



## غربالگری توالی های ثبت شده از نواحی بیان شونده ژنوم پرتقال برای تشخیص نشانگرهای مبتنی بر میکروساتلایت

عبدالکریم زارعی<sup>۱\*</sup>، عزیز ابراهیمی<sup>۲</sup>

<sup>۱\*</sup> گروه بیوتکنولوژی (تولید و ژنتیک گیاهی)، دانشکده کشاورزی، دانشگاه جهرم، جهرم

<sup>۲</sup> گروه جنگلداری و منابع طبیعی، دانشگاه پوردو، وست لافایت، ایالات متحده آمریکا

\*نویسنده مسئول: zarei@jahromu.ac.ir

### چکیده

به منظور شناسایی و تعیین ویژگی انواع توالی های تکراری و توسعه نشانگرهای میکروساتلیتی مبتنی بر آنها، در پژوهش حاضر توالی های رونوشت ثبت شده در بانک داده های اسید نوکلئیک گیاه پرتقال (*Citrus sinensis*) مورد غربالگری و آنالیز قرار گرفت. بطور کلی ۲۱۴۵۹۸ توالی رونوشت مربوط به پرتقال از بانک اطلاعاتی NCBI استخراج و مورد آنالیز قرار گرفت که از بین آنها ۵۴۷۵ عدد حداقل دارای یک ناحیه میکروساتلیتی بودند (بطور میانگین ۲/۵ درصد از توالی های ثبت شده). در مجموع ۵۶۳۲ توالی تکراری در بین توالی های بررسی شده وجود داشت که ۱۱۸ عدد از آنها از نوع توالی های مرکب بودند. از بین توالی های مورد آنالیز ۳۰۹۷ عدد دارای مناطق احاطه کننده مناسب بوده و قابلیت تبدیل شدن به نشانگرهای میکروساتلیتی را دارا بودند. نتایج تجزیه توالی های بررسی شده نشان داد که به ترتیب تکرارهای دوتائی (۴۵/۲٪)، سه تائی (۳۰/۷۲٪) و چهار تائی (۲۳/۱٪) شایع ترین تکرارها در بین تکرارهای موجود در رونوشت های ثبت شده گیاه پرتقال می باشند. روند مشابهی در بین نشانگرهای میکروساتلیتی توسعه یافته مشاهده گردید بطوریکه تعداد نشانگرهایی که برای آنها آغازگر طراحی گردید به ترتیب در مورد تکرارهای دوتائی، سه تائی و چهارتائی، ۱۱۶۸، ۱۱۵۹ و ۷۰ عدد بود. تنوع بالایی هم از نظر نوع توالی در نوع هر تکرار مشاهده گردید، بطوریکه در تکرارهای دوتائی نوع تکرارهای AG/GA بیشترین فراوانی (۱۸۶۴) و تکرارهای GC/CG کمترین فراوانی (۶ عدد) را به خود اختصاص دادند. با توجه به قابلیت بالای انتقال پذیری نشانگرهای میکروساتلیتی مبتنی بر ESTs به نظر می رسد نشانگرهای شناسائی شده توانائی بالائی برای شناسایی و گروه بندی انواع مرکبات دارا بوده و در مطالعات مربوط به شناسایی، نقشه یابی و انتخاب به کمک نشانگر در جنس سیتروس مفید واقع شوند.

**کلمات کلیدی:** آغازگر، پرتقال، توالی های بیان شونده، میکروساتلیت، نواحی تکراری.

### مقدمه

پرتقال (*Citrus sinensis*) گیاهی دیپلوئید ( $2n = 2x = 18$ ) و با اندازه ژنومی معادل ۳۸۰ مگا باز بوده که بعنوان یکی از مهمترین میوه های مناطق گرمسیری و نیمه گرمسیری دنیا اهمیت بالایی در سرتاسر دنیا دارا می باشد (Xu et al., 2013). با توجه به اهمیت اقتصادی بالای این میوه، مطالعات ژنتیکی زیادی روی این گونه انجام گرفته است. پیشرفت های اخیر در زمینه توالی یابی محققین را قادر ساخته تا بتوانند با هزینه کمتر و دقت بالاتر به شناسای انواع نشانگرهای مولکولی اقدام نمایند. توالی های تکراری ساده (SSRs) یکی از مهمترین نشانگرهای مولکولی می باشند که بعنوان یک ابزار چند منظوره توانایی خود را در مطالعات متنوع گیاهی و جانواری به اثبات رسانده اند. شناسایی و توسعه این نشانگرها با استفاده از روش های سنتی بسیار هزینه بر و وقت گیر می باشد. به هر حال توسعه روشهای توالی یابی و در دسترس بودن توالی های ژنومی باعث شده که این توالی ها بتوانند بعنوان منبعی غنی برای شناسایی نشانگرهای میکروساتلیتی مورد استفاده قرار گیرند. توالی های بیان شونده (ESTs) در ژنوم دارای اطلاعات کاربردی بالایی بوده و بعنوان یکی از منابع مناسب برای شناسایی و توسعه نشانگرهای مولکولی در ژنوم موجودات مختلف در بانک های اطلاعاتی مربوط به



اسیدهای نوکلئیک در دسترس می باشند. از مزایای EST-SSRs ها نسبت به SSRهای شناسایی شده از توالی ژنومی، می توان به انتقال پذیری بالای بین گونه ایی، عملکردی بودن تعداد بیشتری از این نوع نشانگرها و تسهیل فرایند انتخاب به کمک نشانگر اشاره کرد (Haq et al. 2014). استفاده از این نشانگرها تاکنون در مورد گونه های گیاهی بسیاری به منظور مطالعات مختلف ژنتیکی مورد استفاده قرار گرفته است و مفید بودن آنها در برای اهداف مختلف از جمله بررسی تنوع ژنتیکی، ژنتیک جمعیت و انتخاب به کمک نشانگر مورد تأیید قرار گرفته است (Wang et al., 2012; Adawy et al., 2013; Haq et al., 2014; Zhou et al., 2015; Sorkheh et al., 2016; Zarei and Ebrahimi, 2017). با هدف شناسایی و تعیین ویژگی نواحی تکراری بیان شونده مربوط به گیاه پرتقال و توسعه نشانگرهای جدید EST-SSRs در این گیاه، در تحقیق حاضر توالی های EST ثبت شده مربوط به این گیاه از بانک های اطلاعاتی استخراج و مورد آنالیز توالی قرار گرفت.

## مواد و روش ها

تعداد ۲۱۴۵۹۸ توالی از نواحی بیان شده مربوط به پرتقال (*Citrus sinensis*) از بانک اطلاعاتی اسید نوکلئیک (NCBI) (<http://www.ncbi.nlm.nih.gov/dbest/>) (NCBI) دانلود و مورد تجزیه و تحیل آماری قرار گرفت. داده های حاصل به منظور جستجوی نواحی تکراری مورد آنالیز قرار گرفتند. به این منظور نرم افزار EST-Trimmer برای حذف انتهای ۵'، ۳' و پلی A از EST های خام استفاده شد و توالی های همپوشان شکل گرفت. توالی های تکراری برای نواحی دارای ۲ جفت باز با تکرار بین ۸-۴۰ مرتبه، ۳ جفت بازی با تکرار ۷-۳۰ مرتبه و ۴ جفت باز با تکرار ۶-۲۰ واحد گزارش شدند. میکروساتلیت های مرکب (که عبارتند از SSR های کنار یکدیگر و یا آنهایی که با حداقل ۱۵ جفت باز از هم جدا شده بودند) در آنالیز قرار نگرفتند. توالی ها با نرم افزار CAP3 به یک کانتیگ تبدیل شده (Huang et al., 1999) و سپس با نرم افزار Primer3 (Ver. 2.3.5) مورد آنالیز قرار گرفته و برای آنها پرایمر طراحی گردید (Untergasser et al., 2012).

## نتایج و بحث

از بین ۲۱۴۵۹۸ توالی بیان شونده مربوط به پرتقال که از بانک های اطلاعاتی استخراج گردید، ۵۴۷۵ عدد حداقل دارای یک توالی تکراری در ساختار خود بودند که معادل ۲/۵۵٪ از کل توالی های آنالیز شده می باشد. در مجموع از توالی های بررسی شده ۵۶۳۲ توالی تکراری میکروساتلیتی شناسایی گردید که حدود ۲/۶۲٪ توالی های شناسایی شده می باشد (جدول ۱).

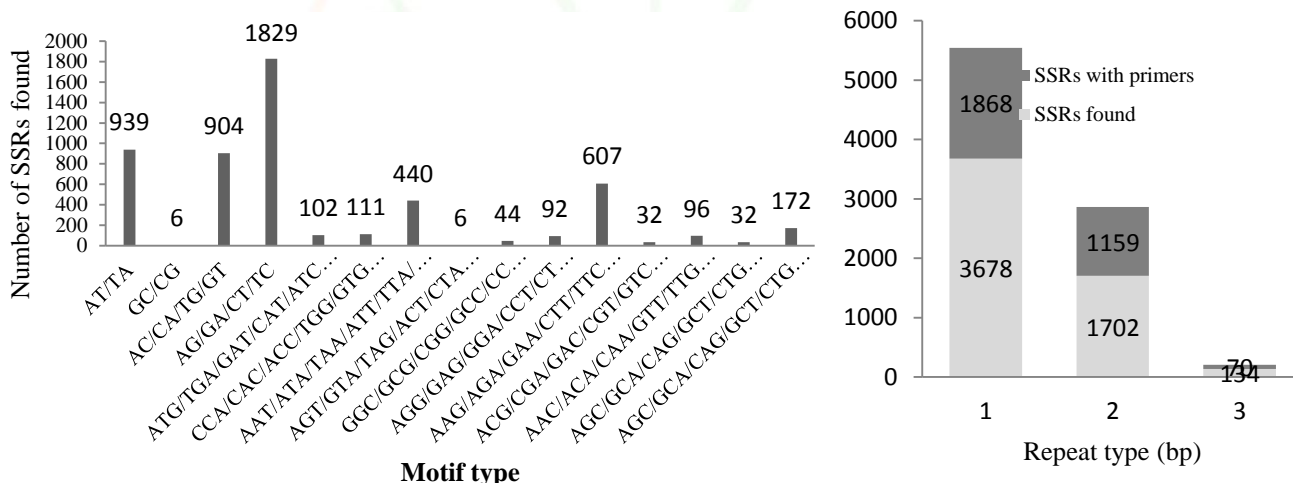
جدول ۱. نتایج حاصل از آنالیز توالی های بیان شونده مربوط به گیاه پرتقال.

شاخص	تعداد
تعداد کل EST های آنالیز شده	۲۱۴۵۹۸
تعداد کل SSR های تشخیص داده شده	۵۶۳۲
تعداد کل توالی هایی که حداقل دارای یک توالی میکروساتلیتی بوده اند	۵۴۷۵
تعداد میکروساتلیت های مرکب	۱۱۸
تعداد EST-SSR هایی که برای آنها آغازگر تهیه شده	۳۰۹۷

از بین تکرارهای های میکروساتلیتی مختلف، تکرارهای دوتائی بالاترین درصد فراوانی را داشته و بیش از ۴۶٪ توالی های تکراری در رونوشت های پرتقال را به خود اختصاص دادند. پس از تکرارهای ۲ واحدی، به ترتیب تکرارهای ۳ و ۴ عددی فراوانترین نوع تکرارها در نواحی بیان شونده ژنوم پرتقال بودند و به ترتیب بیش از ۳۰٪ و ۲۳٪ توالی های



تکراری را شامل می‌شدند. این نتایج با نتایج گزارش شده در مورد توالی‌های تکراری در ژنوم دیگر درختان میوه از قبیل هسته دارها (Sorkhe *et al.*, 2016) و انار (Zarei and Ebrahimi, 2017) مطابقت دارد. این محققین گزارش کردند که تکرارهای دوتائی بیشترین درصد تکرارهای نواحی بیان شونده در ژنوم انواع پرونوس (۲۶/۳۸٪) و انار (۷۰٪) را تشکیل می‌دهند و بطور مشابه تکرارهای سه و چهارتائی در رتبه‌های بعدی وجود داشتند. برعکس نتایج حاصل از این تحقیق، تکرارهای تک نوکلئوتیدی بعنوان شایع‌ترین تکرار (۷۱٪ از کل توالی‌های تکراری) در رونوشت‌های گیاه زیتون گزارش شده است و بعد از آن تکرارهای سه (۱۱/۸۴٪) و دو (۸/۶۲٪) واحدی بیشترین درصد را به خود اختصاص دادند (Adawy *et al.*, 2013). همچنین تکرارهای تک نوکلئوتیدی فراوانترین تکرار در EST-SSR های گیاه *Hevea brasiliensis* بوده و حدود ۳۸/۸۹٪ از تکرارهای رونوشت این گیاه را شامل می‌شود و سپس تکرارهای دو نوکلئوتیدی (۳۶/۹۵٪) و سه نوکلئوتیدی (۱۸/۱۷٪) در رتبه‌های بعدی بودند (Hue *et al.*, 2017). از طرفی Zhou و همکاران (۲۰۱۵) گزارش کردند که تکرارهای سه نوکلئوتیدی شایع‌ترین تکرار (۶۰/۶۶٪ از کل تکرارها) در نواحی بیان شونده گیاه *Elymus sibiricus* L. بوده و بعد از آن تکرارهای دو واحدی (۲۳/۷۹٪) و تک نوکلئوتیدی (۵/۰۶٪) در رده‌های بعدی بودند. با توجه به تنوع نتایج مربوط به گیاهان مختلف به نظر می‌رسد فراوانی و توزیع تکرارها به میزان زیادی وابسته به نوع گیاه مورد مطالعه می‌باشد، هر چند تعداد رونوشت‌های ثبت شده در گیاهان مختلف هم می‌تواند فاکتور موثر و تعیین کننده‌ای در گزارش‌های مختلف باشد.



تصویر ۱. سمت راست: نوع توالی‌های تکراری و میزان فراوانی آنها و توسعه نشانگرهای مبتنی بر ریزماهورها در رونوشت‌های ثبت شده گیاه پرتقال؛ سمت چپ: فراوانی و توزیع میزان توالی‌های میکروساتلیتی شناسایی شده در رونوشت‌های گیاه پرتقال.

همچنین بر اساس نتایج حاصل از این پژوهش مشخص شد که تکرارهای *AG/GA/CT/TC* و *AAG/AGA/GGA* فراوانترین تکرار در بین تکرارهای دو و سه نوکلئوتیدی بوده و تکرارهای *GC/CG* و *AGT/GTA/TAG* کمترین تکرار را در بین تکرارهای مذکور دارا بودند (تصویر ۱). نتایج مشابهی در مورد فراوانی تکرارهای مذکور در مورد گیاه *Hevea brasiliensis* گزارش شده و تکرارهای *AG/CT* بیش از ۸۳٪ از تکرارهای دوتائی موجود در رونوشت‌های این گیاه را



تشکیل داده و تکرارهای AAG/CTT با حدود ۳۸٪ تکرارهای سه تایی بیشترین میزان توالی های تکراری را به خود اختصاص داده اند.

## نتیجه گیری کلی

نتایج حاصل از این پژوهش بیانگر اطلاعاتی در مورد نوع، نحوه توزیع و فراوانی انواع توالی های تکراری در رونوشت های ثبت شده گیاه پرتقال می باشد. همچنین در این پژوهش بیش از ۳۰۰۰ جفت آغازگر اختصاصی برای بررسی EST-SSR های پرتقال شناسایی و گزارش شد که می توان بعنوان یک منبع غنی برای بررسی مطالعات مختلف ژنتیکی در این گونه و همچنین گونه های مشابه نزدیک از آنها استفاده نمود.

## منابع

- Adawy, S.S., Mokhtar, M.M., Alsamman, AM. and Sakr, M.M. 2013. Development of Annotated EST-SSR Database in Olive (*Olea europaea*). International Journal of Science and Research, 4(9): 1063-1073.
- Haq, S.U., Jain, R., Sharma, M., Kachhwaha, S. and Kothari, S.L. 2014. Identification and characterization of microsatellites in expressed sequence tags and their cross transferability in different plants. International Journal of Genomics, Article ID 863948, P. 12.
- Hou, B., Feng, S., and Wu, Y. 2017. Systemic Identification of Hevea brasiliensis EST-SSR Markers and Primer Screening. Journal of Nucleic Acids, Article ID 6590902, p. 9.  
<https://doi.org/10.1155/2017/6590902>
- Huang, X. and Madan, A. 1999. CAP3: A DNA sequence assembly program. Genome Research, 9: 868–877. PMID: 10508846
- Sorkheh, K., Prudencio. A.S., Ghebinejad, A., Kohei Dehkordi, M., Erogu, D., Rubio, M. and Martínez-Gómez, P. 2016. In silico search, characterization and validation of new EST-SSR markers in the genus *Prunus*. BMC Research Notes, 9: 366. DOI 10.1186/s13104-016-2143-y
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G. 2012. Primer3—new capabilities and interfaces. Nucleic Acids Research, 40: 115–115. PMID: 22730293
- Xu, Q., Chen, L.L. and Rutan, X. 2013. The draft genome of sweet orange (*Citrus sinensis*). Nature Genetics, 45: 59-66.
- Zarei, A., Ebrahimi A. 2017. EST-SSR identification in pomegranate expressed sequence tags. 2<sup>nd</sup> International and 10<sup>th</sup> national Biotechnology congress of Islamic Republic of Iran. August 29-31. seed and plant improvement institute, Karaj, Iran.
- Zhou, Q., Luo, D., Ma, L., Xie, W., Wang, Y., Wang, Y. and Liu, Z. 2015. Development and cross-species transferability of EST-SSR markers in Siberian wildrye (*Elymus sibiricus* L.) using Illumina sequencing. Scientific Reports, 6:20549. DOI: 10.1038/srep20549.



## Screening of registered transcript sequences of sweet orange (*Citrus sinensis*) for identification of microsatellite-based markers

Abdolkarim Zarei<sup>1\*</sup>, Aziz Ebrahimi<sup>2</sup>

<sup>1\*</sup> Department of Biotechnology, Agriculture College, Jahrom University, Jahrom, Iran

<sup>2</sup>Department of Forestry and Natural Resources, Purdue University, 715 State Street, West Lafayette, IN 47907, USA

\*Corresponding author: zarei@jahromu.ac.ir

### Abstract

With the aim of identification and characterization of repetitive sequences from expressed sequence tags (ESTs) in *Citrus sinensis* and development of microsatellite-based markers, NCBI deposited sequences of this important fruit species was screened. Data mining in 214589 transcript sequences revealed that 5475 sequences contained at least one SSR (on average 2.5% of the sequences). Altogether 5632 repetitive sequences were identified among which 118 sequences was compound SSRs and 3097 ones had proper flanking regions and hence suitable for primer development. Motif frequencies were highly variable and showed a clear bias toward dinucleotide motifs. On the other hand, dinucleotide repeats were the most frequent (45.2%) repeats in *C. sinensis* transcripts followed by trinucleotide (30.72%), and tetranucleotide (23.11%) repeats. Similarly, developed EST-SSR markers were the highest motif type in the dinucleotide motifs (1168 markers) followed by tri- (1159 markers) and tetra-nucleotides (70 markers). Regarding to the high transferability nature of EST based SSRs, it seems that developed EST-SSRs have a high potential for identification, characterization, and classification in the *Citrus* species and would be highly beneficial for comparative mapping as well as marker assisted selection projects in this genus.

**Keywords:** Primer, Sweet orange, Expressed sequence tags, Microsatellite, Repetitive Sequences.

