰۵	۷۰۵	۷۰۵	۷۰۵
<i>HDS</i>	۲۵۸۲	۲۲۲۹	۲۷۳۵	۲۲۲۹	۲۷۷۳	۲۲۲۹	۲۶۷۴	۲۲۲۹	۳۲۲۷	۲۲۲۹	۲۵۸۵	۲۲۲۹	۲۸۳۲	۲۲۲۹	۲۲۲۹	۲۲۲۹	۲۲۲۹	۲۲۲۹	۲۲۲۹	۲۲۲۹	۲۲۲۹	۲۲۲۹	۲۲۲۹
<i>HDR</i>	۲۱۱۶	۱۳۸۹	۱۱۸۹	۱۳۸۹	۱۴۹۶	۱۳۸۹	۱۹۶۳	۱۳۸۹	۱۷۶۳	۱۳۹۲	۱۴۸۴	۱۳۸۷	۱۹۸۳	۱۳۸۹	۱۳۸۹	۱۳۸۹	۱۳۸۹	۱۳۸۹	۱۳۸۹	۱۳۸۹	۱۳۸۹	۱۳۸۹	۱۳۸۹
<i>IDI</i>	۱۰۸۳	۸۶۲	۱۳۶۱	۹۰۶	۱۱۲۳	۸۶۲	۱۱۶۲	۸۶۲	۹۳۱	۹۱۸	۱۴۵۱	۹۰۶	۱۴۶۷	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵	۹۱۵
<i>IDI2</i>	۲۶۳۵	۸۳۳	۱۱۷۳	۸۴۶	۹۴۷	۸۴۶	۱۲۷۰	۸۴۶	۸۷۵	۷۲۸	۲۶۳۵	۸۸۹	۱۳۴۸	۸۵۸	۸۵۸	۸۵۸	۸۵۸	۸۵۸	۸۵۸	۸۵۸	۸۵۸	۸۵۸	۸۵۸
<i>GDS</i>	۱۶۶۹	۱۲۷۵	۱۷۱۱	۱۲۷۵	۱۶۴۵	۱۲۷۵	۱۶۰۸	۸۱۶	۱۳۴۰	۱۲۷۵	۱۷۹۶	۱۲۷۵	۱۸۶۰	۱۲۷۵	۱۲۷۵	۱۲۷۵	۱۲۷۵	۱۲۷۵	۱۲۷۵	۱۲۷۵	۱۲۷۵	۱۲۷۵	۱۲۷۵
<i>GDS2</i>	-	-	۱۵۶۲	۱۱۳۴	-	-	-	-	-	-	۱۳۱۷	۱۱۳۴	۱۵۲۸	۱۱۳۴	۱۱۳۴	۱۱۳۴	۱۱۳۴	۱۱۳۴	۱۱۳۴	۱۱۳۴	۱۱۳۴	۱۱۳۴	۱۱۳۴
<i>GGPS</i>	۱۱۷۷	۱۰۸۵	۱۴۸۸	۱۰۸۵	۱۳۷۲	۱۰۸۵	۱۳۷۳	۱۰۸۵	۱۳۱۱	۱۰۹۵	۱۳۸۶	۱۰۸۹	۱۶۲۱	۱۰۹۵	۱۰۹۵	۱۰۹۵	۱۰۹۵	۱۰۹۵	۱۰۹۵	۱۰۹۵	۱۰۹۵	۱۰۹۵	۱۰۹۵

جدول ۵. نوع و تعداد TPS شناسایی شده در گونه‌های مختلف آویشن

<i>T. daenensis</i>	<i>T. lancifolius</i>	<i>T. persicus</i>	<i>T. pubescens</i>	<i>T. vulgaris</i>	
۱۸	۲۷	۲۳	۳۲	۲۶	Mono
۵	۱۰	۱۷	۸	۱۵	DI
۵	۱۳	۱۵	۲۰	۱۸	SES
۳	۱۲	۱۲	۱۷	۱۷	TPSa
۱۶	۲۷	۲۲	۲۷	۲۱	TPSb
۰	۳	۳	۳	۷	TPSc
۲	۲	۱۰	۲	۲	TPSe-f
۳	۰	۱	۵	۱	TPSg

و. ۱۰ مشاهده شد. علاوه بر این، موتیف NSE / DTE در انتهای c این زیر خانواده دیده شد. مطالعات قبلی گزارش کرده‌اند که بیوستر دی‌ترین‌ها به عملکرد دو موتیف محافظت‌شده DVDD و DDYFD ارتباط زیادی دارد

ژن‌های کلروپلاستی

از نتایج حاصل از بلاست، تمامی ۸۰ ژن منحصر به فرد موجود در هر دو رفرنس (*Salvia* و *Mentha*) برای تمامی گونه‌ها، استخراج شد. اما از آنجایی که ژن‌های *psbZ*، *rps14* و *rps19*، *petL* و *petN* در تمامی گونه‌ها طولی کمتر از ۱۰۰bp داشتند، از ادامه آنالیز حذف شدند. در نتیجه، ادامه‌ی بررسی‌ها با ۷۵ ژن باقی‌مانده پیش برده شد. پس از هم‌ردیفی و حذف مناطق ضعیف در هم‌ردیفی‌ها، دو ژن *ndhA* و *ycf1* نیز به علت کاهش طول‌هایشان به کمتر از ۱۰۰bp، از مسیر مطالعه حذف شدند. در نهایت، با ادغام توالی‌های ۷۳ ژن باقی‌مانده، مجموعاً ۴۰،۰۷۰bp اطلاعات از ژنوم کلروپلاستی هر گونه استخراج شد. اطلاعات مربوط به طول نهایی تمامی ژن‌ها در جدول ۶ ارائه شده‌است. رنج طولی این ۷۳ ژن، از ۱۰۰bp برای ژن psbT تا ۲۲۵۱bp برای ژن psbA گزارش شد.

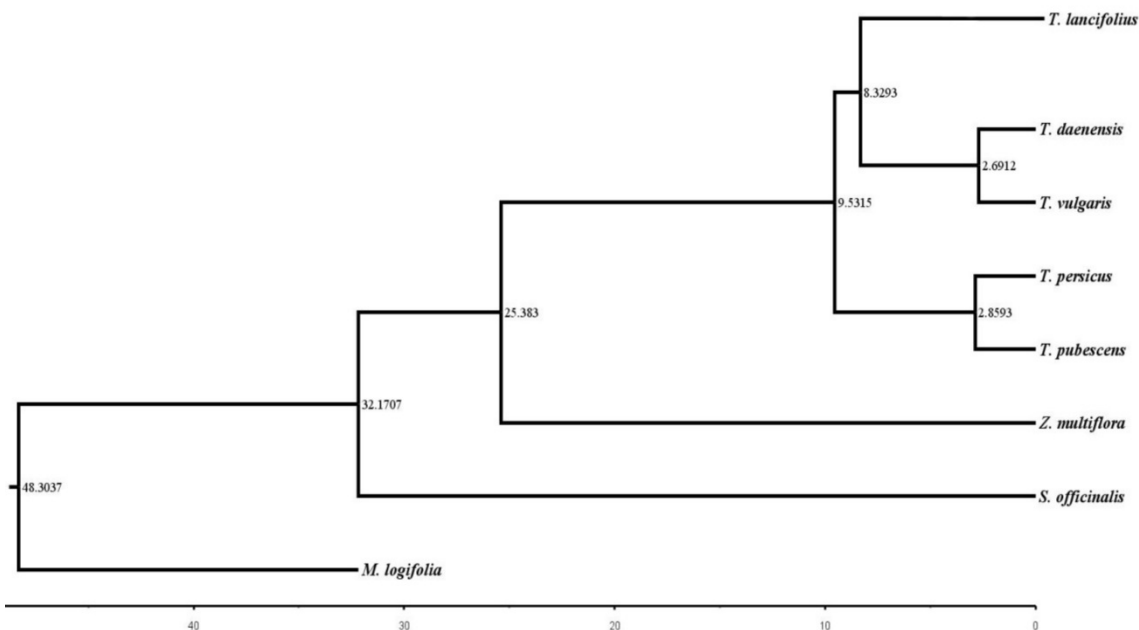
بر اساس مطالعات انجام‌شده ترپن‌سینتازها دارای درجه بالایی از یکسانی و شباهت، موتیف‌ها و دمین‌های حفاظت‌شده هستند که به منظور شناسایی این خانواده آنزیمی تمامی رونوشت‌های مربوط به ترپن‌سینتاز در همه‌ی گونه‌ها نسبت به وجود این موتیف‌ها بررسی شد. اعضای زیر خانواده TPSb حاوی موتیف-های محافظت‌شده DDxxD، RxR و RRx8W با مکان‌های کاملاً محافظت‌شده از آرژنین، سیستئین و هیستیدین بودند. در مطالعه حاضر، موتیف DDxxD تقریباً توسط همه اعضای زیر خانواده TPSb (به عنوان دسته اصلی در بیوستر مونوترپن‌ها) به اشتراک گذاشته شد که به صورت (Y / I / F / A / L) (V / I / F / A / L) نشان داده شد. DDxxD به استثنای زیر خانواده TPS-c، تقریباً در همه زیر خانواده‌های دیگر نیز مشترک است. موتیف محافظت‌شده RDR (RxR) و GTLXEL به ترتیب در بالادست و پایین‌دست موتیف DDxxD قرار داشتند. موتیف RRx8W نیز تقریباً در تمام اعضای زیر خانواده TPSb متنوع بود و بیشتر به صورت (K / R / E / Q) (P / A) RRSNGY (S / T) (L / Y) W مشاهده شد. این موتیف، بیشترین تفاوت بین گونه‌های آویشن و *Z. multiflora* در سایت‌های شماره ۷، ۹

جدول ۶. طول ژن‌های کلروپلاستی در گونه‌های آویشن

نام ژن	طول نهایی قطعات ژنی استخراج شده (bp)	نام ژن	طول نهایی قطعات ژنی استخراج شده (bp)	نام ژن	طول نهایی قطعات ژنی استخراج شده (bp)	نام ژن	طول نهایی قطعات ژنی استخراج شده (bp)
<i>petA</i>	۵۶۳	<i>psaA</i>	۲۲۵۱	<i>accD</i>	۱۱۳۵	<i>ndhK</i>	۶۷۶
<i>petB</i>	۶۴۶	<i>psaB</i>	۲۱۱۷	<i>atpA</i>	۹۴۳	<i>ndhI</i>	۵۰۵
<i>petD</i>	۴۸۱	<i>psaC</i>	۲۴۴	<i>atpB</i>	۱۲۷۲	<i>ndhJ</i>	۴۷۴
<i>petG</i>	۱۱۲	<i>psaI</i>	۱۰۹	<i>atpE</i>	۴۰۰	<i>rpl14</i>	۲۸۷
<i>ycf15</i>	۲۴۷	<i>psaJ</i>	۱۰۷	<i>atpF</i>	۵۵۳	<i>rpl16</i>	۳۹۶
<i>ycf2</i>	۳۱۴	<i>psbA</i>	۱۰۵۷	<i>atpH</i>	۲۴۴	<i>rpl2</i>	۷۸۵
<i>ycf3</i>	۲۸۰	<i>psbB</i>	۸۵۵	<i>atpI</i>	۷۴۲	<i>rpl20</i>	۳۷۹
<i>ycf4</i>	۵۵۲	<i>psbC</i>	۱۲۲۵	<i>ccsA</i>	۷۶۸	<i>rpl22</i>	۴۶۶
<i>rps11</i>	۴۱۵	<i>psbD</i>	۱۰۶۰	<i>cemA</i>	۶۸۳	<i>rpl23</i>	۲۸۰
<i>rps12</i>	۲۳۰	<i>psbE</i>	۲۵۰	<i>clpP</i>	۵۱۸	<i>rpl32</i>	۱۶۰
<i>rps15</i>	۲۷۱	<i>psbF</i>	۱۱۸	<i>infA</i>	۲۳۲	<i>rpl33</i>	۱۹۹
<i>rps16</i>	۱۹۵	<i>psbH</i>	۱۳۹	<i>matK</i>	۱۵۱۵	<i>rpl36</i>	۱۱۲
<i>rps18</i>	۳۰۴	<i>psbI</i>	۱۰۹	<i>ndhB</i>	۷۷۵	<i>rpoA</i>	۶۹۶
<i>rps2</i>	۴۶۹	<i>psbJ</i>	۱۲۱	<i>ndhC</i>	۳۶۱	<i>rpoB</i>	۶۵۰
<i>rps3</i>	۳۵۷	<i>psbK</i>	۱۷۸	<i>ndhD</i>	۱۳۰۷	<i>rpoC1</i>	۸۶۳
<i>rps4</i>	۶۰۴	<i>psbL</i>	۱۱۵	<i>ndhE</i>	۳۰۴	<i>rpoC2</i>	۵۵۷
<i>rps7</i>	۴۵۹	<i>psbM</i>	۱۰۳	<i>ndhF</i>	۸۷۳		
<i>rps8</i>	۲۴۳	<i>psbN</i>	۱۳۰	<i>ndhG</i>	۵۲۰		
<i>rbcL</i>	۱۴۳۲	<i>psbT</i>	۱۰۰	<i>ndhH</i>	۸۷۸		

اکثر توده‌های *T. daenensis* در یک گروه دسته‌بندی شدند، که با نتایج بررسی‌های ما مطابقت دارد. در حالی که توده‌های مختلف گونه‌های *T. fedtschenkoi*، *T. pubescens*، *T. lancifolius* و *T. transcaspicus* در بین گروه‌ها و کلادهای مختلف توزیع شدند. در این مطالعه چنین عنوان شده است که این گونه‌ها توانایی بالایی در پذیرش مواد ژنتیکی از گونه‌های مختلف دارند و احتمال جریان ژنی در این گونه‌ها بسیار بالا است (Mojiri et al., 2014). در مطالعه‌ی دیگر به منظور بررسی روابط فیلوژنتیکی گونه‌های آویشن با استفاده از توالی‌های ITS، تمام گونه‌های *T. daenensis*، *T. pubescens*، *T. persicus* و *T. transcaspicus* در یک کلاد مشترک قرار گرفتند، اما *T. vulgaris* به صورت گروهی جداگانه قرار گرفت. علت این جدایی را وجود یک نوکلئوتید G اضافه در موتیف محافظت شده-ی '5'AAGGAA3' در توالی ITS1 می‌دانند که سبب تفاوت در ساختار ثانویه ITS در این گونه و در نتیجه ایجاد یک کلاد جداگانه در درخت فیلوژنتیکی شده‌است (Abdolahinia et al., 2011). نتایج حاصل از پژوهش که بر روی توالی‌های ITS انجام شده بود نیز گزارشی مشابه ارائه داد (Sonboli et al., 2013).

بررسی روابط فیلوژنتیکی با استفاده از اطلاعات کلروپلاستی درخت حاصل از ابزار Beast گونه‌های جنس آویشن را به دو کلاد تقسیم‌بندی نمود. در این درخت روابط بسیار نزدیک *T. daenensis*، *T. vulgaris* (کلاد ۱) و دو گونه‌ی *T. persicus* و *T. pubescens* (کلاد ۲) به وضوح نشان داده شد (شکل ۱). علاوه بر این، گونه‌های *Z. multiflora*، *S. officinalis* و *M. logifolia* خارج از گروه گونه‌های جنس *Thymus* قرار گرفتند و از این بین، جنس *Zataria* رابطه‌ی نزدیکی را با جنس *Thymus* به نمایش گذاشت. در مطالعه‌ای که بر روی تنوع ژنتیکی و پلی‌مورفیسم شیمیایی گونه‌های آویشن انجام شده است، رابطه‌ی بسیار نزدیک گونه‌های *T. daenensis* و *T. vulgaris* از نظر مورفولوژیکی و محتوای اسانس نیز گزارش شده، به این شکل که گونه‌ی *T. daenensis* حاوی درصد بالایی از تیمول و گونه‌ی *T. vulgaris* حاوی درصد بالایی ترکیبی از تیمول و گاما ترپینین بود (Rustaeie et al., 2013). در مطالعه‌ی دیگری که تنوع ژنتیکی توده‌های آویشن ایران با استفاده از نشانگر ISJ^۱ مورد بررسی قرار گرفت نیز، چنین گزارش شد که تمام توده‌های *T. vulgaris* و



شکل ۱. درخت فیلوژنتیکی حاصل از اطلاعات کلروپلاستی گونه‌های مختلف آویشن

میلیون سال قبل رخ داده‌است. این در صورتی است که گونه‌های جنس آویشن، ۹/۵۳ میلیون سال پیش از یکدیگر انشقاق پیدا کرده‌اند. زمان انشقاق گونه‌های آویشن در زیر کلادها حدود ۲-۳ میلیون سال پیش تخمین زده شد. در مطالعه‌ای که به بررسی روابط فیلوژنتیکی و تخمین زمان‌های انشقاق قبیله‌ی Menthaeae بر اساس داده‌های کلروپلاستی (*ycf1-rps15-ycf1* و *trnL-F* و *rpl32-trnL*) و DNA ریبوزوم هسته‌ای (ITS و ETS) انجام شده چنین گزارش کردند که جنس آویشن حدود ۲/۵ میلیون سال پیش انشقاق یافته‌است (Drew & Sytsma, 2012). در نتایج به‌دست آمده از درخت بیزین حاصل از مجموعه داده‌های کلروپلاستی مطالعه‌ی حاضر نشان داده شد که زمان واگرایی گونه‌های در هر کلاد حدود ۲-۳ میلیون سال پیش است و این خود کلادها هستند که فاصله‌ی زیادی از یکدیگر پیدا کرده‌اند (حدود ۸ تا ۹ میلیون سال پیش). در مطالعه‌ی گذشته‌ی زمان انشقاق گونه‌ی *Z. multiflora*، ۱۷/۹ میلیون سال پیش تخمین زده‌شد، که با نتایج داده‌های کلروپلاستی حاصل از مطالعه‌ی حاضر (۲۵/۳۸ میلیون سال پیش) متفاوت بود (Drew & Sytsma, 2012). با توجه به نتایج به دست آمده و فاصله‌ی زمانی تعریف شده برای زمان انشقاق جنس *Zataria* و *Thymus* این فرضیه که *Zataria* بتواند در یک گروه مشابه با جنس *Thymus* قرار بگیرد کاملاً رد شد. این طور به نظر می‌رسد که جنس *Zataria* بتواند به عنوان یکی از اجداد جنس *Thymus*

در درخت رسم شده، گونه‌ی *Z. multiflora* خارج از گروه جنس آویشن قرار گرفت. این در حالی است که در مطالعات گذشته از *Zataria* (آویشن شیرازی) به عنوان گیاهی مشابه جنس *Thymus* یاد شده‌است، این شباهت که در ظاهر و عطر و بوی جنس‌های *Thymus* و *Zataria* وجود دارد همیشه باعث اشتباه در شناسایی و تشخیص این دو جنس از یکدیگر شده‌است (Izadiyan et al., 2018). لذا چنین به نظر می‌رسد که طبقه‌بندی جنس *Zataria* مشکل باشد. در این راستا نتایج حاصل از این مطالعه چنین نشان داد که جنس *Zataria* در کلادی جداگانه از گونه‌های جنس *Thymus* قرار می‌گیرد. با انجام مطالعات بر روی توالی‌های ITS و *trnK/trnL-F* و بررسی روابط فیلوژنتیکی گونه‌های جنس *Thymus* و *Zataria* چنین گزارش شده است که جنس *Zataria* در گروهی جداگانه اما نزدیک به کلاد گونه‌های آویشن واقع می‌شود (Brauchler et al., 2010). در مطالعه‌ی دیگری که در جهت شناسایی و تأیید مولکولی گونه‌ی *T. persicus* انجام گرفت نیز، گونه‌ی *Z. multiflora* از گونه‌های آویشن جدا شد (Sonboli et al., 2013).

در درخت بیزین حاصل از داده‌های زمان تخمین زده شده برای جنس *Mentha*، ۴۸/۳ میلیون سال پیش و برای جنس *Salvia*، ۳۲/۱۷ میلیون سال پیش بود. نتایج این مطالعه نشان داد که جدایی جنس *Zataria* از گونه‌های جنس آویشن، ۲۵/۳۸

تجزیه و تحلیل چندین ژنوم گیاهی نشان داده است ژن‌های TPS یک خانواده آنزیمی متوسط است، اگرچه همه آن‌ها فعالیت ندارند. بر اساس نتایج این شناسایی، ترپین‌ها در سه خانواده مونو، دی و سزکوئی‌ترپین‌ها و سه کلاس اصلی و پنج زیر کلاس TPSa، TPSb، TPSc، TPSe-f و TPS-g تقسیم شدند. همانگونه که انتظار می‌رفت بیشترین کانتینگ‌ها مربوط به مونوترپین‌ها و کلاس a و کمترین کانتینگ در دی‌ترپین‌ها دیده شد. همچنین با توجه به تحقیقات گذشته که کلاس d را فقط در بازدانگان اعلام کردند در این پژوهش نیز این کلاس دیده نشد که تایید کننده تحقیقات پیشین می‌باشد.

در ادامه، توالی‌های ۷۳ ژن کلروپلاستی (بر اساس ژن‌های کلروپلاستی *Mentha logifolia* و *Salvia officinalis*) استخراج شدند. از اطلاعات کلروپلاستی بدست آمده، در مجموع ۴۰۰۷۰ bp در بررسی روابط فیلوژنتیکی مورد استفاده قرار گرفت. از بررسی درخت فیلوژنتیک مشخص گردید موقعیت و رابطه سه جنس *Thymus*، *Salvia*، *Zataria* و *Mentha* نسبت به جنس *Thymus* ثابت است، و *Zataria* نیز به‌عنوان نزدیک‌ترین جنس به آویشن، در کنار گروه گونه‌های آویشن و خارج از آن قرار گرفت. براساس درختان ترسیم شده در این پژوهش نیز دو گونه‌ی *T. daenensis* و *T. vulgaris* رابطه‌ی خواهری بسیار نزدیکی را نشان دادند که با تحقیقات اساسی گذشته مطابقت داشت.

سپاسگزاری

پژوهش حاضر، با حمایت مالی صندوق حمایت از پژوهشگران و فناوران کشور (INSF) طی اعتبار پژوهشی شماره ۹۶۰۱۴۵۰۳ انجام شده است که مراتب سپاس و قدردانی به عمل می‌آید.

References

Abdolahinia, E. D., Bashir, N. S., Hagnazari, A., Nazemiyeh, H., & Hejazi, M. S. (2011). A comparative phenotypic and ITS based genotypic study in thyme species (*Thymus* L. Lamiaceae). *Vegetos*, 24(2), 102-113.

Andrews, S., Krueger, F., Seconds-Pichon, A., Biggins, F., & Wingett, S. (2015). FastQC. A quality control tool for high throughput sequence data. Babraham Bioinformatics. Babraham Inst.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114-2120.

Bräuchler, C., Meimberg, H., & Heubl, G. (2010). Molecular phylogeny of Menthinae (Lamiaceae, Nepetoideae, Mentheae)—taxonomy,

در مطالعات آینده به عنوان خارج گروه، در رسم درختان ریش‌دار در نظر گرفته‌شود.

نتیجه‌گیری

با توجه به اهمیت گیاه دارویی آویشن، در این مطالعه سعی شد تا با ایجاد اطلاعات ترنسکریپتومی بتوان بخش قابل توجهی از ژن-ها و مسیرهای کلیدی سنتز متابولیت‌های ثانویه در چند گونه مختلف شناسایی و مورد مقایسه قرار گیرد. بدین منظور اطلاعات ترنسکریپتومی گونه مختلف آویشن با استفاده از پلتفرم Illumina حاصل شد. بعد از سرهم‌بندی خوانش‌ها با استفاده از ابزارهای مختلف، بهترین نتیجه برای مراحل بعدی آنالیز مورد استفاده قرار گرفت. تجزیه و تحلیل بیوانفورماتیکی توالی‌های به دست آمده منجر به شناسایی به ترتیب ۶۶۵۹۱، ۸۸۹۵۳، ۶۷۱۷۹ و ۹۴۹۳۹ ژن منفرد در گونه‌های *T. persicus*، *T. lancifolius* و *T. pubescens*، *T. vulgaris* و *T. daenensis* شد. به علت فقدان اطلاعات در آویشن نزدیک به ۴۰٪ از خوانش‌ها هیچ جورشدگی با پایگاه‌های اطلاعاتی را نشان ندادند.

ژن‌های کلیدی در مسیر بیوسنتزی MEP در همه‌ی گونه‌ها شناسایی شدند. برخی از ژن‌ها مانند IDI، DXS و GDS و TPS چندین آیزوفرم در آن‌ها مشخص شد. تقریباً طول کامل اکثر ژن‌ها به دست آمد. که در این بین بیشترین کانتینگ مربوط ژن‌های DXS، HDS، HDR و TPS و کمترین کانتینگ ژن‌های MCT، CMK، GDS و MDS را داشتند. می‌توان نتیجه گرفت ژن‌ها با تعداد کانتینگ بیشتر میزان بیان و فراوانی بیشتری دارند. TPS در سلسه‌ی گیاهی بسیار متنوع است و بسیاری از آن‌ها در بافت‌های خاصی بیان نشان می‌دهند.

biogeography and conflicts. *Molecular phylogenetics and evolution*, 55(2), 501-523.

Bryant, D. M., Johnson, K., DiTommaso, T., Tickle, T., Couger, M. B., Payzin-Dogru, D., ... & Whited, J. L. (2017). A tissue-mapped axolotl de novo transcriptome enables identification of limb regeneration factors. *Cell reports*, 18(3), 762-776.

Bunsawat, J., Elliott, N.E., Hertweck, K.L., Sproles, E. and Alice, L.A. (2004). Phylogenetics of *Mentha* (Lamiaceae): evidence from chloroplast DNA sequences. *Systematic botany*, 29(4), 959-964.

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, 25(15), 1972-1973.

- Drew, B. T., & Sytsma, K. J. (2012). Phylogenetics, biogeography, and staminal evolution in the tribe Menthaeae (Lamiaceae). *American journal of botany*, 99(5), 933-953.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*, 32(5), 1792-1797.
- Federici, S., Galimberti, A., Bartolucci, F., Bruni, I., De Mattia, F., Cortis, P. & Labra, M. (2013). DNA barcoding to analyse taxonomically complex groups in plants: the case of *Thymus* (Lamiaceae). *Botanical Journal of the Linnean Society*, 171(4), 687-699.
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., & Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, 29(7), 644-652.
- Ghobadi, S., Marouf, A. & Majd, M. (2016). Differential expression of the key genes involved in the biosynthesis of monoterpenes in different tissues and in response to abiotic elicitors in Summer savory (*Satureja hortensis*). *Cell and Tissue Journal*, 7, 3.
- Habibi, S., Ehteshamnia, A., Fatehi, F. & Ghaderi, A. (2020). Investigation of γ -Terpinene synthesis and α - Terpinene synthesis reductoisomerase genes expression and its relation to monoterpene carvacrol biosynthesis in *Thymus vulgaris* cv. 'Varico 3'. *Genetic Engineering and Biosafety Journal*, 9(2), 161-169
- Haji Akhondi, A. & Farahani Kia, B. (1386). What do you know about thyme?. *Encyclopedia of Medicinal Plants*, 12: 12-9. (in persian)
- Izadiyan, P., Hemmateenejad, B., Mohsen Taghavi, S., & Izadiyan, M. (2018). Discrimination of Shirazi thyme from thymus species and antioxidant activity prediction using chemometrics and FT-IR spectroscopy. *Journal of the Iranian Chemical Society*, 15, 259-268.
- Jamzad, Z. (1373). *Thyme*. Tehran: Research Institute of Forests and Pastures of the country. (in persian)
- Jamzad, Z. (1388). *Iranian herbs and spices*. Tehran: Research Institute of Forests and Pastures of the country. (in persian)
- Jiang, S. Y., Jin, J., Sarojam, R., & Ramachandran, S. (2019). A comprehensive survey on the terpene synthase gene family provides new insight into its evolutionary patterns. *Genome biology and evolution*, 11(8), 2078-2098.
- Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: a resource for timelines, timetrees, and divergence times. *Molecular biology and evolution*, 34(7), 1812-1819.
- Liu, J., Li, G., Chang, Z., Yu, T., Liu, B., McMullen, R., ... & Huang, X. (2016). BinPacker: packing-based de novo transcriptome assembly from RNA-seq data. *PLoS computational biology*, 12(2), e1004772.
- Mashhady Malekzadeh, A Majdi, M. & Maroufi, A. (2017). Expression analysis of biosynthetic genes of thymol and carvacrol in different tissues of thyme (*Thymus vulgaris*). *Iranian Journal of Rangelands and Forests Plant Breeding and Genetic Research*, 25(1), 160-171.
- Mojiri, F., Esmacili, A., Nazarian Firouz Abadi, F., Maddahi Arefi, H., & Ahmadi, H. (2014). Study of genetic diversity of Iranian *Thymus* accessions, using ISJ semi-random markers. *Agricultural Biotechnology Journal*, 6(2), 147-162. (in persian)
- Omidbeigi, R. (1393). *Production and processing of medicinal plants*. Mashhad: Astan Quds Razavi. (in persian)
- Priya, P., Yadav, A., Chand, J., & Yadav, G. (2018). Terzyme: a tool for identification and analysis of the plant terpenome. *Plant Methods*, 14, 1-18.
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841-842.
- Rambaut, A. & Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol*, 10, e1003537
- Rambaut, A., & Drummond, A. J. (2018). FigTree v1.4.0 in Computer Program and Documentation Distributed by the Author. 2012. Available online: <http://tree.bio.ed.ac.uk/software/figtree/>.
- Rana, S. B., Zadlock IV, F. J., Zhang, Z., Murphy, W. R., & Bentivegna, C. S. (2016). Comparison of de novo transcriptome assemblers and k-mer strategies using the killifish, *Fundulus heteroclitus*. *PloS one*, 11(4), e0153104.
- Rustaiee, A. R., Yavari, A., Nazeri, V., Shokrpour, M., Sefidkon, F., & Rasouli, M. (2013). Genetic diversity and chemical polymorphism of some *Thymus* species. *Chemistry & Biodiversity*, 10(6), 1088-1098.
- Rustaiee, A., Sefidkon, F., Tabatabaei, S. M. F., Omidbaigi, R., & Mirahmadi, S. F. (2011). Chemical polymorphism of essential oils from five populations of *Thymus daenensis* Celak. subsp. *daenensis* endemic to Iran. *Journal of Essential Oil Research*, 23(3), 6-11.
- Salzberg, S. L., Phillippy, A. M., Zimin, A., Puiu, D., Magoc, T., Koren, S., & Yorke, J. A. (2012). GAGE: A critical evaluation of genome assemblies and assembly algorithms. *Genome research*, 22(3), 557-567.

- Schliesky, S., Gowik, U., Weber, A. P., & Bräutigam, A. (2012). RNA-seq assembly—are we there yet?. *Frontiers in plant science*, 3, 220.
- Smith-Unna, R., Bournnell, C., Patro, R., Hibberd, J. M., & Kelly, S. (2016). TransRate: reference-free quality assessment of de novo transcriptome assemblies. *Genome research*, 26(8), 1134-1144.
- Smolik, M., Jadczyk, D. and Korzeniewska, S. (2009). Assessment of morphological and genetic variability in some *Thymus* accessions using molecular markers. *Not. Bot. Hort. Agrobo.* 37: 234-240.
- Sonboli, A., Mirjalili, M. H., Bakhtiar, Z. I. B. A., & Jamzad, Z. I. B. A. (2013). Molecular authentication of *Thymus persicus* based on nrDNA ITS sequences data. *Iran J. Bot*, 19, 179-185.
- Thompson, J. D., Manicacci, D., & Tarayre, M. (1998). Thirty-five years of thyme: a tale of two polymorphisms. *BioScience*, 48(10), 805-815.