

Identification of Microsatellite Molecular Markers in Saffron (*Crocus Sativus* L.) With RNA-Seq Data

Ali Darvishian¹, Farhad Nazarian-Firouzabadi^{2*}, Mostafa Darvishnia³, and Ahamd Ismaili⁴

Submitted: 27 January 2022

Revised: 11 June 2022

Accepted: 5 July 2022

Available Online: 6 July 2022

How to cite this article:

Darvishian, A., Nazarian-Firouzabadi, F., Darvishnia, M., and Ismaili, A. 2022. Identification of Microsatellite Molecular Markers in Saffron (*Crocus Sativus* L.) With RNA-Seq Data. Saffron Agronomy & Technology, 10(2): 149-161.
DOI: 10.22048/jsat.2022.326684.1452

Abstract

Saffron (*Crocus sativus* L.) is the most valuable spice in the world. Due to the lack of genomic information, saffron breeding has encountered many problems. To this end, generating and collecting genetic data using Next Generation Sequencing (NGS) techniques is crucial for saffron traditional and molecular breeding programs. Molecular markers, especially powerful co-dominance markers such as simple sequence repeats (SSRs) markers play an important role in breeding projects. In the present study, for the first time, SSR markers related to genes associated with *F. oxysporum* disease of corms were identified following an RNA-Seq-based transcriptomic approach. To this end, total RNA was extracted from mature corms, and RNA-Seq was performed based on the Novaseq 6000 platform. The De-novo assembly was performed with Trinity software, and the MISA search tool was used to identify SSRs. Based on the results of this study, 357028 transcripts were identified. A total of 70060 transcripts were identified to contain SSR sequences. BLAST algorithm analysis revealed that the highest similarity between saffron SSRs were found with rice and Arabidopsis. Among the identified uni-genes, 18846, 23988, and 10969 genes were identified in UNIPROT, Nr, and GO databases, respectively, in which 10375 uni-genes were common to all databases. Due to the high priority of saffron in Iran as a strategic crop plant, any genetic information, including mining SSRs, is of great importance in studying genetic diversity, constructing genetic maps, linkage, and QTLs analysis in saffron future breeding programs.

Keywords: Transcriptome Analysis, Sequencing, Codominant Markers.

1 - Ph.D candidate, Production Engineering and Plant Genetics Department, Faculty of Agriculture and Natural Resources, Lorestan University, Khorramabad, Iran

2 - Professor, Production Engineering and Plant Genetics Department, Faculty of Agriculture and Natural Resources, Lorestan University, Khorramabad, Iran

3 - Associate Professor, Plant Protection Department, Faculty of Agriculture and Natural Resources, Lorestan University Khorramabad, Iran

4- Production Engineering and Plant Genetics Department, Faculty of Agriculture and Natural Resources, Lorestan University, Khorramabad, Iran

(*-Corresponding author: Nazarian.f@lu.ac.ir)



مقاله پژوهشی

شناسایی نشانگرهای مولکولی ریزماهواره در گیاه زعفران (*Crocus sativus* L.) با استفاده از داده‌های RNA-Seq

علی درویشیان^۱، فرهاد نظریان فیروزآبادی^{۲*}، مصطفی درویش‌نیا^۳ و احمد اسماعیلی^۴

تاریخ دریافت: ۷ بهمن ۱۴۰۰

تاریخ بازنگری: ۲۱ خرداد ۱۴۰۱

تاریخ پذیرش: ۱۴ تیر ۱۴۰۱

درویشیان، ع.، نظریان فیروزآبادی، ف.، درویش‌نیا، م.، و اسماعیلی، ا. ۱۴۰۱. شناسایی نشانگرهای مولکولی ریزماهواره در گیاه زعفران (*Crocus sativus* L.) با استفاده از داده‌های RNA-Seq. زراعت و فناوری زعفران، ۱۰(۲): ۱۶۱-۱۴۹.

چکیده

زعفران ارزشمندترین گیاه ادویه‌ای و دارویی جهان است. به‌نژادی زعفران به دلیل اطلاعات اندک از ساختار ژنوم آن با مشکلات زیادی همراه است. از این‌رو، دستیابی به اطلاعات ژنتیکی این گیاه از اهمیت بسیاری برای برنامه‌های به‌نژادی سنتی و مولکولی آن برخوردار می‌باشد. نشانگرهای مولکولی به‌ویژه نشانگرهای قدرتمند هم‌بازی همانند نشانگرهای ریزماهواره (SSRs) جایگاه مهمی در برنامه‌های به‌نژادی دارند. در پژوهش حاضر برای اولین بار نشانگرهای ریزماهواره‌ی SSR بنه گیاه زعفران از ترانسکریپتوم بنه تحت تأثیر بیماری قارچی *Fusarium oxysporum* استخراج شدند. در این راستا پس از استخراج RNA، توالی‌یابی بر اساس چارچوب فنی Novaseq6000 انجام شد. یکپارچه‌سازی نوپدید با بهره‌گیری از نرم‌افزار Trinity صورت گرفت و برای شناسایی SSRها از ابزار جستجوی MISA استفاده شد. بر اساس نتایج این مطالعه، از ۳۵۷۰۲۸ رونوشت شناسایی شده، تعداد ۷۰۰۶۰ رونوشت دربرگیرنده‌ی توالی SSR بودند. همچنین از مجموع ۸۹۸۸۸ SSR شناسایی شده، تکرارهای ۱ و ۲ نوکلئوتیدی بیشترین فراوانی (۵۰٪ و ۲۶٪) را به خود اختصاص دادند. بر اساس نتایج حاصل از الگوریتم BLAST، بیشترین میزان شباهت رونوشت‌های حاوی SSRها با برنج و آراییدوپسیس به دست آمد. در مجموع از میان یونی‌ژن‌ها، تعداد ۱۸۸۴۶، ۲۳۹۸۸ و ۱۰۹۶۹ به ترتیب در پایگاه‌های داده‌ی UNIPROT، Nr و GO شناسایی شدند که تعداد ۱۰۳۷۵ یونی‌ژن در بین همه‌ی پایگاه‌ها مشترک بود. به دلیل اهمیت این گیاه در ایران به‌عنوان یک محصول راهبردی، توسعه نشانگرهای هم‌باز با چند شکلی بالا مانند نشانگرهای SSR، برای بررسی تنوع ژنتیکی، ساخت نقشه‌های ژنتیکی، تجزیه و تحلیل پیوستگی و نقشه‌یابی QTLs در برنامه‌های به‌نژادی این گیاه مهم خواهد بود.

کلمات کلیدی: تجزیه و تحلیل ترانسکریپتوم، توالی‌یابی، نشانگرهای هم‌باز

۱- دانشجوی دکتری اصلاح نباتات، ژنتیک مولکولی و مهندسی ژنتیک، گروه مهندسی تولید و ژنتیک گیاهی، دانشکده کشاورزی و منابع طبیعی، دانشگاه لرستان، خرم‌آباد، ایران
۲- فرهاد نظریان فیروزآبادی، استاد گروه مهندسی تولید و ژنتیک گیاهی، دانشکده کشاورزی و منابع طبیعی، دانشگاه لرستان، خرم‌آباد، ایران
۳- دانشیار گروه گیاهپزشکی، دانشکده کشاورزی و منابع طبیعی، دانشگاه لرستان، خرم‌آباد، ایران
۴- استاد گروه مهندسی تولید و ژنتیک گیاهی، دانشکده کشاورزی و منابع طبیعی، دانشگاه لرستان، خرم‌آباد، ایران
(#-نویسنده مسئول: Nazarian.f@lu.ac.ir)

مقدمه

توالی‌های تکراری ساده (SSRs: Simple Sequence Repeats) یا ریزماهواره‌ها (Microsatellites) نشانگرهای مولکولی با توالی‌های تکراری ۱ تا ۶ نوکلئوتیدی هستند که در ژنوم یوکاریوت‌ها پراکنش دارند (Zietkiewicz et al., 1994). ویژگی‌هایی چون تکرار پذیری زیاد، فراوانی و پراکنش ژنومی بالا، وراثت هم‌بارز، تشخیص ساده و کم هزینه بودن سبب استفاده گسترده و اعتماد بیشتر به نشانگرهای SSR نسبت به سایر نشانگرهای مولکولی شده است (Thiel et al., 2003). با این روش امکان شناسایی نشانگرهای مولکولی براساس توالی کامل ژنوم و همچنین تجزیه و تحلیل ترانسکریپتوم وجود دارد. با این حال، شناسایی به واسطه‌ی توالی‌یابی ترانسکریپتوم قادر است نشانگرهایی را شناسایی کند که با ژن‌های کارکردی مرتبط هستند (Hiremath et al., 2011). مزیت دیگر شناسایی نشانگرهای مولکولی و به‌ویژه نشانگرهای SSR بر اساس تجزیه و تحلیل ترانسکریپتوم، قابلیت وراثت پذیری بیشتر SSRهای حاصل برای سایر گونه‌ها است. در واقع توالی‌یابی ترانسکریپتوم قادر است با شناخت مسیرهای متابولیکی گیاهان ارزشمند دارویی-اقتصادی و نیز شناسایی نشانگرهای مولکولی، به افزایش اطلاعات پژوهش‌گران مشغول در به‌نژادی مولکولی کمک نماید (Li et al., 2002; Nemati et al., 2012; Chen et al., 2017). جنس *Crocus* بیش از ۸۰ گونه چندساله را شامل می‌شود که تا کنون ۹ گونه از این جنس در ایران گزارش شده است (Alavi-Kia et al., 2008). در این بین، گونه‌ی زراعی آن با نام علمی *Crocus sativus* L. با ویژگی‌های ظاهری کللاه‌ی بلند، قرمز و معطر از سایر گونه‌های این جنس متمایز است که به‌عنوان گران‌ترین محصول کشاورزی و دارویی در جهان شناخته می‌شود (Koocheki et

al., 2011; Alipoor Miandehi et al., 2014; Babaei et al., 2014). به دلایلی مانند: نیاز آبی کم، سهولت در عملیات زراعی، عملکرد بالای مزارع زعفران و مصرف گسترده در صنایع غذایی، پزشکی و نساجی، کشت این گیاه از صرفه‌ی اقتصادی بالایی برخوردار می‌باشد (Festuccia et al., 2014). بررسی‌های کاربوتیپی انجام شده بر روی زعفران زراعی نشان می‌دهد که زعفران در سلول‌های سوماتیک خود $2n=3x=24$ کروموزوم دارد، از این‌رو زعفران زراعی یک اتوتتریپلوئیدی عقیم بوده و تکثیر آن صرفاً از طریق بنه انجام می‌گیرد (Nemati et al., 2019).

از نشانگرهای SSR به‌طور گسترده در تعیین فاصله یا قرابت ژنتیکی بین ژنوتیپ‌ها (Nasiri et al., 2019)، مکان‌یابی ژن‌های کنترل‌کننده صفات کمی، تهیه نقشه ژنومی و تجزیه پلی ژنیک صفات مختلف کمی در گیاهان زراعی استفاده می‌شود (Millan et al., 2006). از جمله‌ی راهکارهایی که می‌تواند به افزایش اطلاعات ژنومی و ترانسکریپتومی گیاهان غیر مدلی هم‌چون زعفران کمک کند، روش‌های توالی‌یابی نسل بعد (NGS) همانند فناوری RNA-Seq است. در توالی‌یابی RNA توالی‌یابی ناحیه‌های کدکننده صورت می‌گیرد و هم‌زمان اطلاعات مربوط به توالی نشانگرها نیز گزارش می‌شود. درواقع یکی از مهم‌ترین دستاوردهای استفاده از تجزیه و تحلیل ترانسکریپتوم بر اساس فناوری RNA-Seq، شناسایی نشانگرهای مولکولی است (Strickler et al., 2012). برای مثال، مقایسه‌ی ترانسکریپتوم زعفران در مرحله‌ی گل‌دهی و قبل از گل‌دهی در سایت MISA منجر به شناسایی تعداد ۷۹۰۲۸ نشانگر SSR یک تا شش نوکلئوتیدی گردید که در این بین SSRهای تک نوکلئوتیدی ۷۱/۲ درصد کل SSRهای شناسایی شده را به خود اختصاص دادند. در مرحله بعد،

مطالعه‌ی ریزماهوره‌های SSR با استفاده از داده‌های حاصل از توالی‌یابی RNA می‌پردازد تا نشانگرهای SSR مرتبط با ژن‌های تحت تأثیر پاتوژن قارچی *F. oxysporum* را شناسایی کند. بدیهی است در مرحله بعد می‌توان از این نشانگرهای مولکولی برای شناسایی QTLs مرتبط با مقاومت یا حساسیت به بیماری قارچی *F. oxysporum* در جمعیت‌های نقشه‌یابی استفاده نمود.

مواد و روش‌ها

آماده‌سازی نمونه‌ی گیاهی و اعمال تیمار

بنه‌های مورد آزمایش از نظر بروز علائم بیماری‌های قارچی به کمک آزمایشگاه قارچ‌شناسی دانشگاه لرستان ارزیابی گردیدند. پس از انتخاب بنه‌های سالم، بنه‌ها ضدعفونی سطحی شدند و سپس اعمال تیمار قارچی صورت گرفت. برای انجام تیمار، کشت تازه از قارچ *F. oxysporum* تهیه شد. چهار قطعه با اندازه‌ی 1 cm^2 از کشت جامد قارچ در محیط حاوی ۲۵۰ mL آب مقطر استریل و ۲۵g کاه و کلش (محیط کشت تکثیر اسپور قارچ) رشد داده شد و بر روی شیکر با ۱۲۲ دور در دقیقه به مدت ۴ روز قرار داده شد (Wegener, 1992). برای رسیدن به غلظت اسپور مورد نیاز در انجام تیمار به روش زیر عمل شد: (۱) برای شمارش هاگ‌های قارچ مقدار ۲ μL از سوسپانسیون هاگ‌های قارچ بر روی لام هموسیتمتری یا نتوبار (گلبول شمار) قرار داده شد. تعداد ۵ مربع به صورت تصادفی از مربع‌های لام انتخاب و شمارش هاگ‌های موجود در آن‌ها در زیر میکروسکوپ با بزرگ‌نمایی ۴۰ صورت گرفت. (۲) عدد حاصل از میانگین شمارش ۵ مربع لام در فرمول‌های زیر جای‌گذاری شدند و میزان غلظت هاگ‌های قارچ مشخص شد (Davari et al., 2015).

نشانگرهای SSR دو و سه نوکلئوتیدی به ترتیب با ۱۵/۶۹ و ۱۱/۹۱ درصد، بیشترین فراوانی را داشتند (Qian et al., 2019). همچنین توالی‌یابی ژنوم زعفران نیز منجر به شناسایی تعداد ۹۶۴۲۳۱ نشانگر SSR در پایگاه MISA شد که تکرارهای تک نوکلئوتیدی (۵۰/۴ درصد)، دو نوکلئوتیدی (۳۰/۵ درصد) و سه نوکلئوتیدی (۵/۲ درصد) بیشترین فراوانی را داشتند (Ambardar et al., 2021). با وجود ارزش اقتصادی بالای زعفران اطلاعات ژنتیکی بسیار کمی از این گیاه در دسترس پژوهش‌گران قرار دارد. بنابراین و با در نظر گرفتن این مطلب که زعفران گیاهی خودناسازگار و تریپلوئید است و تکثیر آن به صورت رویشی و از طریق بنه‌های زعفران صورت می‌گیرد، در اختیار داشتن اطلاعات ژنتیکی آن از جمله اطلاعات نشانگرهای مولکولی هم‌چون SSR می‌تواند در برنامه‌های به‌نژادی این گیاه از اهمیت به‌سزایی برخوردار باشد. از آن‌جا که تکثیر زعفران به صورت رویشی و از طریق بنه‌های زعفران صورت گرفته و بنه‌ها چندین سال در مزارع باقی می‌مانند، امکان حمله‌ی بیمارگرهای قارچی، باکتریایی، ویروسی و نماتدی وجود دارد. از این‌رو حفظ بنه‌های زعفران از بیمارگرها و سالم ماندن بنه‌ها چالشی است که مزارع کشت زعفران برای حفظ کیفیت و عملکرد محصول با آن روبرو هستند. پوسیدگی طوقه‌ای حاصل از آلودگی قارچ *Fusarium oxysporum* f. sp. *gladio* مخرب‌ترین بیماری در کشت زعفران است که کاهش شدید در عملکرد را موجب می‌شود. قارچ *F. oxysporum* پاتوژنی خاک‌زاد با دامنه‌ی میزبانی وسیع است که پژمردگی آوندی را در گیاه موجب می‌گردد. پژوهش‌های پیشین در رابطه با بررسی نشانگرها و ازجمله نشانگرهای SSR بعد از رشد گل و برگ‌های زعفران صورت گرفته است اما تا به حال پژوهشی بر روی ترانسکریپتوم بنه‌های زعفران به‌ویژه در مواجهه با بیماری حاصل از حمله *F. oxysporum* صورت نپذیرفته است. از این‌رو پژوهش حاضر به

(۱)

$$\text{Average spores per large square (N)} = \frac{\text{The total number of spores in five large random squares}}{5 \text{ or number of large random squares}}$$

(۲)

$$\text{Average spores per small square (n)} = \frac{\text{Average spores per large square}}{16 \text{ or number of small squares in large squares}}$$

(۳)

Spore concentration per milliliter = $n * 4 * 10^6$

پایان به جای تزریق سوسپانسیون قارچی، آب مقطر استریل به نمونه‌ی شاهد تزریق شد. پس از سپری شدن مدت دو هفته وضعیت رشدی قارچ بر روی بنه‌های آلوده شده با روش چشمی بررسی گردید (شکل ۱) و سپس برای بنه‌های تیمار شاهد و تیمار قارچی، RNA کل در دو تکرار بیولوژیک استخراج و برای توالی‌یابی به روش RNA-seq ارسال گردید.

در اعمال تیمار بنه‌ها با قارچ، آماده‌سازی سوسپانسیون قارچی در زمان انجام تیمار صورت گرفت و از سوسپانسیون تازه‌ی قارچ *F. oxysporum* استفاده شد. برای انجام تیمار قارچی بر روی هر یک از بنه‌های زعفران زخم ایجاد شد و با سرنگ میزان ۵ µL از محلول سوسپانسیون قارچ به بنه‌ها تزریق گردید. همه‌ی مراحل برای نمونه‌ی شاهد نیز انجام گرفت و در



شکل ۱- بنه‌های زعفران آلوده شده به قارچ *Fusarium oxysporum* بنه‌های سمت راست و بنه سمت چپ شاهد است
Figure 1- The infected saffron corms with *Fusarium oxysporum*. Right corm, infected, and left control.

استخراج RNA و توالی‌یابی

نانومتر نمونه‌های استخراجی با استفاده از دستگاه نانودراپ (Thermo Scientific NanoDrop 2000) تعیین شد و نمونه‌ها برای توالی‌یابی به شرکت (Beijing Genomes Institute) ارسال شدند. اولین مرحله در فرایند توالی‌یابی بررسی کیفیت و تعیین معیار عدم خردشدگی RNAها یا تعیین عدد

از بافت بنه‌ی زعفران، استخراج RNA کل با استفاده از کیت ترایزول (شرکت NEB: New England Bio labs) صورت گرفت. کیفیت RNAهای استخراج شده به کمک الکتروفورز ژل آگارز ۱ در صد بررسی گردید. همچنین کمیت و نسبت جذب نور در طول موج‌های ۲۶۰ به ۲۸۰ و ۲۶۰ به ۲۳۰

تفسیر کارکردی یونی‌ژن‌های یکپارچه‌سازی شده‌ی زعفران دارای SSR

با بهره‌گیری از نرم‌افزار TransDecode توالی‌های نوکلئوتیدی به توالی‌های پروتئینی ترجمه شدند و فرایند BLAST در بانک‌های اطلاعاتی برای یافتن شباهت صورت گرفت. با این هدف تک رونوشت‌های یکپارچه‌سازی شده‌ی زعفران در پایگاه‌های اطلاعاتی Nr^۲ و UniprotKB^۳ با بهره‌گیری از نرم‌افزار BLASTX با احتمال خطای ۱۰^{-۵} و توالی‌های پروتئینی ترجمه شده نیز در بانک اطلاعاتی UniprotKB با بهره‌گیری از نرم‌افزار BLASTP با احتمال خطای ۱۰^{-۵} مورد ارزیابی قرار گرفتند. همچنین پیش‌بینی دومین‌های پروتئینی با نرم‌افزار HMMER علیه پایگاه اطلاعات Pfam^۴ صورت پذیرفت و پیش‌بینی دومین‌های ترانس‌ممبرانی^۵ و پپتیدهای سیگنال^۶ با استفاده از پایگاه SignalP انجام شد. در ادامه برای مشخص کردن مسیر بیولوژیکی از سرور GhostKOALA (<https://www.kegg.jp/blastkoala>) و KAAS^۷ (<https://www.genome.jp/kegg/kaas>) استفاده گردید. همچنین از وب سرور WEGO^۸ (<http://wego.genomics.org.cn>) برای انجام هستی‌شناسی ژن^۹ و طبقه‌بندی کارکردی "عبارت‌ها"^{۱۰} استفاده گردید و بر این اساس "عبارت‌ها" در سه دسته بیولوژی، سلولی و مولکولی جای گرفتند.

نتایج و بحث

میانگین محتوای درصد GC برای نمونه‌ها ۴۳/۴۴ درصد بود (جدول ۱) که این درصد از میزان GC با میزان درصد GC در

RIN^۱ بود که با استفاده از دستگاه بیوانالایز مدل ۲۱۰۰ (شرکت Agilent) محاسبه شد که این شاخص در نمونه‌های مورد آنالیز ≥ 7 بود. در ادامه cDNA با استفاده از کیت Truseq mRNA stranded سنتز شد. پس از سنتز cDNA، توالی‌یابی بر اساس چارچوب فنی Novaseq6000، با روش خوانش‌های جفتی (Paired) با طول خوانش ۱۵۰ نوکلئوتید و ۶ گیگابایت دیتا صورت پذیرفت.

سنجش کیفیت و ویرایش داده‌ها

سنجش کیفیت داده‌ها با استفاده از نرم‌افزار FastQC (نسخه ۵.۰.۱۱) انجام شد (Andrews, 2010). از نرم‌افزار Trimmomatic برای حذف توالی‌های آداپتورها، بررسی کیفیت بازها در انتهای ۳' و ۵'، حذف نوکلئوتیدها و خوانش‌های با کیفیت پایین استفاده شد. یکپارچه‌سازی خوانش‌ها با استفاده از نرم‌افزار Trinity با K-mer=25 (Haas et al., 2013) و حداقل طول رونوشت‌ها ۲۰۰ جفت باز صورت پذیرفت و در نهایت تک رونوشت‌هایی (Unitranscripts) به دست آمدند که در ادامه از گروه‌بندی آن‌ها تک ژن‌ها حاصل شدند.

شناسایی نشانگرهای SSR

با استفاده از ابزار جستجوی پایگاه MISA (<https://webblast.ipk-gatersleben.de/misa>) جایگاه‌های ریزماهورهای زعفران شناسایی شدند و حداقل تکرار برای نشانگرهای تک نوکلئوتیدی "۱۰"، دو نوکلئوتیدی "۶"، سه، چهار، پنج و شش نوکلئوتیدی "۵" در نظر گرفته شد (Beier et al., 2017).

۶-Signal peptide

۷-KEGG Automatic Annotation Server

۸-Web Gene Ontology Annotation Plot

۹-Gene Ontology

10- Terms

۱- RNA Integrity Number

۲-Non-redundant protein

۳-Database of protein sequence and functional information

۴-Protein families database

۵-Transmembrane domain

سایر گزارش‌های ارائه شده در مورد گیاه زعفران مطابقت داشت (Jain et al., 2016; Ambardar et al., 2021).

جدول ۱- نتایج رونوشت‌های یکپارچه‌سازی شده در نرم‌افزار Trinity
Table 1- Results of transcripts integrated into Trinity software

آماره Parameter	ترینیتی Trinity
کل ژن‌های شناسایی شده Total identified genes	1684478
کل رونوشت‌های شناسایی شده Total identified transcripts	357028
درصد GC GC percent	43.44
N50	951
میانگین طول کانتیگ Contig average	750
کل بازهای سرهم‌بندی شده Total assembled bases	267602092

وضعیت تکرار در نشانگرهای SSR

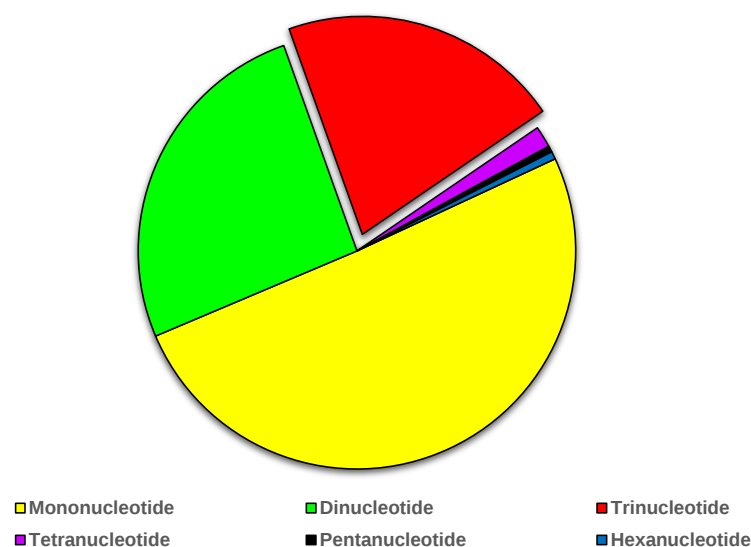
نشانگرهای EST-SSR ریزماهواره‌هایی هستند که روی قطعات کدکننده ژن قرار گرفته‌اند و ممکن است با نواحی عملکردی مرتبط باشند همچنین قادر به ارزیابی تکامل مولکولی ژنی هستند که روی آن قرار دارند. از طرفی حفاظت بیشتر نواحی بالا و پایین دست توالی EST-SSR، این نشانگرها را در بررسی تنوع ژنتیکی بین گونه‌ها نسبت به نشانگرهای ریزماهواره‌ی ژنومی کارآمدتر می‌سازد (Luro et al., 2008; Jiang et al., 2018). در این پژوهش، بررسی نشانگرها با استفاده از ابزار جستجوی MISA تعداد ۸۹۸۸۸ نشانگر SSR را در ۷۰۶۰ توالی شناسایی کرد (جدول ۲) که در نتیجه‌ی این تجزیه و تحلیل، تکرارهای ۱ نوکلئوتیدی بیشترین فراوانی (۵۰٪) را داشتند (شکل ۲). تکرارهای ۲، ۳، ۴، ۵ و ۶ نوکلئوتیدی به ترتیب با ۲۶٪، ۲۰/۹٪، ۱/۶٪، ۵۶٪ و ۰/۴۶٪، فراوانی سایر SSRهای تجزیه و تحلیل شده را به خود اختصاص دادند. با بررسی ترانسکریپتوم زعفران در مراحل قبل و بعد از گل‌دهی تعداد ۷۹۰۲۸ نشانگر SSR یک تا شش نوکلئوتیدی شناسایی

شدند که در این بین SSRهای تک نوکلئوتیدی ۷۱/۲ درصد کل SSRهای شناسایی شده را به خود اختصاص داده و SSRهای دو و سه نوکلئوتیدی به ترتیب با ۱۵/۶۹ و ۱۱/۹۱ درصد بیشترین فراوانی را داشتند (Qian et al., 2019). همچنین توالی‌یابی ژنوم زعفران نیز ۹۶۴۲۳۱ نشانگر SSR را در ابزار جستجوی MISA شناسایی کرد و تکرارهای تک نوکلئوتیدی (۵۰/۴ درصد)، دو نوکلئوتیدی (۳۰/۵ درصد) و سه نوکلئوتیدی (۱۵/۲ درصد) بیشترین فراوانی را داشتند (Ambardar et al., 2021). بر اساس نتایج به‌دست آمده و گزارش‌های پیشین (Jain et al., 2016; Qian et al., 2019; Ambardar et al., 2021) نشانگرهای SSR تک نوکلئوتیدی نوع غالب این گروه از نشانگرها در گیاه زعفران هستند. در بررسی دیگری روی تعداد ۶۷۴۵ توالی EST در زعفران، تعداد ۱۰۸ توالی EST-SSR شناسایی شدند که در کل ژنوم توزیع شده بودند. در این پژوهش توالی‌های دو و سه نوکلئوتیدی بیشترین فراوانی را در بین نشانگرهای EST-SSR داشتند (Guo-Qing, 2011).

جدول ۲- نشانگرهای SSR شناسایی شده ی بنه ی زعفران با استفاده از ابزار جستجوی MISA

Table 2- Detected SSR markers of saffron corm using MISA search tool

تعداد	موارد بررسی شده
Number	Items reviewed
357028	تعداد کل رونوشت ها Total number of transcripts
267602092	تعداد کل جفت بازهای ارزیابی شده The total number of evaluated base pairs
89888	تعداد کل SSRهای شناسایی شده The total number of detected SSRs
70060	تعداد توالی های دربرگیرنده ی SSR Number of sequences containing SSRs
15056	تعداد توالی های دربرگیرنده ی بیش از یک SSR Number of sequences containing more than one SSR
9619	تعداد SSRهای موجود در فرم های ترکیبی Number of SSRs in combined forms



شکل ۲- فراوانی نسبی نشانگرهای SSR شناسایی شده (یک، دو، سه، چهار، پنج و شش نوکلئوتیدی) در بنه زعفران با استفاده از ابزار جستجوی MISA

Figure 2- Using the MISA search tool, the saffron corm shows the relative frequency of identified SSR markers (Mono, Di, Tri, Tetra, Penta, and Hexa nucleotides).

براساس تجزیه و تحلیل انجام شده AAG/CTT, GAA/TTC و TCT/AGA در میان SSRهای با سه تکرار بیشترین فراوانی را به خود اختصاص دادند. نتایج به دست آمده با نتایج سایر پژوهش ها مطابقت داشت به گونه ای که تکرارهای AAG/CTT و AAC/GTT و AGG/CCT به عنوان پرتکرارترین SSRهای

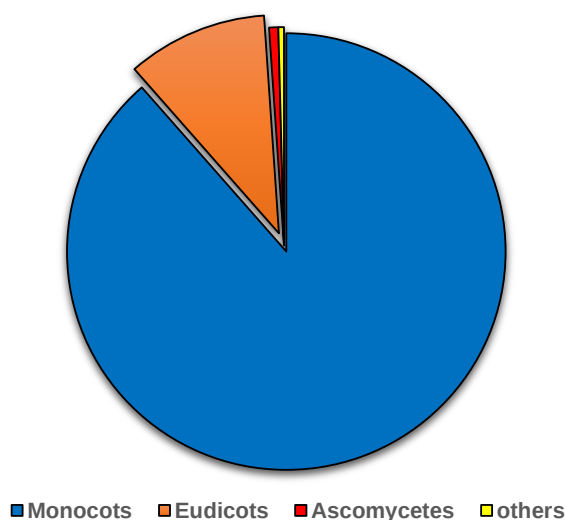
در بین تکرار های دو نوکلئوتیدی تکرار های AG/CT و GA/TC بیشترین فراوانی را داشتند که با نتایج جین و همکاران (Jain et al., 2016) همخوانی داشت. همچنین بیشترین فراوانی SSRهای دو نوکلئوتیدی در مطالعه ی جیان و همکاران دی نوکلئوتیدهای CA و TG بودند (Qian et al., 2019).

Catalytic activity و Binding و در دسته‌ی فرآیندهای بیولوژیکی Metabolic process و Cellular process بیشترین درصد را به خود اختصاص دادند. نتایج حاصل با نتایج حاصل از تجزیه و تحلیل ترانسکریپتوم در مطالعه‌ی بابا و همکاران (Baba et al., 2015) مشابه بود و در دسته‌بندی گزارش شده در این مطالعه نیز عبارت‌های Membrane و Cell part در دسته‌ی اجزای سلولی، Catalytic activity و Binding در دسته‌ی عملکرد مولکولی و Metabolic process و Cellular process در دسته‌ی فرآیندهای بیولوژیکی بیشترین فراوانی را داشتند. همچنین بر اساس نتایج حاصل از الگوریتم BLAST، بیشترین میزان شباهت رونوشت‌های حاوی SSRها با برنج و آرابیدوپسیس به‌دست آمد و در مجموع از میان یونی‌ژن‌های شناسایی شده تعداد ۱۸۸۴۶، ۲۳۹۸۸ و ۱۰۹۶۹ یونی‌ژن به ترتیب در پایگاه‌های داده‌ی UNIPROT، Nr و GO شناسایی شدند که تعداد ۱۰۳۷۵ یونی‌ژن در بین همه‌ی پایگاه‌ها مشترک بود (شکل ۵).

سه نوکلئوتیدی زعفران گزارش شدند (Jain et al., 2016).

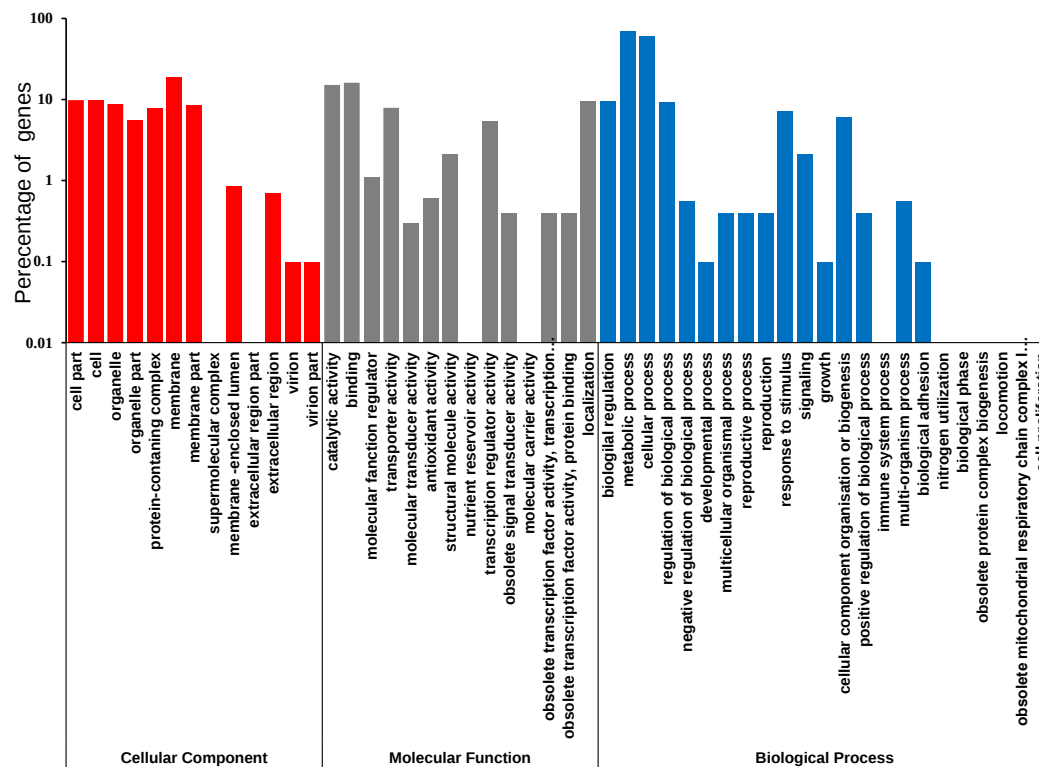
هستی‌شناسی و تفسیر عملکردی یونی‌ژن‌های یکپارچه‌سازی شده‌ی زعفران دارای SSR

در داده‌های مستندسازی در GhostKOALA، ۸۸/۵ درصد رونوشت‌های دارای نشانگر SSR به رده‌ی تک‌لپه‌ها، ۱۰/۳۶ درصد به دولپه‌ای‌ها، ۰/۶۷ درصد به آسکومیست‌ها و ۰/۴۱ درصد به سایر گونه‌ها اختصاص یافت (شکل ۳). با توجه به این که زعفران از نظر طبقه‌بندی گیاه‌شناسی به رده‌ی تک‌لپه‌ای‌ها تعلق دارد، این نتیجه دور از انتظار نبود و صحت تجزیه و تحلیل‌ها تا این مرحله را نشان داد. همچنین از ابزار وب WEGO برای مشخص کردن ماهیت رونوشت‌ها استفاده شد. بر این اساس، ۵۰ دسته‌ی کارکردی حاصل شد که "عبارت‌ها"ی حاصل در سه دسته‌ی اجزای مولکولی، عملکرد مولکولی و فرآیندهای بیولوژیکی جای گرفتند (شکل ۴). در دسته‌ی اجزای سلولی Membrane و Cell part در دسته‌ی عملکرد مولکولی

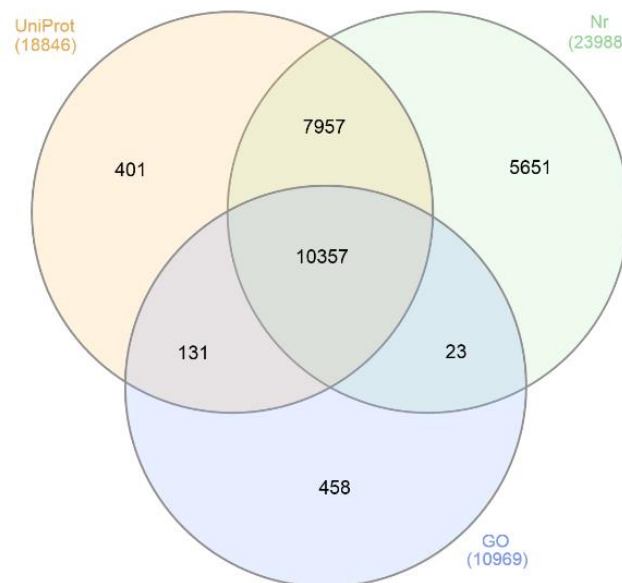


شکل ۳- مستندسازی رونوشت‌های دارای نشانگر SSR شناسایی شده بیه زعفران در ابزار مستندسازی GhostKOALA در مقایسه با اطلاعات موجود از گیاهان تک‌لپه‌ای، دولپه‌ای و آسکومیست‌ها

Figure 3- Annotation of saffron corm SSRs by GhostKOALA documentation compared to the available data on monocotyledons, dicotyledons plants, and ascomycetes.



شکل ۴- دسته‌بندی رونوشت‌های دارای نشانگر SSR شناسایی شده در بنه زعفران توسط ابزار وب WEGO
Figure 4- Classification of saffron corms transcripts identified by WEGO web tool.



شکل ۵- نمودار ون^۱ مربوط به رونوشت‌های سرهم‌بندی شده ترانسکریپتوم بنه زعفران در پایگاه‌های GO، Uniprot KB و non redundant (Nr) با استفاده از الگوریتم BLAST
Figure 5- Van diagram of transcripts assembled of saffron corm in GO, Uniprot KB, and non-redundant (Nr) databases using the BLAST algorithm.

۱- Venn diagram

(et al., 2019).

نتیجه گیری

ادویه‌ی زعفران گران‌بها‌ترین و ارزشمندترین ادویه‌ی جهان است که با وجود اهمیت اقتصادی آن برای پیشبرد اهداف به‌نژادی آن، اطلاعات زیادی از ژنتیک آن در اختیار نیست. نشانگرهای مولکولی هم‌بارز از جمله نشانگرهای SSR در پروژه‌های به‌نژادی از اهمیت ویژه‌ای برخوردار هستند. در پژوهش حاضر تجزیه و تحلیل ترانسکریپتوم زعفران نشان داد که بیشترین میزان شباهت با برنج و آرابیدوپسیس بود. در مجموع، تعداد ۸۹۸۸ توالی SSR با استفاده از ابزار جستجوی MISA شناسایی شدند که در این بین توالی‌های تکرارهای یک و دو نوکلئوتیدی بیشترین فراوانی را به خود اختصاص دادند، بر اساس نتایج به‌دست آمده از این مطالعه مشخص شد که تکرارهای یک و دو نوکلئوتیدی (AG/CT و GA/TC) بیشترین فراوانی را در بین نشانگر SSR داشتند. پژوهش حاضر توانسته است با در اختیار داشتن توالی کل mRNAهای بنه‌ی زعفران، SSRهای مرتبط با ژن‌های بیان شونده را شناسایی و معرفی نماید که انتظار می‌رود در ساخت نقشه‌ی ژنتیکی و پیوستگی زعفران کاربرد داشته باشند.

در برنامه‌های به‌نژادی، استفاده از نشانگرهای مولکولی با هدف تخمین تنوع ژنتیکی در بررسی روابط خویشاوندی و تشخیص ژرم‌پلاسم‌ها استفاده می‌شود. شناسایی نشانگرهای مرتبط با توالی رونوشت‌های بیانی به این دلیل که به صورت مستقیم با ژن‌های کارکردی در ارتباط هستند در پروژه‌ها و اهداف به‌نژادی از اهمیت به‌سزایی برخوردار است (Li et al., 2002). از این‌رو نشانگرهای SSR به سبب دارا بودن ویژگی‌هایی از جمله هم‌بارزی، تکرارپذیری، توزیع مناسب در ژنوم و آسانی تجزیه و تحلیل، نسبت به سایر نشانگرها توجه بیشتری را به خود معطوف کرده‌اند. از آن‌جا که نشانگرهای SSR خاص هر گونه هستند، برای دستیابی به توالی آن نیاز به شناخت کامل توالی گیاه یا موجود مورد مطالعه است (Shokeen et al., 2011; Kumar & Bhatia, 2016). در این راستا، مطالعات ترانسکریپتومی قادر است SSRها را در مقیاس گسترده و به نسبت سریع و ارزان شناسایی کند. به‌طوریکه در اختیار داشتن توالی کامل ترانسکریپتوم با روش RNA-Sequencing گونه گیاهی مورد مطالعه می‌تواند این هدف را تسهیل کرده و تعداد زیادی نشانگر ریزماهواره را با پوشش کلی‌تری در مقایسه با ESTها که بخش بسیار کوچکی از اگزون‌ها را در برمی‌گیرند، فراهم آورد و در نتیجه تخمین فراوانی تکرارهای SSRها و خصوصیات آن‌ها کمتر دچار خطا می‌شود (Soltani

منابع

- Alavi-Kia, S., Mohammadi, S., Aharizad, S., and Moghaddam, M. 2008. Analysis of genetic diversity and phylogenetic relationships in *Crocus* genus of Iran using inter-retrotransposon amplified polymorphism. *Biotechnology and Biotechnological Equipment* 22: 795-800.
- Alipoor Miandehi, Z., Mahmoodi, S., Behdani, M.A., and Sayyari, M.H. 2014. Effects of corm weight and application of fertilizer types on some growth characteristics and yield of saffron (*Crocus sativus* L.) under Mahvelat conditions. *Journal of Saffron Research* 2: 97-112.
- Ambardar, S., Vakhlu, J., and Sowdhamini, R. 2021. De-novo draft genome sequence of *Crocus sativus* L, saffron, a golden condiment. *BioRxiv*.

- Andrews, S. 2010. FastQC: A quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Baba, S.A., Mohiuddin, T., Basu, S., Swarnkar, M.K., Malik, A.H., Wani, Z.A., Abbas, N., Singh, A.K., and Ashraf, N. 2015. Comprehensive transcriptome analysis of *Crocus sativus* for discovery and expression of genes involved in apocarotenoid biosynthesis. BMC Genomics 16: 1-14.
- Babaei, S., Talebi, M., Bahar, M., and Zeinali, H. 2014. Analysis of genetic diversity among saffron (*Crocus sativus*) accessions from different regions of Iran as revealed by SRAP markers. Scientia Horticulturae 171: 27-31.
- Beier, S., Thiel, T., Münch, T., Scholz, U., and Mascher, M. 2017. MISA-web: a web server for microsatellite prediction. Bioinformatics 33: 2583-2585.
- Chen, J., Li, R., Xia, Y., Bai, G., Guo, P., Wang, Z., Zhang, H., and Siddique, K.H. 2017. Development of EST-SSR markers in flowering Chinese cabbage (*Brassica campestris* L. ssp. *chinensis* var. *utilis* Tsen et Lee) based on de novo transcriptomic assemblies. PloS One 12: e0184736.
- Festuccia, C., Mancini, A., Gravina, G.L., Scarsella, L., Llorens, S., Alonso, G.L., Tatone, C., Di Cesare, E., Jannini, E.A., and Lenzi, A. 2014. Antitumor effects of saffron-derived carotenoids in prostate cancer cell models. BioMed Research International 2014: 1-12.
- Guo-Qing, C. 2011. Analysis of SSR information in EST resource of saffron (*Crocus sativus*). Guangxi Plant 31: 43-46.
- Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., Couger, M.B., Eccles, D., Li, B., and Lieber, M. 2013. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols 8 (8): 1494-1512.
- Hiremath, P.J., Farmer, A., Cannon, S.B., Woodward, J., Kudapa, H., Tuteja, R., Kumar, A., BhanuPrakash, A., Mulaosmanovic, B., and Gujaria, N. 2011. Large-scale transcriptome analysis in chickpea (*Cicer arietinum* L.), an orphan legume crop of the semi-arid tropics of Asia and Africa. Plant Biotechnology Journal 9: 922-931.
- Jain, M., Srivastava, P.L., Verma, M., Ghangal, R., and Garg, R. 2016. De novo transcriptome assembly and comprehensive expression profiling in *Crocus sativus* to gain insights into apocarotenoid biosynthesis. Scientific Reports 6: 1-13.
- Jiang, Y., Li, H., Zhang, J., Xiang, J., Cheng, R., and Liu, G. 2018. Whole genomic EST-SSR development based on high-throughput transcript sequencing in proso millet (*Panicum miliaceum*). International Journal of Agriculture and Biology 20: 617-620.
- Koocheki, A., Siah, M.A., Azizi, G., and Jahani, K.M. 2011. The effect of high density and depth of planting on agronomic characteristics of Saffron (*Crocus sativus* L.) and corms behavior. Journal of Agroecology 3 (1): 36-45.
- Kumar, S., and Bhatia, S. 2016. A polymorphic (GA/CT)_n-SSR influences promoter activity of Tryptophan decarboxylase gene in *Catharanthus roseus* L. Don. Scientific Reports 6: 33280.
- Li, Y.C., Korol, A.B., Fahima, T., Beiles, A., and Nevo, E. 2002. Microsatellites: genomic distribution, putative functions and mutational mechanisms :a review. Molecular Ecology 11:

- 2453-2465.
- Luro, F.L., Costantino, G., Terol, J., Argout, X., Allario, T., Wincker, P., Talon, M., Ollitrault, P., and Morillon, R. 2008. Transferability of the EST-SSRs developed on *Nules clementine* (*Citrus clementina* Hort ex Tan) to other *Citrus* species and their effectiveness for genetic mapping. *BMC Genomics* 9: 1-13.
- Millan, T., Clarke, H.J., Siddique, K.H., Buhariwalla, H.K., Gaur, P.M., Kumar, J., Gil, J., Kahl, G., and Winter, P. 2006. Chickpea molecular breeding: new tools and concepts. *Euphytica* 147: 81-103.
- Nemati, Z., Harpke, D., Gemicioglu, A., Kerndorff, H., and Blattner, F.R. 2019. Saffron (*Crocus sativus*) is an autotriploid that evolved in Attica (Greece) from wild *Crocus cartwrightianus*. *Molecular Phylogenetics and Evolution* 136: 14-20.
- Nasiri, E., Sabouri, A., Forghani, A., and Esfahani, M. 2019. Grouping of rice genotypes based on grain iron, zinc, manganese, and protein and performance measurement of linked microsatellite markers. *Plant Genetic Researches* 5: 73-84.
- Nemati, Z., Zeinalabedini, M., Mardi, M., Pirseyediand, S.M., Marashi, S.H., and Khayam Nekoui, S.M. 2012. Isolation and characterization of a first set of polymorphic microsatellite markers in saffron, *Crocus sativus* (Iridaceae). *American Journal of Botany* 99: 340-343.
- Qian, X., Sun, Y., Zhou, G., Yuan, Y., Li, J., Huang, H., Xu, L., and Li, L. 2019. Single-molecule real-time transcript sequencing identified flowering regulatory genes in *Crocus sativus*. *BMC Genomics* 20: 1-18.
- Shokeen, B., Choudhary, S., Sethy, N.K., and Bhatia, S. 2011. Development of SSR and gene-targeted markers for construction of a framework linkage map of *Catharanthus roseus*. *Annals of Botany* 108: 321-336.
- Soltani, H.M., SADAT, N.S.A., Shariati, V., and Amiripour, M. 2019. Large scale identification of SSR molecular markers in Ajowan (*Trachyspermum ammi*) using RNA sequencing. *Journal of Plant Genetics Researches* 6 (1): 31-46.
- Strickler, S.R., Bombarely, A., and Mueller, L.A. 2012. Designing a transcriptome next-generation sequencing project for a nonmodel plant species1. *American Journal of Botany* 99: 257-266.
- Thiel, T., Michalek, W., Varshney, R., and Graner, A. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 106 (41): 411-422.
- Wegener, M. 1992. Optimierung von saatgutpillierungen mit mikrobiellen antagonistien zur biologischen biologischen bekämpfung von *Fusarium culmorum* (WG Smith) Sacc. in weizen. Weizen. Diplomarbeit, Universität Gottingen.
- Zietkiewicz, E., Rafalski, A., and Labuda, D. 1994. Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification. *Genomics* 20: 176-183.