Iranian Journal of Horticultural Science
Vol 49, No 2, Summer 2018 (313-321)
DOI: 10.22059/ijhs.2017.201730.952

علوم باغبانی ایران
دورهٔ ۴۹، شمارهٔ ۲، تابستان ۱۳۹۷ (ص ۳۲۱-۳۱۳)

ردیابی و معرفی نشانگرهای ریزماهواره در گیاه شاهدانه با کاوش در ترنسکریپتوم

ابوذر سورنی[۱]، سید علیرضا سلامی[۲*] و محمدرضا فتاحی مقدم[۳]
۱، ۲ و ۳. دانشجوی دکتری، استادیار و استاد، پردیس کشاورزی و منابع طبیعی دانشگاه تهران، کرج
(تاریخ دریافت: ۱۳۹٤/۱۲/۲۳ - تاریخ پذیرش: ۱۳۹۵/٦/۱)

چکیده

گیاه شاهدانه یکی از مهم‌ترین گیاهان ازنظر اقتصادی برای تولید دارو، غذا، فیبر و روغن است. بااین‌حال نبود نشانگرهای کافی و مؤثر ریزماهواره، گسترش بررسی‌های ژنتیکی این گیاه را محدود کرده است. در این پژوهش داده‌های ترنسکریپتوم برای شناسایی و معرفی ریزماهواره‌ها به‌منظور بررسی تنوع ژنتیکی و تفکیک رقم‌ها و جمعیت‌های فیبری و دارویی استفاده شد. بر پایهٔ این داده‌ها ۳۳۸۸ مکان ریزماهواره از ۱٤۰۲ توالی در رقم فیبری و ۱۰۳۸۱ مکان از ٤۷٤۳ توالی در رقم دارویی شناسایی شد. در میان این نشانگرها موتیف‌های سه نوکلئوتیدی و پس از آن تک و دو نوکلئوتیدی بالاترین فراوانی را نشان دادند. آغازگرهای مناسب برای افزونش ۲۳٤ ریزماهواره از ۱۵۲ توالی منحصر در رقم فیبری و ۱۵٤۳ ریزماهواره از ۱۳۷۲ توالی مختص در رقم دارویی طراحی شدند. این بررسی نخستین ارزیابی در مقیاس گسترده برای شناسایی نشانگرهای ریزماهواره از توالی‌های ترنسکریپتوم در گیاه شاهدانه است. تأثیر این نشانگرهای مولکولی می‌تواند با استفاده از جمعیت‌های موجود در ایران آزمون و ارزیابی شده و کارایی آن‌ها در اصلاح این جمعیت‌ها بررسی شود.

واژه‌های کلیدی: **بررسی‌های ژنتیکی، تیپ دارویی، ریزماهواره‌ها و فیبری، موتیف دو نوکلئوتیدی.**

Tracing and introducing SSR markers in *Cannabis sativa* by exploring in transcriptome

Abouzar Soreni[1], Seyed Alireza Salami[2*] and Mohammadreza Fattahi Moghddam[3]
1, 2, 3. Ph. D. Candidate, Assistant Professor and Professor, University College of Agriculture & Natural Resources, University of Tehran, Karaj, Iran
(Received: Mar. 13, 2016 - Accepted: Aug. 22, 2016)

ABSTRACT

Cannabis sativa L. is an important economic plant for the production of medical, food, fiber and oils. Nevertheless, lack of sufficient simple sequence repeat (SSR) markers has limited the development of cannabis genetic research. In this study, transcriptome sequences were used for identification and introduction of SSR markers in order to assess genetic diversity and separation of fiber and drug types. Based on the cannabis RNA-Seq data, 3383 SSR were identified from 1402 reads in fiber type whiles this number was 10381 for 4743 reads in drug type. Among these markers, trinucleotide repeat motifs were the most abundant, followed by mononucleotide and dinucleotide. Primers were successfully designed for amplification of 234 SSR markers from 152 individual sequences in fiber and 1543 SSR markers from 1372 individual sequences in drug using Primer3 with default parameters. This study outlines the first large-scale development of SSR markers for cannabis. The effectiveness of these molecular markers should be tested using different Iranian Marijuana populations and could be particularly useful for cannabis populations breeding.

Keywords: Dinucleotide motifs, fiber and drug types, genetic research, microsatellites.

* Corresponding author E-mail: asalami@ut.ac.ir

Archive of SID
www.SID.ir

مقدمه

نشانگرهای مولکولی از ابزار کاربردی در بررسی تغییر ژنتیکی در گیاهان هستند. در میان این نشانگرها، ریزماهواره‌ها به دلیل تکرارپذیری بالا، طبیعت چند آللی، وراثت‌پذیری هم‌بارز، فراوانی نسبی بالا و پوشش خوب ژنگان (ژنوم)، کاربردهای فراوانی در ژنتیک و اصلاح گیاهان دارند (Varshney *et al*., 2005). مهم‌ترین کاربرد نشانگرهای ریزماهواره، بررسی تنوع ژنتیکی، بررسی تبارزایی (فیلوژنی) و تکاملی، تهیهٔ نقشهٔ ژنگانی و نشانگری، همسانه کردن ژنی و نقشه‌برداری ژنتیکی، پژوهش‌های بوم‌شناختی (اکولوژیکی) و تکامل جمعیت‌ها، انتخاب نشانگرها همراه با صفت مورد نظر، خالص‌سازی مواد ژنتیکی و انجام بررسی‌های ژنتیک یاخته‌ای (سیتوژنتیکی) برای شناسایی رگه (لاین)هایی با کروموزوم‌های اضافی و رگه‌هایی با جایگزینی کروموزومی است (Naghavi *et al*., 2005). یکی از موارد کاربرد نشانگرهای ریزماهواره شناسایی نشانگرهای آگاهی‌بخش از راه تشخیص نشانگرهای ریزماهواره‌ای است که با صفات ریخت‌شناختی (مورفولوژیکی) خاصی همبستگی نشان دهند. چنین نشانگرهایی داده‌های ارزشمند ژنتیکی برای هدف‌های اصلاحی فراهم می‌آورند (Kadkhodaei *et al*., 2011).

کاربرد نشانگرهای مولکولی برای شناسایی و اصلاح شاهدانه به‌منظور کاربرد در تحقیقات جنایی، اصلاح شاهدانهٔ فیبری و یا دستیابی به رقمی خاص با ترکیب شیمیایی مطلوب به‌تازگی بررسی شده است. از آغاز بررسی‌های مولکولی مربوط به شاهدانه بیشتر مرتبط با چهار جنبهٔ فیزیولوژیکی و زیستی (بیولوژی)، متابولیت‌های ثانویهٔ فیبر و جنسیت بوده است و در نتیجه کاربرد روش‌های مولکولی روی این گیاه صرف بررسی و بهبود این صفات می‌شود. در دسترس بودن نشانگرهای مولکولی مربوط به صفات خاص که جایگاهشان روی نقشهٔ ژن پیوستگی (لینکاژی) شاهدانه مشخص باشد می‌تواند نقش برجسته‌ای بر روند اصلاح این گیاه داشته باشد، برای مثال گسترش نشانگرهای مرتبط با جنسیت موجب ایجاد روشی سریع برای تشخیص گیاهان نر و ماده در اوایل مرحلهٔ رشد شده است. بنابراین توسعهٔ نقشه‌های مولکولی اشباع‌شده برای یافتن نشانگرهای مرتبط با دیگر صفات مهم اصلاحی و نقشه‌برداری از این QTLها جز مهم‌ترین پروژه‌های آینده خواهد بود (Ranalli & Venturi, 2004).

تا پیش از سال ۲۰۱۴ بررسی گسترده‌ای درزمینهٔ شناسایی و معرفی ریزماهواره‌ها با هدف‌های ارزیابی، شناسایی و اصلاحی وجود نداشت به‌جز یکی دو مورد در مقیاس کوچک. به‌عنوان‌مثال شناسایی پانزده مکان ریزماهواره از ۴۸ نمونهٔ به‌دست‌آمده از پنج نمونهٔ بانک ژن (اکسشن) مختلف با میانگین ناخالصی (هتروزیگوتی) ۰/۶۸ نتیجهٔ کار پژوهشی Gilmore & Peakall (2002) است که توانستند میانگینی از ۱۰ آلل (۲۸-۲) را برای هر مکان ژنی (لوکوس) ردیابی کنند.

در سال ۲۰۱۴ بررسی به نسبت گسترده‌ای برای کاربرد ریزماهواره‌ها به‌منظور ارزیابی تنوع میان ۱۱۵ رقم (واریتهٔ) شاهدانه توسط Gao *et al*. (2014) انجام شد. در این بررسی از توالی‌های کوتاه بیان‌شوندهٔ (ESTs) موجود در بانک اطلاعات برای شناسایی و معرفی ریزماهواره‌های جدید استفاده شد. Gao *et al*. (2014) ۴۵۳۷ مکان ریزماهواره از ۳۶۲۴ توالی EST شناسایی کردند و با ارزیابی ۱۱۷ آغازگر، ۸۷ آغازگر را برای بررسی چندشکلی میان نمونه‌ها معرفی کردند.

شاید بهترین و معتبرترین دادهٔ موجود در بانک اطلاعات برای شناسایی و معرفی ریزماهواره‌های جدید در گیاه شاهدانه اطلاعات ترنسکریپتوم اندام‌های مختلف این گیاه در بررسی Bakel *et al*. (2012) باشد. در این بررسی DNA ژنگانی و RNA نژادگان (ژنوتیپ) دارویی Purple Kush با روش‌های خواندن کوتاه توالی‌یابی شد. یک توالی ۵۳۴ Mbp از پیش نویس (درفت) ژنگان هاپلوئید و یک ترنسکریپتوم شامل ۳ هزار ژن گزارش شد. رخداد انحصاری Δ9-تتراهیدرو کانابینولیک اسید سینتاز در ترنسکریپتوم Purple Kush و جایگزینی آن‌ها با کانابیدیولیک اسید سینتاز در Finola ممکن است توصیف کند که چرا THC در رقم دارویی تولیدشده اما در رقم فیبری این‌گونه نبوده است. توالی‌یابی دوبارهٔ رقم‌های فیبری Finola و USO-31 تفاوت ناچیزی در شمار نسخهٔ ژن

Archive of SID

www.SID.ir

آنزیم‌های مسیر کانابینوئید نشان دادند درحالی‌که تجزیهٔ SNP سطح بالای نسبی از تنوع بین چهار نوع شاهدانه را آشکار ساخت و جدایی تیپ‌های مخدر از فیبری را حمایت کرد (Bakel *et al*., 2012).

در این بررسی برای نخستین بار از داده‌های ترنسکریپتوم موجود در بانک اطلاعاتی NCBI برای شناسایی و معرفی مکان‌های ریزماهواره در گیاه شاهدانه به‌صورت مقایسه‌ای در دو تیپ فیبری و دارویی استفاده شد.

مواد و روش‌ها

داده‌های مورد استفاده

در این پژوهش از اطلاعات ترنسکریپتوم موجود در بانک اطلاعات NCBI قرار داده شده توسط Bakel *et al*. (2012) برای دو نژادگان دارویی Purple Kush و فیبری Finola به‌صورت جداگانه استفاده شد. این اطلاعات ترنسکریپتومی مربوط به اندام‌های مختلف گیاه مانند گل، ریشه و ساقه بودند که با شمارهٔ دستیابی از SRR352932 تا SRR352210 در بخش SRA بانک داده‌های قابل دسترسی هستند. این خوانش‌ها محصول پلت‌فرم Illumina Hiseq 2000 به‌صورت خوانش دو سویه ۱۰۰bp هستند. درمجموع ۳۴۰۰۲۶۶۴۰ میلیون خوانده برای نژادگان دارویی و ۱۹۰۴۷۲۶۶۵ میلیون خوانده برای نژادگان فیبری به‌منظور سرهم‌بندی توالی‌ها استفاده شد.

تجزیهٔ داده‌ها

مهار کیفیت و سرهم‌بندی توالی‌ها

پس از دانلود فایل‌های SRA مربوط به ترنسکریپتوم، فایل‌ها با استفاده از نرم‌افزار Fastq-dump v. 2.3.5 به فایل‌های fastq تبدیل شدند. پس از بررسی کیفیت توالی‌ها توسط نرم‌افزار FastQC v0.11.2، از نرم‌افزار fastq-mcf v. 1.04.807 برای ویرایش توالی‌ها با کمترین کیفیت ۳۰ و کمترین طول توالی ۵۰bp استفاده شد.

درنهایت از نرم‌افزار Trinity-v2.0.6 (Grabherr *et al*. 2011) برای سرهم‌بندی توالی‌ها استفاده شد. در آغاز فایل‌های ویرایش شده با یکدیگر ادغام شد (فایل‌های مربوط به خوانش چپ با یکدیگر و فایل‌های مربوط به خوانش راست باید با یکدیگر) تجزیهٔ سرهم‌بندی داده‌ها به‌طور جداگانه برای توالی‌های رقم فیبری و دارویی انجام شد.

بلاست نوکلئوتید برای شناسایی توالی‌های همسان در رقم فیبری و دارویی

محصول سرهم‌بندی خوانش‌ها فایلی با فرمت fasta است که کانتیگ‌ها را در خود جای داده است. ازآنجایی‌که شمار و طول خوانش‌ها، عمق و درصد پوشش ژنگانی اطلاعات برای دو نژادگان متفاوت بوده، به‌منظور ایجاد یکنواختی و پالایش (فیلتر) کردن کانتیگ‌های همسان که نقاط یکسانی از ژنگان را هدف قرار می‌دهند از راهکار بلاست کردن استفاده شد. به کمک بلاست کردن می‌توان، کانتیگ‌های همسان در هر دو نژادگان را رهگیری کرد چراکه وجود خوانش‌ها یا کانتیگ‌ها در یکی از نژادگان‌ها دلیلی بر وجود ۱۰۰ درصدی در آن نژادگان و نبود آن در نژادگان دیگر نیست (به دلیل حجم و کیفیت متفاوت اطلاعات به‌دست‌آمده از توالی‌یابی) و همین امر می‌تواند منجر به شناسایی ریزماهواره‌هایی شود که شاید به اشتباه تصور شود در یکی از نژادگان‌ها حضور دارد و در دیگری حضور ندارد. برای این منظور در آغاز یکی از فایل‌های fasta. (رقم فیبری) با استفاده از نرم‌افزار formatbd فرمت شده تا قابلیت استفاده از آن در بلاست فراهم شود. سپس از نرم‌افزار blastn برای بلاست کردن استفاده شد. فایل خروجی فایلی با ساختار جدولی است که در میان ستون‌ها، دو ستون مهم میزان همسانی و طول بلاست وجود دارد. از این دو ستون برای پلایش کردن توالی‌ها بر پایهٔ دو فراسنجهٔ (پارامتر) همسان بالای ۷۰ درصد و طولی به نسبت ۷۰ درصد طول کانتیگ استفاده شد.

شناسایی ریزماهواره‌ها

از فایل fasta به‌دست‌آمده از سرهم‌بندی توالی‌ها برای شناسایی ریزماهواره‌ها استفاده شد. در آغاز از نرم‌افزار SciRoKo (Kofler *et al*., 2007) یکی از رایج‌ترین و آسان‌ترین نرم‌افزارها درزمینهٔ شناسایی ریزماهواره‌ها استفاده شد. برای شناسایی ریزماهواره‌ها به کمک این نرم‌افزار از فراسنجه‌های پیش‌فرض استفاده شد. اما دو

Archive of SID

www.SID.ir

فراسنجهٔ مهم کمترین تکرار و کمترین طول ریزماهواره به ترتیب با مقادیر ۳ و ۱۵ تنظیم شدند. بااین‌وجود، پالایش کردن توالی‌ها بر پایهٔ طول ریزماهواره و مقایسهٔ همسانی توالی میان دو نژادگان نیازمند استفاده از نرم‌افزار MISA (MIcroSAtellite, http://www.pgrc.ipk-gatersleben.de/misa) (Aggarwal *et al.*, 2007) بود. در مرحلهٔ مقدماتی توالی‌ها با حذف دم پلی A در انتهای توالی (3'-end) پردازش شدند. یکی از فایل‌های ایجادشده توسط این نرم‌افزار پسوند misa. دارد که محتوای اطلاعاتی در مورد نوع و موقعیت هر ریزماهوارهٔ منحصربه‌فرد است.

طراحی آغازگر برای ریزماهواره‌های پیش‌بینی‌شده نیازمند دسترسی به توالی‌های اطراف (flank) ریزماهواره است. باوجود شمار بالای ریزماهواره‌ها برای آسانگری کار از اسکریپت P3_in.pl استفاده شد تا فایل مورد نیاز را برای طراحی آغازگرها با استفاده از نرم‌افزار Primer3 (http://www.fokker.wi.mit.edu/primer3) فراهم شود. ازآنجایی‌که ریزماهواره‌های پیش‌بینی‌شده منحصربه‌فرد نیستند (تفاوت تنها در چند نوکلئوتید در سمت راست یا چپ) و ممکن است در فایل misa. برای یک ID چند ریزماهواره وجود داشته باشد لازم است فایل مورد نظر بر پایهٔ IDهای واحد دسته‌بندی شود. با استفاده از راهنمای نرم‌افزار و به‌کمک اسکریپت‌های ارائه‌شده درنهایت فایل حاوی اطلاعات آغازگرها مانند توالی، موقعیت و دمای اتصال آن‌ها و اندازهٔ محصول PCR فراهم شد. از مهم‌ترین فراسنجه‌ها استفاده‌شده برای طراحی آغازگرها دمای اتصال ۶۰-۵۷ درجهٔ سلسیوس، طول آغازگر ۲۷bp-۱۸ و طول افزونش ۳۰۰bp-۲۰۰ بودند.

ژن‌های در برگیرندهٔ ریزماهواره

در آغاز برای شناسایی ژن‌ها از blastx (Gish & States, 1993) به کمک نتایج به‌دست‌آمده از سرهم‌بندی با اطلاعات موجود در بانک اطلاعاتی uniprot استفاده شد. در ادامه برای شناسایی ژن‌هایی که حاوی ریزماهواره هستند از مقایسهٔ اطلاعاتی مانند نام کانتیک استفاده شد. فایل در هم ادغام‌شدهٔ داده‌های ریزماهواره‌ها با اطلاعات بلاست به پیوست ارائه شده است (پیوست ۱).

ارزیابی شماری از آغازگرهای طراحی‌شده

برای ارزیابی آغازگرها با توجه محدودیت‌های آزمایشگاهی و فنی، هشت آغازگری را که توالی مربوط به آن همسانی بالای ۸۰ درصد را در نتایج بلاست نشان می‌داد و کدکنندهٔ ژن بودند انتخاب شدند. دو فراسنجهٔ دیگری که برای انتخاب آغازگرها استفاده شد طول ریزماهواره و شمار نوکلئوتید موتیف بود. سعی بر آن شد که آغازگرها برای ریزماهواره‌هایی انتخاب شوند که طولی در محدودهٔ ۱۶ تا ۲۲ نوکلئوتید داشته و ساختاری با موتیف دو و سه نوکلئوتیدی داشته باشند که فراوانی بیشتری داشته و اغلب کارایی بالاتری در ارزیابی‌های ژنتیکی دارند.

نمونهٔ گیاهی و استخراج دی.ان.ای

با استفاده از روش Pirttila *et al.* (2001) از برگ‌های جوان، سالم و تازهٔ ۳۹ فرد از دو نژادگان ۸۹۱۳۸۵ (دارویی) و ۹۲۱۰۱۸ (فیبری) به‌طور جداگانه دی.ان.ای استخراج شد. کمیت و کیفیت دی.ان.ای با استفاده از دستگاه نانودراپ و الکتروفورز در ژل آگارز ۱ درصد تعیین و به کمک آن‌ها غلظت یکسان از دی.ان.ای نمونه‌ها (۱۰ نانوگرم در میکرو لیتر) آماده شد.

شرایط واکنش زنجیره‌ای پلیمراز (پس.سی.آر)

واکنش زنجیره‌ای پلیمراز با استفاده از دستگاه ترموسایکلر (Bio Rad, i Cycler, USA) انجام شد. هر مخلوط واکنش پس.سی.آر شامل ۲ میکرو لیتر دی.ان.ای ژنگانی با غلظت ۱۰ نانوگرم در میکرولیتر، ۱ میکرولیتر از هر آغازگر رفت‌وبرگشت با غلظت ۱۰ پیکو مول و ۷/۵ میکرولیتر کیت پس.سی.آر با غلظت 2X از شرکت سیناژن بود که درنهایت با اضافه کردن ۳ میکرولیتر آب مقطر دو بار تقطیر سترون (استریل) حجم مخلوط واکنش پس.سی.آر به ۱۵ میکرولیتر رسانده شد. چرخه‌های گرمایی شامل ۹۴°C برای واسرشت‌سازی اولیه به مدت پنج دقیقه، شمار ۳۵ چرخه به‌صورت ۹۴°C به مدت یک دقیقه، دمای اتصال ۶۰°C به مدت ۳۰ ثانیه و ۷۲°C به مدت دو دقیقه برای افزونش قطعه‌ها و درنهایت یک چرخهٔ ۷۲°C به مدت هفت دقیقه برای تکمیل بسط انجام شد.

Archive of SID

www.SID.ir

الکتروفورز محصول پس.سی.آر

پس از انجام واکنش زنجیره‌ای پلیمراز، محصول واکنش در چاهک‌های ژل متافور ۳ درصد تهیه‌شده بر پایهٔ دستورکار شرکت سازنده (Lonza) در بافر TBE بارگیری و به مدت ۱۵۰ دقیقه و شدت ۱۱۰ ولت الکتروفورز شد. بدین‌صورت که برای برآورد طول قطعه‌های افزونش شده از نشانگر اندازه (سایز مارکر) ۱kb مربوط به شرکت فرمنتاز (Fermentas) استفاده شد. به‌منظور رنگ‌آمیزی نمونه‌ها از رنگ GelRed استفاده شد.

تجزیهٔ داده‌ها و تجزیهٔ آماری

پس از انجام مراحل آزمایشگاهی، برای بررسی مراحل چندشکلی بین نمونه‌ها به حضور یک نوار عدد یک و به نبود آن عدد صفر داده شد. پس از تشکیل ماتریس صفر و یک، ماتریس تشابه نژادگان‌ها با استفاده از نرم‌افزار NTSYSpc (Ver 2.02) و استفاده از ضریب تشابه جاکارد و به روش UPGMA محاسبه شد.

نتایج و بحث

نتایج سرهم‌بندی خوانش‌ها

درمجموع نرم‌افزار Trinity ۲۵۴۵۰ کانتیگ با کمترین طول ۲۲۴bp و بیشترین طول ۲۸۵۷bp با مجموع شمار نوکلئوتیدهای ۱۱۳۱۳۹۹۱bp برای نژادگان فیبری و ۸۵۷۷۳ کانتیگ با کمترین طول ۲۲۴bp و بیشترین طول ۱۲۱۴۳bp با مجموع شمار نوکلئوتیدهای ۳۶۳۷۷۲۷۴bp برای نژادگان دارویی خلق کرد. طول N50 به ترتیب برای نژادگان فیبری و دارویی ۴۷۸bp و ۴۲۲bp محاسبه شد. آمارهٔ N50 به‌طورمعمول برای ارزیابی نتایج سرهم‌بندی استفاده می‌شود که در اصل نوعی ارزیابی برای طول کانتیگ است. این فراسنجه برای سرهم‌بندی ژنگان اهمیت بالایی دارد چراکه یک ارزش بالایی از آن نشان می‌دهد نقاط شکستهٔ کمی در سرهم‌بندی ایجاد شده است. با این وجود N50 بزرگ به‌حتم حکایت از عملکرد بهتر برای سرهم‌بندی ترنسکریپتوم ندارد چراکه ترنسکریپتوم‌ها توالی‌های تکه‌تکه شده دی.ان.آ هستند (Kumar & Blaxter, 2010; Schliesky *et al*., 2012).

ردیابی و معرفی ریزماهواره‌ها و آغازگرها

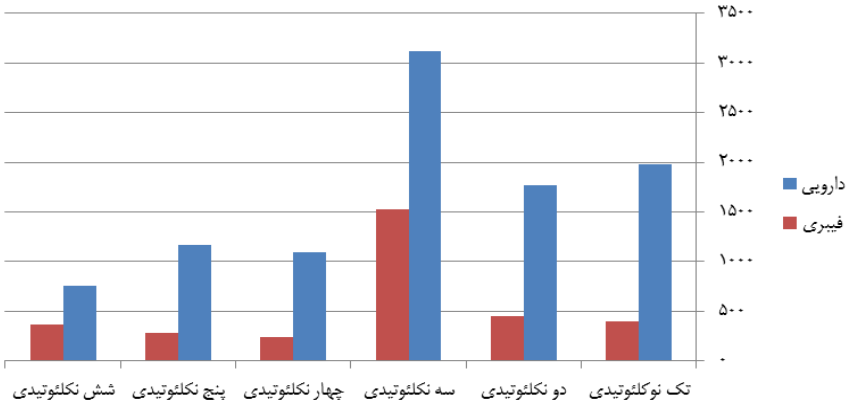
در این بررسی مکان‌های ریزماهواره در گیاه شاهدانه از توالی‌های ترنسکریپتوم ردیابی شدند. از توالی به‌دست‌آمده از سرهم‌بندی توالی‌های ترنسکریپتوم در رقم فیبری (FN)، ۲۳۲۴۲ توالی برای شناسایی ریزماهواره‌ها استفاده شد که در این میان ۱۴۰۲ توالی ۳۳۸۸ مکان ریزماهواره‌ها را در خود جای دادند درحالی‌که از ۸۵۷۷۳ توالی در رقم دارویی (PK)، ۷۵۶۰۰ توالی برای شناسایی استفاده شد که ۴۷۴۳ توالی ۱۰۳۸۱ مکان را در بر داشتند. فراوانی رخداد مکان‌ها به ازای هر ریزماهواره‌ها حدود ۳Kb برآورد شد. در رقم فیبری ۶۸ توالی دو مکان و سه توالی سه ریزماهواره را در بر می‌گرفت درحالی‌که در رقم دارویی ۳۱۴ توالی دو ریزماهواره و نه توالی بیش از سه ریزماهمراه را شامل می‌شد.

از ۲۳۲۴۲ توالی موجود در فایل سرهم‌بندی رقم فیبری ۹۱۴ توالی با همسان کمتر از ۷۰ درصد به‌عنوان توالی‌های منحصر در رقم فیبری جداسازی شدند که از میان این شمار توالی ۵۲۸ توالی توانستند ۳۴۰۱ ریزماهواره را در خود جای دهند. شمار توالی منحصربه‌فرد در رقم دارویی ۲۱۸۹۸ توالی برآورد شد که ۳۸۴۷ توالی ۱۰۴۱۸ مکان ریزماهواره را نشان دادند. بالاترین میزان موتیف در هر دو رقم فیبری و دارویی را موتیف‌های سه نوکلئوتیدی تشکیل می‌دادند (شکل‌های ۱ و ۲).

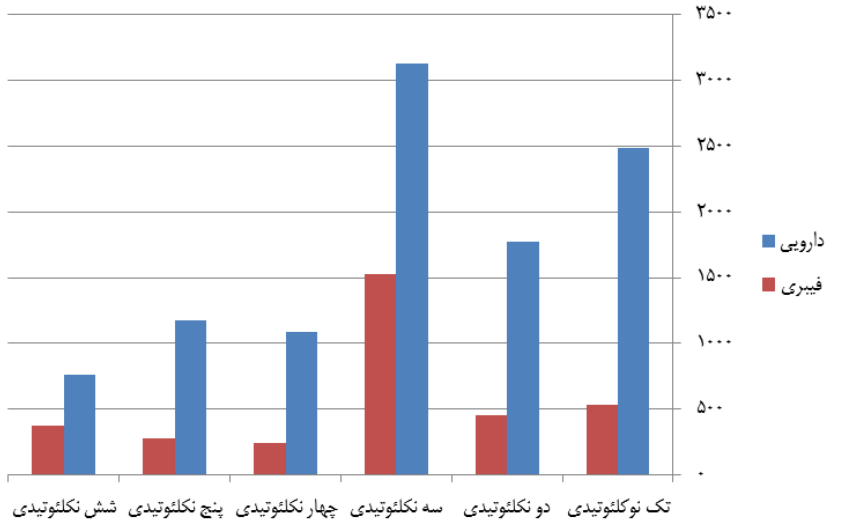
در میان موتیف‌های ترکیبی موتیف AAG-ATC و به دنبال آن AAG-ACC با فراوانی هفت و سه به ترتیب فراوان‌ترین موتیف‌ها در رقم فیبری بودند که این فراوانی برای موتیف AAG-ATC در رقم دارویی پنج بود. موتیف AT-AG با فراوانی ۲۳ بالاترین شمار موتیف ترکیبی را در رقم دارویی به خود اختصاص داد و به دنبال آن AC-AG و AT-AC به ترتیب با فراوانی نه و هشت بالاترین سطح فراوانی را داشتند. از دیگر موتیف‌های ترکیبی شناسایی‌شده در رقم دارویی می‌توان به موتیف‌های AAG-AGG، AG-AAAG، ATC-ACT، AAT-ATC، AAT-AAG و AAC-AGC با فراوانی کمتر از پنج اشاره کرد.

Archive of SID

www.SID.ir



شکل ۱. توزیع موتیف‌های ریزماهواره از همهٔ توالی‌های به‌دست‌آمده از همردیفی در رقم فیبری و دارویی گیاه شاهدانه
Figure 1. Distribution of SSR motifs was counted form assembly of total reads in fiber and drug type



شکل ۲. توزیع موتیف‌های ریزماهواره از توالی‌های منحصربه‌فرد در رقم فیبری و دارویی گیاه شاهدانه
Figure 2. Distribution of SSR motifs was counted form assembly of individual reads in fiber and drug type

به‌طور عموم موتیف‌های سه نوکلئوتیدی همه بیست نوع اسیدآمینه به‌جز والین، آرجنین، تیروزین و سیتوزین را کد می‌کنند. موتیف‌های شش نوکلئوتیدی GCCGCA، GAAGTC و CAACAG به ترتیب کدکنندهٔ آلاآلا (AlaAla)، گلووال (Glu-Val) و گلن‌گلن (GlnGln) هستند. نتایج تحقیقات بیانگر آن است که موتیف‌های سه نوکلئوتیدی توان جداسازی بالاتری دارند و کمتر احتمال دارد که منجر به نتایج مثبت کاذب شوند (Diwan & Cregan 1997).

برای شمار گسترده‌ای از این ریزماهواره‌ها به‌طور موفقیت‌آمیز برای توالی‌های منحصربه‌فرد در رقم فیبری و دارویی آغازگرهای مناسب طراحی شد. نرم‌افزار Primer3 توانست با توجه به فراسنجه‌های یادشده برای ۲۳۴ ریزماهواره از ۱۵۲ توالی در رقم فیبری آغازگر مناسب طراحی کند درحالی‌که در رقم دارویی برای ۱۵۴۳ ریزماهواره از ۱۳۷۲ توالی آغازگر طراحی شد (پیوست ۲).

غالب ترنسکریپت‌ها (۸۳٪) یک اصابت بلاست در گیاهان دیگر داشتند بااین‌حال نتایج بلاست و مقایسهٔ آن با نتایج شناسایی ریزماهواره‌ها نشان داد که از میان همهٔ ریزماهواره‌ها، ۸۱۹ ریزماهواره وجود داشت که توالی مربوط به آن‌ها کدکنندهٔ ژن بوده و در هر نژادگان حضور دارند.

www.SID.ir

ارزیابی کارایی آغازگرها

بر پایهٔ نتایج به‌دست‌آمده از توسعهٔ آغازگرها، هشت جفت آغازگر برای بررسی کارایی و ارزیابی چندشکلی و ارتباط میان ۳۹ نمونه از دو نژادگان دارویی و فیبری انتخاب شدند (جدول ۱). از هشت آغازگر طراحی‌شده شش آغازگر قادر به ایجاد چندشکلی میان نمونه‌ها بودند. شمار آلل‌ها به ازای هر مکان ژنی از ۲ تا ۵ در میان ۳۹ نمونهٔ استفاده‌شده متغیر بود. نتایج بررسی کارایی این شمار از آغازگرها سطح مناسبی از چندشکلی را در میان شماری از نژادگان‌های شاهدانه نشان می‌دهد. بااین‌حال استفاده از این آغازگرها می‌تواند با هدف مشخص در سطوح گسترده‌تر استفاده شود. نتایج خوشه‌بندی (کلاستر) UPGMA (شکل ۳) گروه‌بندی و جداسازی دو نژادگان دارویی و فیبری را در سطح تشابه ۰/۸ نشان می‌دهد. بیان این مطلب که این آغازگرها توانسته‌اند دو نژادگان فیبری و دارویی را به‌طورکامل از هم جداسازی کنند بسیار دشوار و نادر است. چراکه نخست همسانی یا ارتباط ژنتیکی فراسنجه متغیر است و به‌واسطهٔ شاخص‌های زیادی مانند شمار نمونهٔ مورد بررسی و سطح بررسی ژنگان یا شمار و نوع نشانگرها در نوسان است. از سوی دیگر جداسازی صفات کمی نیازمند ارزیابی شمار زیادی از نشانگرهاست تا بتواند پوشش مناسبی در سطح ژنگان ایجاد کند. در این صورت شانس یافتن آلل یا آلل‌هایی با فراوانی یک (که به‌ندرت رخ می‌دهد) یا نزدیک به یک‌دریک تیپ و فراوانی صفر یا نزدیک به صفر در دیگری وجود خواهد داشت. چراکه تاکنون بررسی‌های زیادی درزمینهٔ شناسایی و جداسازی تیپ‌های دارویی و فیبری در مقیاس کوچک (چند نشانگر) انجام شده است که هیچ‌یک به قادر به شناسایی و جداسازی مطلق این دو تیپ نبوده است (Howard *et al*., 2008).

اگرچه تنوع ژنتیکی با کارایی بالا و بیشترین ظرفیت و توان برای هدف‌های اصلاحی استفاده نمی‌شود بااین‌حال فرصتی عالی برای اصلاح رقم (واریته)ها یا جمعیت‌های گیاه شاهدانه برای تولید و معرفی نمونه‌های با عملکرد دارویی و یا فیبری بالا و صفات مطلوب زراعی و باغی است. ثابت شده است که نشانگرهای مولکولی استفاده منحصربه‌فردی در اصلاح گیاهان و بهبود صفات آن دارند. آغازگرهای معرفی‌شده در این بررسی از ژن‌های بیان شونده مشتق شده‌اند بنابراین برای بررسی‌های نقشه‌برداری و اصلاح سودمند هستند. یکی دیگر از برتری‌های نشانگرهای برگرفته‌شده از ترنسکریپتوم این است که آن‌ها درون ژن‌های عملکردی واقع شده‌اند که می‌توانند در میان گونه‌های گیاهی وابسته حفاظت شده باشند در نتیجه به‌صورت انتقالی استفاده شوند. تأثیر این نشانگرهای مولکولی می‌تواند با استفاده از جمعیت‌های موجود در ایران آزمون و ارزیابی شده و کارایی آن‌ها در اصلاح این جمعیت‌ها بررسی شود. بسیاری از پرورش‌دهندگان گیاه شاهدانه از اصلاح برای تولید رقمی با محتوای THC بالا استفاده می‌کنند (Coyle *et al*., 2003).

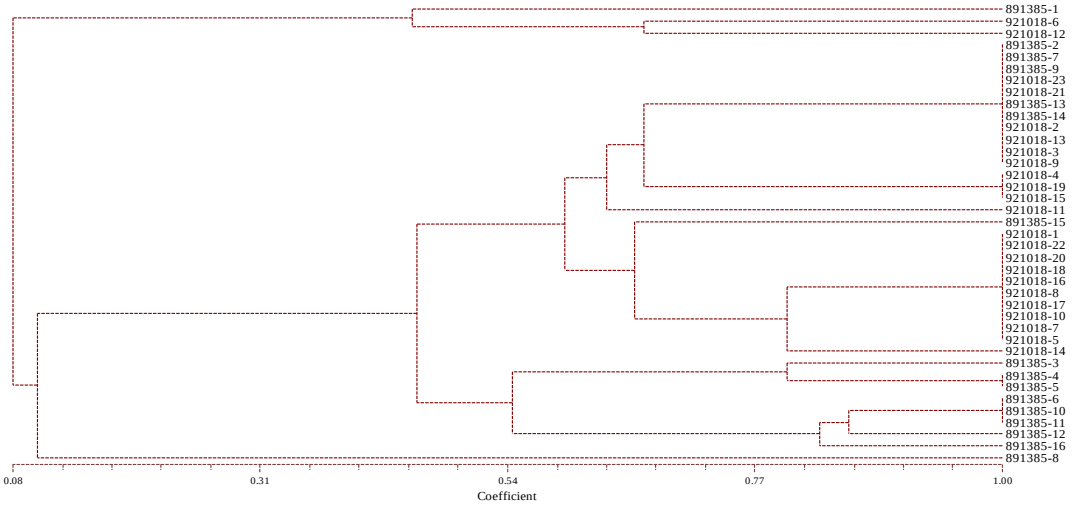
جدول ۱. توالی و درصد چندشکلی آغازگرهای مورد استفاده

Table 1. Sequence and percentage of polymorphism in used primers

آغازگر	توالی آغازگر	دمای اتصال (°C)	شمار کل قطعات افزونش‌شده	شمار قطعات چند شکل	درصد چند شکلی
comp12718-F	5'TGGCTTTTGGGCTGAGATAG3'	۶۰	۱	۰	۰
comp12718-R	5'CCACAGACCGGATCATTAGG3'				
comp22029-F	5'TCGAATATGCAAGAAAAATGGAGG3'	۶۰	۳	۲	٪۶۶
comp22029-R	5'CACACCTTGCTATGCCTGTC3'				
comp29550-F	5'AAAGGAAGTCCTGGGAAAACA3'	۶۰	۲	۱	٪۵۰
comp29550-R	5'TCGAGTTCGACTCCCAATCT3'				
comp34678-F	5'CCAAACCCCAACAACACACT3'	۶۰	۲	۱	٪۵۰
comp34678-R	5'TTGGCCTTCGAAGAAGAGAA3'				
comp37083-F	5'GGTTGAGTCGGTGGAAACTG3'	۶۰	۴	۳	٪۷۵
comp37083-R	5'CTCCCCACCCCTACTACCAT3'				
comp40091-F	5'ACGTTCCAGTAGGAGGTGGTT3'	۶۰	۲	۱	٪۵۰
comp40091-R	5'GATCTTGGCCAGTACCTGCT3'				
comp48241-F	5'GGCGGTAAATAGTGGTGCTG3'	۶۰	۲	۱	٪۵۰
comp48241-R	5'CGCAAGCAATTTTAATGGCGG3'				
comp101781-F	5'AGCCATGGCCAGTGACATA3'	۶۰	۱	۰	۰
comp101781-R	5'GCGTGTTTGTGTTTGAGGAG3'				

Archive of SID

www.SID.ir



شکل ۳. گروه‌بندی نژادگان‌های شاهدانهٔ مورد بررسی بر پایهٔ ماتریس تشابه به روش UPGMA و با استفاده از ضریب تشابه جاکارد

Figure 3. UPGMA dendrogram for studied samples of *cannabis sativa* based on Jaccard's similarity

اصلاح بر مبنای انتخاب گیاهانی با محتوای بالای THC می‌تواند به دو روش به‌دست آید. روش اول استفاده از ریزازدیادی نمونه‌های گیاهی با محتوای بالای THC است که در این روش گیاه مادری و گیاهان به‌دست‌آمدهٔ رخ‌نمای (پروفایل) ژنتیکی یکسانی دارند. روش دوم تلاقی و تولید بذر است که در این حالت هر گیاه به‌دست‌آمده از بذر رخ‌نمای ژنتیکی منحصربه‌فردی دارد. نشانگرهای توسعه داده شده در این بررسی می‌تواند برای بررسی ارتباط ژنتیکی میان این نتاج به‌کار رود.

REFERENCES

1. Aggarwal, R. K., Hendre, P. S., Varshney, R. K., Bhat, P. R., Krishnakumar, V. & Singh, L. (2007). Identification, characterization and utilization of EST-derived genic microsatellite markers for genome analyses of coffee and related species. *Theoretical and Applied Genetics*, 114, 359-372.
2. Coyle, H. M., Shutler, G., Abrams, S., Hanniman, J., Neylon, S., Ladd, C., Palmbach, T. & Lee, H. C. (2003). A simple DNA extraction method for marijuana sample used in amplified fragment length polymorphism (AFLP) analysis. *Journal of Forensic Sciences*, 48(2), 343-347.
3. Diwan, N. & Cregan, P. (1997). Automated sizing of fluorescent-labelled simple sequence repeat (SSR) markers to assay genetic variation in soybean. *Theoretical and Applied Genetics*, 95, 723-733.
4. Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. & Regev, A. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, 29, 644-652.
5. Gilmore, S. & Peakall R. (2003). Isolation of microsatellite markers in Cannabis sativa L. (marijuana). *Molecular Ecology Notes*, 3(1), 105-107
6. Gao, Ch., Xin, P., Cheng Ch., Tang, Q., Chen, P., Wang, Ch., Zang, G. & Zhao, L. (2014). Diversity Analysis in Cannabis sativa based on Large-Scale Development of Expressed Sequence Tag-Derived Simple Sequence Repeat Markers. *PLoS One*, 9(10), e110638.
7. Gish, W. & States, D. J. (1993). Identification of protein coding regions by database similarity search. *Nature Genet*, 3, 266-272.
8. Howard, C., Gilmore, S., Robertson, J. & Peakall, R. (2008). Application of new DNA markers for forensic examination of Cannabis sativa seizures. *National Drug Law Enforcement Research Fund*, 61 p.
9. Kadkhodaei, S., Nekouei, N. K., Shahnazari, M., Etminani, H., Imani, A., Ghaderi-Zefrehei, M., Elahy, M. & Ariff, B. (2011). Molecular tagging of agronomic traits using simple sequence repeats: Informative markers for almond ('*Prunus dulcis*') molecular breeding. *Australian Journal of Crop Science*, 5(10), 1199-1209.
10. Kofler, R., Schlötterer, C. & Lelley, T. (2007). SciRoKo: a new tool for whole genome microsatellite search and investigation. *Bioinformatics*, 13, 1683-1685.

Archive of SID

www.SID.ir

11. Kumar, S. & Blaxter, M. (2010) Comparing de novo assemblers for 454 transcriptome data. *BMC Genomics*, 11, 571.
12. Schliesky, S., Gowik, U., Weber, A. P. & Bräutigam, A. (2012). RNA-seq assembly–are we there yet?. *Frontiers in Plant Science*, 3, 220.
13. Naghavi, M. R., Ghareyazi, B. & Hosseini, Gh. (2005). Molecular markers. Tehran university press. Tehran. Iran. pp: 88-100.
14. Ranalli, P. & Venturi, G. (2004). Hemp as a raw material for industrial applications. *Euphytica*, 140(1-2), 1-6.
15. Varshney, R. K., .Graner, A. & Sorrells, M. E. (2005). Genic microsatellite markers in plants: features and applications. *Trends Biotechnol*, 23(1), 48-55.

Archive of SID

www.SID.ir