ژنتیک نوین
دوره هشتم، شماره ۲، تابستان ۱۳۹۲
صفحه ۱۳۰-۱۱۷

ابداع نشانگرهای مبتنی بر توالی‌های بیان شده (ESTs) و غربال ژنتیکی برای چندشکلی تک نوکلئوتیدی (SNP) در ژنوم گیاهان

EST-based marker discovery and SNP genotyping in plants genome

فهیمه شاهین نیا *[۱]، بدرالدین ابراهیم سید طباطبایی [۲]
۱- عضو هیات علمی موسسه تحقیقات ژنومیکس گیاهی (ACPFG)، دانشگاه آدلاید
۲- استاد دانشکده کشاورزی دا نشگاه صنعتی اصفهان

Shahinnia F[*1], Sayed Tabatabaei BE[2]
1. Australian Center for Plant Functional Genomics, School of Agriculture, University of Adelaide, SA 5064, Australia
2. Professor, College of Agriculture, Isfahan University of Technology, Isfahan

* نویسنده مسئول مکاتبات، پست الکترونیکی: fahimeh.shahinnia@acpfg.com.au
(تاریخ دریافت: ۹۱/۱۱/۲۸- تاریخ پذیرش: ۹۲/۲/۳)

چکیده

در اصلاح نباتات سنتی، تنوع ژنتیکی اغلب بر اساس صفات ظاهری مشخص می‌شود. اخیرا، با پیشرفت نشانگرهای مولکولی در بسیاری از محصولات زراعی، شناسایی تنوع ژنتیکی در سطح مولکولی و بر اساس تفاوت موجود در DNA و همچنین اثر آن بر صفات فنوتیپی امکان‌پذیر است. نشانگرهای مولکولی وابسته به DNA از جمله ابزارهای قدرتمند تشخیصی هستند که نه تنها برای شناسایی چندشکلی در سطح ژنوم بلکه در رابطه با مکان‌های ژنتیکی خاص نیز بکار می‌روند. در گذشته چنین نشانگرهایی با استفاده از کتابخانه‌های ژنومی (RFLPs و SSRs) یا از طریق تکثیر تصادفی PCR از DNA ژنومی (RAPDs) و یا هر دو روش (AFLPs) تهیه می‌شد. اخیرا دسترسی به DNA ژنومی و توالی‌های cDNA (ESTs) در بانک‌های اطلاعاتی عمومی سبب شده است تا ابداع نشانگرها به طور مستقیم و مقرون به صرفه انجام شود. از جمله کاربردهای عمومی این نشانگرها در بخش کشاورزی می‌توان به غربال مولکولی ژرم پلاسم، تهیه نقشه ژنی، مکان‌یابی ژنی صفات کمی و همچنین انتخاب به کمک نشانگر (MAS) اشاره کرد. در این مقاله، پیشرفت‌های حاصل در ابداع نشانگر و غربال ژنتیکی برای چندشکلی تک نوکلئوتیدی (SNP) با استفاده از نشانگرهای مبتنی بر توالی‌های بیان شده (ESTs) مورد بحث قرار گرفته است.

واژه‌های کلیدی

انتخاب به کمک نشانگر (MAS)، نشانگرهای مبتنی بر توالی های بیان شده (ESTs)، غربال ژنتیکی (SNPs)

Archive of SID

مقدمه

نشانگرهای مولکولی از جمله عناصر زیستی هستند که به عنوان کاوشگرهای آزمایشگاهی برای یافتن و مشخص کردن یک فرد، بافت، سلول، هسته سلولی، کروموزوم یا ژن بکار می‌روند و به صورت اشکال مختلف آللی معرفی می‌شوند. در ژنتیک کلاسیک، چندشکلی بر اساس تنوع در همردیف‌های ژنی است در حالیکه در ژنتیک نوین، چندشکلی به عنوان تنوع در هر مکان ژنی روی ژنوم معرفی می‌شود (Xu 2010).

بیشتر نشانگرهای مولکولی، تنها تنوع موجود در DNA را بروز می‌دهند و به منظور تسهیل مطالعات توارثی مورد استفاده قرار می‌گیرند. نشانگرهای ژنتیکی مناسب دارای مشخصات زیر هستند: ۱) فراوانی چندشکلی ژنتیکی، ۲) ویژگی همبارزی جهت تشخیص هتروزیگوت‌ها، ۳) ترسیم سیمای آللی به طوری که آلل‌های مختلف را بتوان تشخیص داد، ۴) توزیع مناسب روی ژنوم، ۵) عدم اثر پلیوتروپی، ۶) شناسایی آسان و امکان اجرا به طور اتوماتیک، ۷) هزینه کم در تهیه نشانگر و غربالگری ژنتیکی، ۸) تکرارپذیری فراوان و امکان جمع آوری داده‌ها در آزمایشگاه‌های مختلف (Xu 2010). نشانگرهای مولکولی را می‌توان بر اساس روش‌های مولکولی به پنج دسته اصلی تقسیم کرد (Vershney 2010; Xu 2010):

۱) نشانگرهای مبتنی بر لکه‌گذاری ساترن شامل RFLP[1]، SSCP-RFLP[2] و DGGE-RFLP[3]؛

۲) نشانگرهای مبتنی بر PCR شامل RAPD[4]، SSR[5]، STS[6]، SCAR[7]، RP-PCR[8]، AP-PCR[9]، OP-[10]، PCR، SSCP-PCR[11]، SODA[12]، DAF[13]، AFLP[14]،

۳) نشانگرهای مبتنی بر توالی‌های تکراری شامل ماهواره‌ها[18]، ماهوارک‌ها[19] و ریزماهواره‌ها[20] که مورد اخیر با اسامی مختلفی همچون SSR، SSLP[21]، SRS[22] و TRS[23] معرفی می‌شود؛

۴) نشانگرهای مبتنی بر mRNA شامل DD[24]، RT-PCR[25]، DDRT-PCR[26]، RDA[27]، EST[28]، STS و SAGE[29]؛

۵) نشانگرهای مبتنی بر توالی‌یابی یا ریزآرایه‌ای شامل SNP[30]، DArTs[31] و SFP[32].

نشانگرهای مولکولی بویژه آنهایی که متعلق به دو گروه اول هستند، در بیشتر گیاهان زراعی تهیه شده‌اند. این نشانگرها، DNA ژنومی را بدون توجه به مناطق بیان شونده یا غیر بیان شونده و همچنین بدون توجه به فعالیت بخش‌های ژنومی، مورد تفحص قرار می‌دهند. با این حال طی سال‌های اخیر، امکان تهیه نشانگرهای مبتنی بر مناطق شناخته شده ژنوم یا ژن‌ها فراهم شده است که به عنوان نشانگرهای کارکردی (Anderson and Lübberstedt 2003; Gupta and Rustgi 2004)، نشانگرهای مولکولی وابسته به ژن (Varshney et al. 2007) یا نشانگرهای مبتنی بر بیان ژن (West et al. 2006) خوانده می‌شوند. تا چند سال پیش، گسترش نشانگرهای وابسته به ژن محدود به برخی از گونه های زراعی می‌شد ولی معرفی فناوری‌های نسل جدید توالی‌یابی سبب گسترش نشانگرهای وابسته به ژن حتی در گونه‌های زراعی کمتر شناخته‌شده‌ای شد که فاقد اطلاعات ژنومی

[1] Restriction fragment length polymorphism
[2] Single strand conformation polymorphic RFLP
[3] Denaturing gradient gel electrophoresis RFLP
[4] Randomly amplified polymorphic DNA
[5] Simple sequence repeat
[6] Sequence tagged site
[7] Sequence characterized amplified region
[8] Random primer-PCR
[9] Arbitrary primer-PCR
[10] Oligo primer-PCR
[11] Single strand conformation polymorphism-PCR
[12] Small oligo DNA analysis
[13] DNA amplification fingerprinting
[14] Amplified fragment length polymorphism
[15] Sequence-related amplified polymorphism
[16] Target region amplified polymorphism
[17] Insertion/deletion polymorphism
[18] Satellite DNA
[19] Mini-satellite DNA
[20] Microsatellite DNA
[21] Simple sequence length polymorphism
[22] Short repeat sequence
[23] Tandem repeat sequence
[24] Differential display
[25] Reverse transcription PCR
[26] Differential display reverse transcription PCR
[27] Representational difference analysis
[28] Expression sequence tags
[29] Serial analysis of gene expression
[30] Single nucleotide polymorphism
[31] Diversity array technology
[32] Single feature polymorphism

Archive of SID
www.SID.ir

بودند (Varshney et al. 2009; Varshney 2010). در این نوشتار مروری، پیشرفت‌های حاصل در نشانگرهای مبتنی بر mRNA یا EST و همچنین کاربرد آن در غربال ژنتیکی بر اساس SNP و اصلاح مولکولی به منظور بهبود محصولات زراعی ارائه می‌شود.

معرفی توالی‌های نشانمند بیان شده (ESTs)

زمانی که توالی ژنومی یک موجود قابل دسترسی باشد، مجموعه توالی‌های cDNA وسیله مناسبی برای تشخیص ژن‌ها در توالی DNA خواهند بود. صدها جفت باز حاصل از توالی انتهای cDNA یا توالی قطعات کوچک همسانه‌سازی شده mRNA به عنوان EST شناخته می‌شود. طول چنین قطعاتی که با استفاده از روش های توالی‌یابی حاصل می‌شوند در حدود ۵۰۰ تا ۸۰۰ نوکلئوتید است. با توجه به اینکه این قطعات DNA همسانه‌سازی شده مکمل mRNA هستند، لذا ESTها را می‌توان بخشی از ژن‌های بیان شده دانست. بطور کلی مشخصات این قطعات به عنوان توالی cDNA/mRNA یا مکمل معکوس mRNA در بانک‌های اطلاعاتی ثبت می‌شود (Adams et al. 1991; Pontius et al. 2002).

فناوری توالی‌یابی cDNA با بازدهی بالا از سال ۱۹۹۱ شروع شد. در این فناوری انتخاب تصادفی همسانه‌های cDNA و توالی‌یابی در یک یا دو انتهای قطعات صورت می‌گیرد. بر این اساس واژه EST به گروه جدیدی از قطعات توالی‌یابی شده کوتاه و نسبتا دقیق (کمتر از دو درصد خطا) اطلاق می‌شود. نیاز به انجام تنها یکبار توالی‌یابی قطعات، از ویژگی‌های بارز و ارزشمند در معرفی این گروه از نشانگرهاست. در بیشتر موارد، نیازی به تشخیص و شناسایی همسانه‌ها وجود ندارد، بلکه تنها با استفاده از مقایسه بخش کوچکی از داده‌های توالی‌یابی شده با توالی‌های ژن‌های شناخته شده یا ESTهای دیگر، می‌توان اقدام به شناسایی همسانه نمود. همچنین بسیاری از همسانه‌ها به عنوان بخشی از نمونه‌های شناخته شده قبلی بوده و تنها معدودی از آنها حاوی تفاوت‌های بارزی می‌باشند. توجه به هزینه توالی‌یابی با کیفیت مناسب به منظور تایید همسانه‌ها نیز از موارد قابل تامل است (Pontius et al. 2002).

پس از بررسی‌های اولیه در مورد هزینه‌های تهیه EST و میزان کاربرد آن در ژنوم انسان، پروژه‌های مشابه بسیاری انجام و تعداد متنابهی EST در انسان معرفی شدند (Okubo et al. 1992). علاوه بر آن، پروژه‌های گسترده دیگری در مورد تهیه EST در موجودات دیگر نیز آغاز شد. در سال ۱۹۹۲، بانک اطلاعاتی dbEST به منظور جمع‌آوری ESTها ایجاد شد (Boguski et al. 1993; Boguski et al. 1994) ولی بعدا به صورت بخش EST در بانک اطلاعاتی GenBank تثبیت شد (Benson et al. 2002; Pontius et al. 2002). بخش EST بانک اطلاعاتی GenBank حاوی حدود دو سوم تمامی ESTهای ثبت شده است. بیشترین ESTهای ثبت شده در این بانک اطلاعات (http://www.ncbi.nlm.nih.gov/dbEST/dbEST_summary.html) تا ماه ژانویه ۲۰۱۳ حدود ۷۴۱۸۶۶۹۲ توالی نشانمند بیان شده می باشد (جدول ۱).

جدول ۱- بیشترین ESTهای ثبت شده بانک اطلاعاتی GenBank در ۱۳ ژنوم مختلف جانوری و گیاهی تا ماه ژانویه ۲۰۱۳

نام علمی	تعداد ESTهای ثبت شده
Homo sapiens (human)	۸۷۰۴۷۹۰
Mus musculus + *domesticus* (mouse)	۴۸۵۳۵۷۰
Zea mays (maize)	۲۰۱۹۱۳۷
Sus scrofa (pig)	۱۶۶۹۳۳۷
Bos taurus (cattle)	۱۵۵۹۴۹۳
Arabidopsis thaliana (thale cress)	۱۵۲۹۷۰۰
Danio rerio (zebrafish)	۱۴۸۸۲۷۵
Glycine max (soybean)	۱۴۶۱۷۲۲
Triticum aestivum (wheat)	۱۲۸۶۳۷۲
Xenopus (Silurana) *tropicalis* (western clawed frog)	۱۲۷۱۴۸۰
Oryza sativa (rice)	۱۲۵۳۵۵۷
Ciona intestinalis	۱۲۰۵۶۷۴
Rattus norvegicus + sp. (rat)	۱۱۶۲۱۳۶

الف) جستجو برای یافتن توالی EST مورد نظر

تجزیه‌وتحلیل داده‌های EST می‌تواند سبب تسهیل شناسایی ژن و تعیین ساختار آن، تفسیر کل ژنوم، شناسایی چند شکلی‌های تک نوکلئوتیدی (SNPs) و همچنین بررسی ترانسکریپتوم و پروتئوم شود. کشف ژن‌های فعال بر اساس توالی‌های EST نیازمند تلاش بسیاری است. توالی EST ممکن است حاوی قطعات ساده، شیمرهای فراوان، توالی‌های تکراری و زائد همچون قطعات

Archive of SID
www.SID.ir

آداپتور و بخشی از توالی‌های ناقل استفاده شده باشند که بایستی قبل از تجزیه‌وتحلیل مشخص شوند. با توجه به اینکه توالی‌های EST قطعات بریده شده از cDNA هستند، لذا بایستی این قطعات به هم متصل شوند تا توالی رونوشت mRNA و به دنبال آن ژن رمز کننده مشخص شود؛ ولی به دلیل وجود توالی‌های شیمری که موجب ایجاد پیچیدگی در مرتب‌سازی قطعات مختلف می‌شود، بازسازی رونوشت با مشکل مواجه شده و به سهولت امکان‌پذیر نیست. در واقع توالی‌های شیمری حاصل سبب گمراهی در اتصال قطعات و تفسیر ژن‌های غیر صحیح می‌شود. در نتیجه حذف EST شیمری برای پی بردن به رونوشت‌های صحیح بر اساس توالی‌های EST ضروری است (Lee et al. 2007). با توجه به اینکه احتمال خطا در مرتب‌سازی قطعات ESTها به شدت زیاد است، لذا استفاده از روش‌های کامپیوتری به منظور افزایش دقت در آنالیز، گردآوری فراگیر، دسته‌بندی و تفسیر توام ESTها ضروری می باشد (Nagaraj et al. 2007). تجزیه و تحلیل مجموعه داده‌های EST توسط نرم‌افزارهای مختلفی انجام می‌شود که از آن جمله می‌توان به ESTpass (Lee et al. 2007)، ESTAP (Mao et al. 2003) و ESTAnnotator (Hotz-Wagenblatt 2003) اشاره کرد. لازم به ذکر است که بیشتر نرم‌افزارها تنها تفسیر ساده‌ای را ارائه می‌کنند و عمدتا به‌طور محدود در دسترس بوده و دائما با مدل‌های جدید جایگزین می‌شوند و نیاز به روزآوری می‌باشند. به این دلایل برخی خدمات شبکه‌ای مانند PESTAS برای تجزیه و تحلیل ESTها و توالی‌یابی ارائه شده‌اند (Nam et al. 2009).

ب) منابع مرتبط با داده‌های EST

همانطور که ذکر شد، عظیم‌ترین مجموعه مربوط به داده‌های EST در بخش dbEST در مرکز ملی اطلاعات بیوتکنولوژی ایالات متحده امریکا[1] (NCBI) به آدرس http://www.ncbi.nlm.nih.gov/dbEST/dbEST/ و شامل ۷۴۱۸۶۶۹۲ داده EST (تا ماه ژانویه ۲۰۱۳) از موجودات مختلف است. همچنین بخش UniGene در بانک اطلاعاتی NCBI ژن‌های بی‌همتا (UniGene) را ذخیره‌سازی می‌کند و حاوی گروهی از ژن‌های مشخص غیر تکراری حاصل توالی‌های EST می‌باشد (Boguski and Schuler 1995; Pontius et al. 2002). برخی کلکسیون‌های اختصاصی داده‌های EST نیز وجود دارند که از جمله می‌توان به موسسه سرطان دانا فاربر[2] (www.dana-farber.org/)، شاخصه‌های ژن گیاهی[3] (http://compbio.dfci.harvard.edu/tgi/tgipage.html) (Lee at al. 2005; Nagaraj et al. 2006)، پروژه داده‌های EST موش صحرایی[4] در دانشگاه آیوا ایالات متحده امریکا (https://genome.uiowa.edu/home.old/index.html) و پروژه آناتومی ژنوم سرطان[5] (http://cgap.nci.nih.gov/) اشاره کرد (Boguski et al. 1993; Boguski and Schuler 1995; Nagaraj et al. 2006).

ج) تجزیه و تحلیل توالی‌های EST

داده‌های خام EST، حاوی اطلاعات زیستی ناچیزی هستند. تجزیه و تحلیل این داده‌ها با استفاده از ابزارهای محاسباتی ویژه سبب رفع این نقیصه شده و نتایج حاصل از بررسی صحیح انبوهی از توالی‌های EST می‌تواند در نهایت منجر به شناسایی ترانسکریپتوم موجود مورد مطالعه شود. مراحل معمول در تجزیه‌وتحلیل داده‌های EST به شرح زیر می‌باشد (Nagaraj et al. 2006):

۱) حذف ناخالصی ناقل، بخش‌های تکراری و همچنین توالی‌های بسیار کوتاه و فاقد کیفیت مناسب در توالی‌های EST

۲) دسته‌بندی و اتصال ESTها به منظور تهیه توالی‌های مورد توافق (رونوشت‌های فرضی)

۳) جستجو در بانک اطلاعاتی DNA برای یافتن توالی‌های همخوان و برآورد فعالیت بالقوه آن

۴) یافتن پپتیدهای فرضی بر اساس نتایج حاصل از ترجمه توالی‌های مورد توافق

۵) جستجو در بانک اطلاعاتی پروتئین به منظور مشخص کردن عملکرد فرضی آن

۶) تجزیه‌وتحلیل‌های تکمیلی جهت تفسیر عملکرد EST و رویت‌سازی نتایج

[1] National Center for Biotechnology Information

[2] Dana Farber Cancer Institute
[3] Plant Gene Indices
[4] Rat EST Project
[5] Cancer Genome Anatomy Project

Archive of SID
www.SID.ir

د) محاسبات مقدماتی توالی های EST

سلسله محاسبات مقدماتی سبب کاهش خطاهای موجود در داده‌های EST شده و افزایش کارایی تجزیه‌وتحلیل‌های بعدی را بدنبال دارد. ناخالصی‌های حاصل از ناقل در توالی‌یابی بایستی قبل از دسته‌بندی ESTها حذف شوند، به این صورت که می‌توان با استفاده از ابزارهایی همچون BLAST یا cross-map به مقایسه داده‌های EST با داده‌های موجود در بانک‌های اطلاعاتی ناقلین (مانند UniVec و EMVEC) پرداخت و پس از شناسایی توالی‌های ناقل در داده‌های EST، اقدام به حذف آنها نمود. نواحی تکراری در داده‌های EST را می‌توان توسط روش‌هایی مانند DUST (در NCBI) یا nseg شناسایی کرد. عناصر تکرار شونده مانند LINEها[1]، SINEها[2]، LTRها[3] و SSRها با استفاده از RepeatMasker یا MaskerAid قابل شناسایی هستند (Bedell et al. 2000; Nagaraj et al. 2006).

ه) گروه‌بندی و مرتب‌سازی توالی‌های EST

ضروری است تا ESTهای دارای کیفیت مناسب بر مبنای تشابه توالی، دسته‌بندی شده و به نحوی در کنار هم قرار گیرند که بیشترین شباهت توالی مورد توافق حاصل شود. این مرحله با استفاده از تعداد زیادی از توالی‌های کوتاه به منظور افزایش طول توالی و سازمان‌دهی بهتر ESTها انجام می‌گیرد. در این رابطه بیشترین برنامه‌های کاربردی، برنامه‌های Phrap و CAP3 می‌باشند (Nagaraj et al. 2006). از برنامه PEACE خصوصا برای دسته‌بندی و مرتب کردن قطعات توالی‌یابی شده حاصل از روش‌های توالی‌یابی NGS [4] استفاده می‌شود. این برنامه در شبکه اطلاعاتی http://www.peace-tools.org قابل دسترسی است.

و) پیش‌بینی ORF و ترجمه EST

پیش‌بینی نواحی کد کننده پروتئین بر اساس توالی‌های توافقی EST سبب ارتقای روند شناسایی نسخه‌برداری و ترجمه در ژن می‌شود. داده‌های EST می‌تواند با داده‌های پروتئین‌های شناخته شده مقایسه شوند و از این طریق عملکرد محصول حاصل از ترجمه ژن نسخه‌برداری شده مشخص شود (Nagaraj et al. 2006).

ز) فرایند ذخیره‌سازی اطلاعات حاصل از توالی‌یابی EST

در پروژه‌های گسترده توالی‌یابی ژنوم که هزاران EST به‌طور روزانه تولید می‌شوند، داده‌های تولید شده بایستی در یک فرایند خودکار ذخیره، دسته‌بندی و قابل تفسیر شوند. بنابراین بایستی برای انتقال داده‌ها از دستگاه توالی‌یاب به بانک اطلاعات یک دستورالعمل کارا، اتوماتیک و دقیق تهیه و تنظیم شود. در چنین فرایندی، لازم است تا ESTها به‌طور خودکار از بانک داده‌ها یا کروماتوگرام‌های حاصل از دستگاه توالی‌یاب دریافت شوند. پس از حذف توالی‌های نامناسب، داده‌های مطلوب گروه‌بندی شده و در کنار یکدیگر قرار داده می‌شوند تا توالی نهایی حاصل شود. در نهایت این توالی‌ها ترجمه شده و عملکرد آنها بر اساس مقایسه با DNAهای مختلف و پروتئین‌های مشابه مشخص می‌شود (Nagaraj et al. 2006). برخی از ابزارهای مناسب برای این منظور عبارتند از: ۱) ESTAP (http://staff.vbi.vt.edu/estap/)، ۲) PartiGene (http://www.nematodes.org/bioinformatics/PartiGene/)، ۳) ParPEST (http://www.cab.unina.it)، ۴) EGassembler (http://egassembler.hgc.jp/)، ۵) ESTWeb (http://bioinfo.iq.usp.br/estweb/) و ۶) ESTIMA (http://titan.biotec.uiuc.edu/ESTIMA/).

ح) تفسیرهای عملکردی

به محض اینکه پلی‌پپتید فرضی حاصل شد، فعالیت آن بر مبنای توالی پروتئین‌های اصلی شناخته شده، موتیف‌ها و پروتئین‌های هم‌خانواده با استفاده از یک ابزار ترکیبی حاصل از مجموعه اطلاعات مانند InterProScan (Zdobnov and Apweiler 2001) پیش‌بینی می‌شود. توالی‌های پروتئینی الگوهای مناسب‌تری در تفسیر عملکردی هستند، به‌طوری که در تهیه همردیف‌سازی چندگانه، تجزیه و تحلیل چند ژنی، تهیه کتابخانه‌های انگشت‌نگاری پروتئینی به منظور کاربردهای پروتئومیکس مورد استفاده قرار می‌گیرند. در این خصوص ابزارهای Pfam و SMART برای تجزیه‌وتحلیل دُمین و موتیف و همچنین ارائه الگوی تفسیر عملکردی مناسب توالی‌ها، از کارایی بالایی برخوردارند (Nagaraj et al. 2006). حصول چنین تفسیرهایی

[1] Long Interspersed Nucleotide Elements
[2] Short Interspersed Nucleotide Elements
[3] Long Terminal Repeats
[4] Next Generation or Sanger Sequencing

Archive of SID
www.SID.ir

با استفاده از ابزارهایی مانند ۱) Gene Ontology (http://www.geneontology.org/)، ۲) InterProScan (http://www.ebi.ac.uk/InterProScan/)، ۳) KEGG (http://www.genome.jp/kegg/)، ۴) BLAST2GO (www.blast2go.org/) و ۵) Generic EST Analysis (http://biolinfo.org/EST/) نیز امکان‌پذیر می‌باشد.

در پایان این بخش اشاره می شود که بطور خلاصه، بیشتر برنامه‌های مناسب برای فرایندسازی مقدماتی EST، چیدمان و دسته‌بندی، سازگاری داده‌ها و تفسیر عملکردی در سایت http://biolinfo.org/EST قابل دسترسی است (Nagaraj et al. 2006). علاوه بر این، ESTExplorer به عنوان مجموعه‌ای کامل از تجزیه‌وتحلیل EST به منظور تفسیر نوکلئوتیدی و پروتئینی، یکی از برنامه‌هایی است که توسط بسیاری از کاربران مورد استفاده قرار می‌گیرد (Nagaraj et al. 2007).

شناسایی نشانگرهای مبتنی بر EST

مشخصات توالی DNA بسیاری از ژن‌ها به‌طور کامل شناسایی و توالی کامل همسانه‌های cDNA آنها در مورد بسیاری از گونه‌های گیاهی در بانک داده‌ها قرار داده شده است، به طوری که امکان دریافت توالی داده‌های EST، ژن‌ها و همسانه‌های cDNA از بانک ژن (GenBank) وجود داشته و می‌توان از آنها برای تعیین نشانگرهای مبتنی بر EST استفاده کرد.

الف) نشانگرهای EST-SSR

آغازگرهای حاصل از نواحی مجاور موتیف‌های موجود در بخش ژنی می‌تواند به منظور تکثیر بخش‌های ریز ماهواره‌ای ژن‌ها مورد استفاده قرار گیرند. به عنوان مثال در ذرت، نشانگرهای SSR بدین روش تهیه شدند و توالی آغازگر آنها را می‌توان از آدرس www.maizeGDB.org دریافت کرد. نشانگرهای SSR وابسته به ژن نسبت به نشانگرهای SSR ژنومی دارای مزایای بسیاری هستند به نحوی که می‌توان آنها را با استفاده از روش‌های زیست‌رایانه‌ای سریعا در نواحی بیان شونده ژنوم شناسایی کرد و سپس بر مبنای نواحی محافظت شده و با ثبات، طراحی آغارگرهای مناسب انجام شود. این نشانگرها قابلیت کاربرد در بسیاری از گونه‌ها را دارند (Xu 2008; Varshney 2010). بر این اساس به نظر می‌رسد که بتوان نشانگرهای EST-SSR را برای گونه‌های خویشاوندی که اطلاعات کمی در مورد SSR یا EST آنها در دسترس است، مورد استفاده قرار داد. بطور مثال از نشانگرهای EST-SSR حاصل از جو ، برای مقایسه نقشه ژنتیکی در گندم، چاودار و برنج استفاده شده است (Yu et al. 2004; Varshney et al. 2005). علاوه بر این نشانگرهای SSR وابسته به ژن، گزینه‌های مناسب‌تری به عنوان نشانگرهای اورتولوگ با ثبات در ژنتیک و اصلاح گونه‌های مختلف هستند. برای مثال، مشخص شده است که یک دسته ۱۲ تایی از نشانگرهای EST-SSR جو دارای شباهت قابل قبولی با نشانگرهای EST در چهار گونه تک‌لپه‌ای (گندم، ذرت، سورگوم و برنج) و همچنین دو گونه دولپه‌ای (آرابیدوپسیس و یونجه) می‌باشد به طوری که می‌توان از آن‌ها در گونه‌های مذکور استفاده نمود (Varshney et al. 2005). شایان ذکر است که اگرچه نشانگرهای EST-SSR به عنوان نشانگرهای با کیفیت شناخته می‌شوند، ولی میزان چندشکلی آنها کمتر از نشانگرهای SSR ژنومی است (Cho at al. 2000; Eujayl et al. 2002; Thiel et al. 2003).

ب) نشانگرهای EST-SNP

توالی‌های EST، منابعی قابل استفاده در شناسایی چندشکلی‌های تک نوکلئوتیدی (SNP) می باشند (Picoult-Newberg et al. 1999; Kota et al. 2003). نشانگرهای SNP جدیدترین نشانگرهای وابسته به DNA بوده و به عنوان نشانگر ژنتیکی دو آللی شناخته می‌شوند. علاوه بر این، SNPها دارای فراوانی زیادی بوده و در مقایسه با SSR کمتر تحت تاثیر موتاسیون قرار می‌گیرند (Shahinnia and Sayed-Tabatabaei 2009). اگرچه SNPها به عنوان فناوری جدید نشانگری معرفی می‌شوند، ولی مبتنی بر دیدگاه جدیدی نیستند. از زمانی که توالی‌یابی DNA امکان‌پذیر گشت، امکان مقایسه آلل‌های یک مکان ژنی فراهم شد و اختلافات توالی آنها آشکار گشت. چنین نشانگرهایی را می‌توان در نقشه‌یابی فیزیکی و ژنتیکی و همچنین در انتخاب همسانه‌های BAC به منظور بررسی توالی کامل ژن‌ها بکار برد (Somers 2004).

در بیشتر گیاهان مهم زراعی، تعداد بسیاری از نشانگرهای EST بر اساس چندین ژنوتیپ از یک گونه تهیه شده‌اند و در بانک داده‌ها ثبت شده‌اند. دسته‌های فراوانی از نشانگرهای EST در یک گونه

Archive of SID
www.SID.ir

مورد مطالعه را می‌توان انتخاب و با استفاده از برنامه‌های زیست‌رایانه‌ای و ابزارهای تشخیص SNP در شرایط آزمایشگاهی، به شناسایی چندشکلی‌های تک نوکلئوتیدی (SNP) آنها پرداخت. تعداد بسیار زیادی از فرایندها و ابزارهای زیست‌رایانه‌ای جهت شناسایی SNPها معرفی شده‌اند (Varshney 2010). اساسا تمامی این ابزارها یا مسیرهای فرایندی، موجب گروه‌بندی نشانگرهای EST می‌شوند به نحوی که ESTهای مربوط به یک ژن در یک دسته قرار داده می‌گیرند و شرایط برای شناسایی SNPها فراهم می‌شود. نشانگرهای SNP حاصل از این روش بایستی توسط کروماتوگرام توالی ESTهای حاصل از مطالعات حقیقی در آزمایشگاه مورد تایید قرار گیرند زیرا احتمال خطا در SNPهای شناسایی شده حاصل از تجزیه‌وتحلیل زیست‌رایانه‌ای وجود دارد (Varshney 2010). بیشترین SNP شناخته شده بر این مبنا مربوط به ژنوم انسان است بطوریکه تا سال ۲۰۰۳ حدود ۱/۸ میلیون SNP گزارش شد. این در حالی است که بیشترین پیشرفت شناسایی SNP در غلات مربوط به ذرت بوده (Tenaillon et al. 2001) و این اطلاعات در مورد گندم، جو و برنج نیز به سطح قابل قبولی رسیده است. قابلیت اجرای اتوماتیک در شناسایی SNPها وجود دارد و روش‌های مختلف شناسایی در این خصوص ارائه شده است (Gupta and Varshney 2000). تصور می‌شود که در غلات فراوانی SNP معادل یک اتفاق در ۱۰۰ تا ۶۰۰ نوکلئوتید در بخش DNA بیان‌شونده باشد (Kanazin 2002; Somers et al. 2003). SNPهای حاصل از نشانگرهای EST، قابلیت کاربرد در آزمایشات تشخیصی و بالینی پزشکی را نیز دارا می باشند.

ج) نشانگرهای EST-TRAP

نشانگر TRAP [1] حاصل یک روش کارا و سریع مبتنی بر PCR است، که به عنوان نشانگر چندشکل در نزدیکی ژن هدف مورد استفاده قرار می‌گیرد (Hu and Vick 2003). در این روش، دو آغازگر ۱۸ نوکلئوتیدی برای ایجاد نشانگر استفاده می‌شود. یکی از آغازگرهای مورد استفاده در تهیه نشانگرهای TRAP مبتنی بر توالی EST هدف (بر مبنای توالی برگرفته از بانک اطلاعاتی) طراحی می‌شود و آغازگر دوم دارای توالی تصادفی است به نحوی که فاقد ناحیه غنی از AT یا GC (به ترتیب نواحی اینترونی یا اگزونی) باشد. روش TRAP می‌تواند در غربال ژنتیکی ژرم پلاسم در بانک ژن و همچنین در نشانمند کردن ژن‌هایی که دارای ارزش زراعی هستند، مفید باشد (Xu 2010).

د) نشانگرهای EST-Genic

با استفاده از اطلاعات مبتنی بر ترانسکریپتوم و ژن‌های اختصاصی یک موجود می توان نشانگرهای جدیدی ابداع نمود که از آن جمله می‌توان به چندشکلی‌های EST (بر اساس داده‌های EST)، نشانگرهای حاصل از اورتولوگ‌های حفاظت شده (بر اساس مقایسه توالی‌های ژنوم هدف با توالی‌های گونه‌های نزدیک به آن)، نشانگرهای تکثیریافته از ژنتیک توافقی (بر اساس ژن‌های شناخته شده در گونه‌های خویشاوند مدل)، نشانگرهای ژن‌های اختصاصی (بر اساس آغازگرهای حاصل از توالی ژن)، آنالوگ‌های ژن‌های مقاومت (بر مبنای آغازگر حاصل از دُمین‌های شناخته شده در مقاومت)، چندشکلی حاصل از ترانسپوزون‌های اگزونی (بر اساس آغازگرهای حاصل از LTRهای ترانسپوزونی یا اولیگونوکلئوتیدهای تصادفی ریزماهواره‌ای) و همچنین نشانگرهای مبتنی بر PCR برگرفته از نواحی اگزون، اینترون و راه‌انداز ژن‌های شناخته شده، اشاره کرد (Gupta and Rustgi 2004; Xu 2010).

ه) نشانگرهای EST-Genome wide

در گونه‌هایی که توالی EST یا ژنومی آنها موجود است، می‌توان ژنوتیپ‌های مورد نظر با خصوصیات مطلوب (مانند ژنوتیپ‌های والدی و نتاج حاصل از یک جمعیت در تهیه نقشه ژنتیکی) را جهت انجام توالی‌یابی مجدد (resequencing) به منظور شناسایی تنوع ژنتیکی بین افراد، مورد مطالعه قرار داد. همچنین توالی‌یابی جدیدی را می‌توان بر مبنای cDNA یا DNA ژنومی افراد مختلف انجام داد و داده‌های حاصل از این توالی‌ها را با توالی ژنوم مرجع ردیف‌سازی و مقایسه کرد. در نتیجه، تنوع ژنومی بین ژنوتیپ‌های مختلف یا تفاوت آنها با ژنوتیپ مرجع آشکار می‌شود (Varshney 2010). برای مثال، بر اساس ۴۵۴ توالی حاصل از ترانسکریپتوم مریستم های انتهای ساقه در لاین‌های خالص B73 و Mo17 ذرت (به ترتیب ۲۶۰۰۰۰ و ۲۸۰۰۰۰ توالی EST)، بیشتر از ۳۶۰۰۰ تنوع تک نوکلئوتیدی (SNPs) بین افراد

[1] Target Region Amplification Polymorphism

Archive of SID
www.SID.ir

مشخص شد که حاصل توالی‌یابی ۹۹۸۰ قطعه ژنومی در لاین B73 بود. تجزیه‌وتحلیل‌های بعدی به منظور شناسایی چند شکلی‌های حقیقی سبب کاهش تعداد SNPها به بیشتر از ۷۰۰۰ تفاوت شد. در این مطالعه متجاوز از ۸۵ درصد (۹۴ از ۱۱۰) از مجموع SNPهای مفروض توسط توالی‌یابی به روش سانجر در آزمایشگاه مورد تایید قرار گرفت (Barbazuk et al. 2007; Varshney 2010).

غربال ژنتیکی بر مبنای نشانگرهای مبتنی بر EST

به طور کلی، به منظور غربال ژنتیکی با استفاده از نشانگرهای مبتنی بر EST، بایستی توالی‌های EST مورد نظر در بیش از دو ژنوتیپ ردیف‌سازی شود. پس از مشخص شدن SNPها در بین ژنوتیپ‌ها، انتخاب روشی کارآمد به منظور غربال ژنتیکی SNPها ضروری به نظر می‌رسد. با توجه به ارائه و معرفی روش‌های مختلف غربال ژنتیکی مبتنی بر SNP (بیش از ۳۰ روش تاکنون)، به نظر می‌رسد که انتخاب این روش‌ها نه تنها بر مبنای پتانسیل نشانگری آنها بلکه بر اساس توجیه اقتصادی، سهولت کار و دقت انجام می‌شود (Shahinnia and Sayed-Tabatabaei 2009). در همین زمینه Bagge and Lübberstedt در سال ۲۰۰۸ اقدام به مقایسه دقیق روش‌های غربال ژنتیکی SNP نمودند و گزارش کردند که توجیه اقتصادی و موضوع مورد مطالعه از عوامل اصلی در انتخاب و استفاده از یک روش غربال مناسب است. بطور کلی در این خصوص به دو روش غربال ژنتیکی محدود و گسترده به تفصیل اشاره می‌شود.

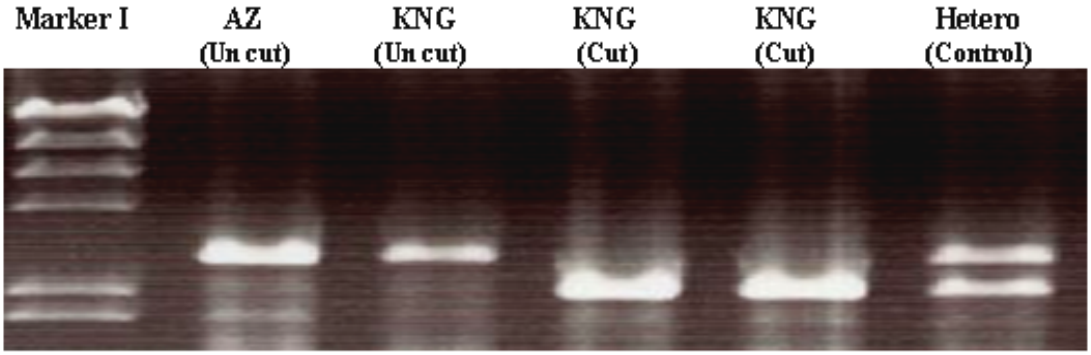
الف) غربال ژنتیکی محدود

زمانی که محدودیت منابع مالی و همچنین تعداد کمی نشانگر مبتنی بر EST-SNP برای غربال ژنتیکی وجود داشته باشد، ارزیابی ژنتیکی SNP با هزینه کم قابل اجراست. روش‌های موجود غربال ژنتیکی SNP را می‌توان بر اساس دو سیستم آشکارسازی مبتنی بر ژل و مستقل از آن تقسیم کرد. اکثر سیستم‌های مستقل از ژل مانند SnaPshot، Pyrosequencing وBiplex invader مبتنی بر اطلاعات توالی‌یابی بوده و نیاز به تسهیلات اتوماتیک اولیه زیادی دارند (Pati et al. 2004). این در حالی است که روش‌های مبتنی بر ژل، هزینه کمتری داشته و بازدهی متوسطی دارند. در این راستا، نشانگرهای CAPs معمول‌ترین روش مورد استفاده شناخته شده‌اند (Thiel et al. 2003; Shahinnia et al. 2009). در صورتی می‌توان نشانگرهای SNP را به CAPs تبدیل کرد که چندشکلی موجود در توالی نوکلئوتیدی، در مکان بری یک آنزیم انحصاری قرار داشته باشد (شکل ۱).

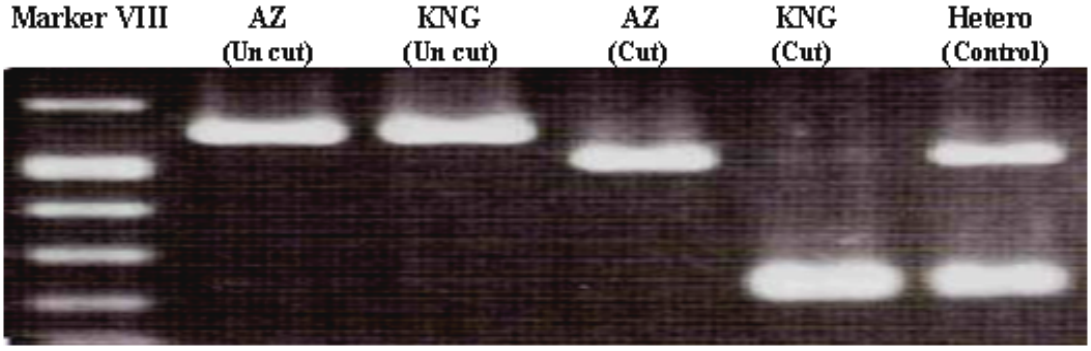
علاوه بر این Michaels and Amasino (1998) و Neff et al. (1998) نشان دادند که می‌توان چندشکلی‌های نوکلئوتیدی فاقد تاثیر در مکان برش آنزیمی را با استفاده از روش dCAPs به نشانگرهای مبتنی بر PCR تبدیل کرد. در این روش، با استفاده از آغازگرهایی که دارای یک یا چند نوکلئوتید غیر منطبق بر DNA الگو[1] و در ناحیه SNP است، می توان اقدام به ایجاد یک مکان برشی در محصول PCR نمود. سپس محصول PCR حاصل تحت تیمار آنزیم برشی مورد نظر واقع شده و پس از توالی یابی وجود یا فقدان SNP مورد ارزیابی قرار می‌گیرد (Shahinnia and Sayed-Tabatabaei 2009). روش dCAPs همانند CAPs روشی ساده، کارامد و نسبتا ارزان است (Michaels and Amasino 1998). با استفاده از این روش می‌توان تعداد نشانگرهای مفید در تهیه نقشه ژنتیکی و همسانه‌سازی مبتنی بر نقشه ژنی را فزونی داد. به عنوان مثال، جهت تهیه نقشه ژنتیکی فشرده در ناحیه ژنی *int-c*[2] واقع بر بازوی کوتاه کروموزوم 4H جو به منظور همسانه‌سازی ژن، تعداد ۲۸ نشانگر مولکولی EST بر مبنای نقشه‌های ژنتیکی ارائه شده در جو یا بر حسب نواحی مشابه آن در کروموزوم شماره ۳ برنج تهیه شد (Shahinnia et al. 2009). بر اساس SNPهای یافت شده حاصل از همریف‌سازی ESTها در دو رقم جو ژاپنی "Azumamugi" و "Kanto Nakate Dold" و تبدیل آنها به نشانگرهای مبتنی بر PCR به روش CAPs یا dCAPs، غربال ژنتیکی تعداد ۹۹ لاین خالص نوترکیب (RILs) حاصل از تلاقی این دو والد انجام شد. در مطالعه‌ای مشابه (Shahinnia et al. 2011)، نقشه ژنتیکی فشرده در ناحیه ژنی *dsp.ar*[3] جو با استفاده از ۳۶ نشانگر EST حاصل از مقایسه نواحی مشابه جو، برنج، سورگوم و براکی‌پودیوم[4] و تبدیل آنها به

[1] Mismatch
[2] Intermediate-spike c
[3] Dense spike
[4] *Brachypodium*

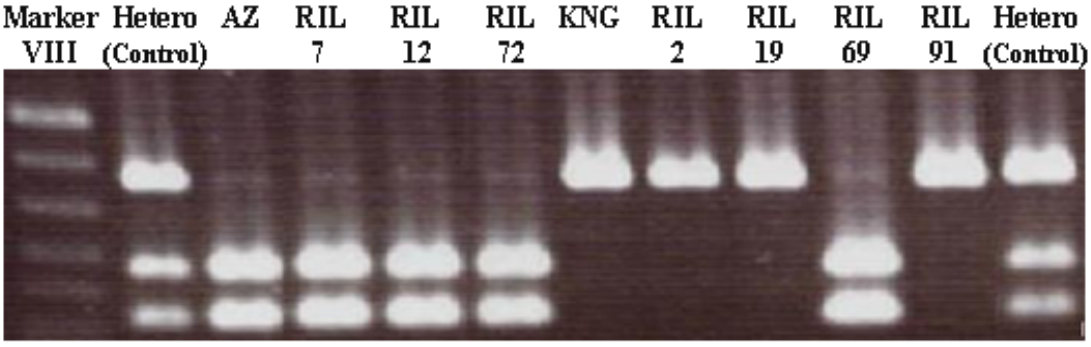
Archive of SID
www.SID.ir



(الف)



(ب)



(ج)

شکل ۱- تولید نوارهای پلی مورفیک با استفاده از نشانگرهای EST وتبدیل آنها به نشانگرهای CAPs و الکتروفورز ژل آگارز. الف) نوارهای حاصل از برش نشانگر *AV833030* در والد KNG با آنزیم *Mwo*I؛ ب) نوارهای حاصل از برش نشانگر *CA000177* در والدین AZ و KNG با آنزیم *Hind* III؛ ج) نوارهای حاصل از برش نشانگر *BJ473916* در والد AZ با آنزیم *Sac*II . در این شکل تمایز نوارهای والدی و شاهد هتروزیگوت به سهولت امکان پذیر است (shahinnia 2007; shahinnia et al. 2009).

www.SID.ir

نشانگرهای مبتنی بر PCR به روش CAPs تهیه شد. نتایج حاصل از غربال ژنتیکی تعداد ۱۹۹۶ فرد F_2 حاصل از تلاقی رقم "Bowman" و لاین جهش یافته "BW256" نشان داد که ژن مورد نظر در فاصله ۰/۳۷ سانتی مورگان نشانگرهای واقع در ناحیه سانترومر کروموزوم 7H قرار دارد.

چون امکان آشکارسازی چندشکلی حاصل از تنها یک نوکلئوتید با استفاده از الکتروفورز ژل پلی‌آکریل‌آمید وجود ندارد، استفاده از روش SSCPs [1] به عنوان راهکار دیگری در غربال ژنتیکی محدود و برای تشخیص چندشکلی‌های بسیار کوچک (یک یا دو نوکلئوتیدی) در محصولات PCR معرفی شد (Varshney 2010). این روش مبتنی بر تفاوت ساختار ثانویه حاصل از رشته‌های منفرد محصولات PCR است که در یک یا تعداد محدودی نوکلئوتید اختلاف دارند. به منظور آزمون چندشکلی به روش SSCPs، محصول PCR ژنوتیپ‌های مختلف حاوی SNP واسرشت شده و با استفاده از ژل آکریل‌آمید خنثی (غیر واسرشت ساز) الکتروفورز می‌شوند. به دلیل وجود SNP در ژنوتیپ‌ها، نوارهای چندشکل بر روی ژل قابل شناسایی است (Varshney 2010). این روش در مورد چندین گونه همچون کاج (Germano and Klein 1999)، ارزن (Bertin et al. 2005) و کاساوا (Castelblanco and Fregene 2006) مورد استفاده قرار گرفته است.

ب) غربال ژنتیکی گسترده

به دلیل در دسترس بودن حجم زیادی از داده‌های ژنومی، غربال ژنتیکی گسترده توالی‌های منفرد EST در سطح وسیعی از ژنوم و تعداد زیادی ژنوتیپ به عنوان نشانگرهای SNP به راحتی قابل اجراست، به نحوی که از نظر اقتصادی و زمان نیز قابل توجیه می‌باشد. انجام چنین روش‌هایی، به صورت خدمات آزمایشگاهی توسط شرکت‌ها و مراکز ژنتیکی مجهز به امکانات مورد نظر ارائه می‌شود و نیازی به تهیه امکانات بصورت مستقل در هر آزمایشگاه نیست.

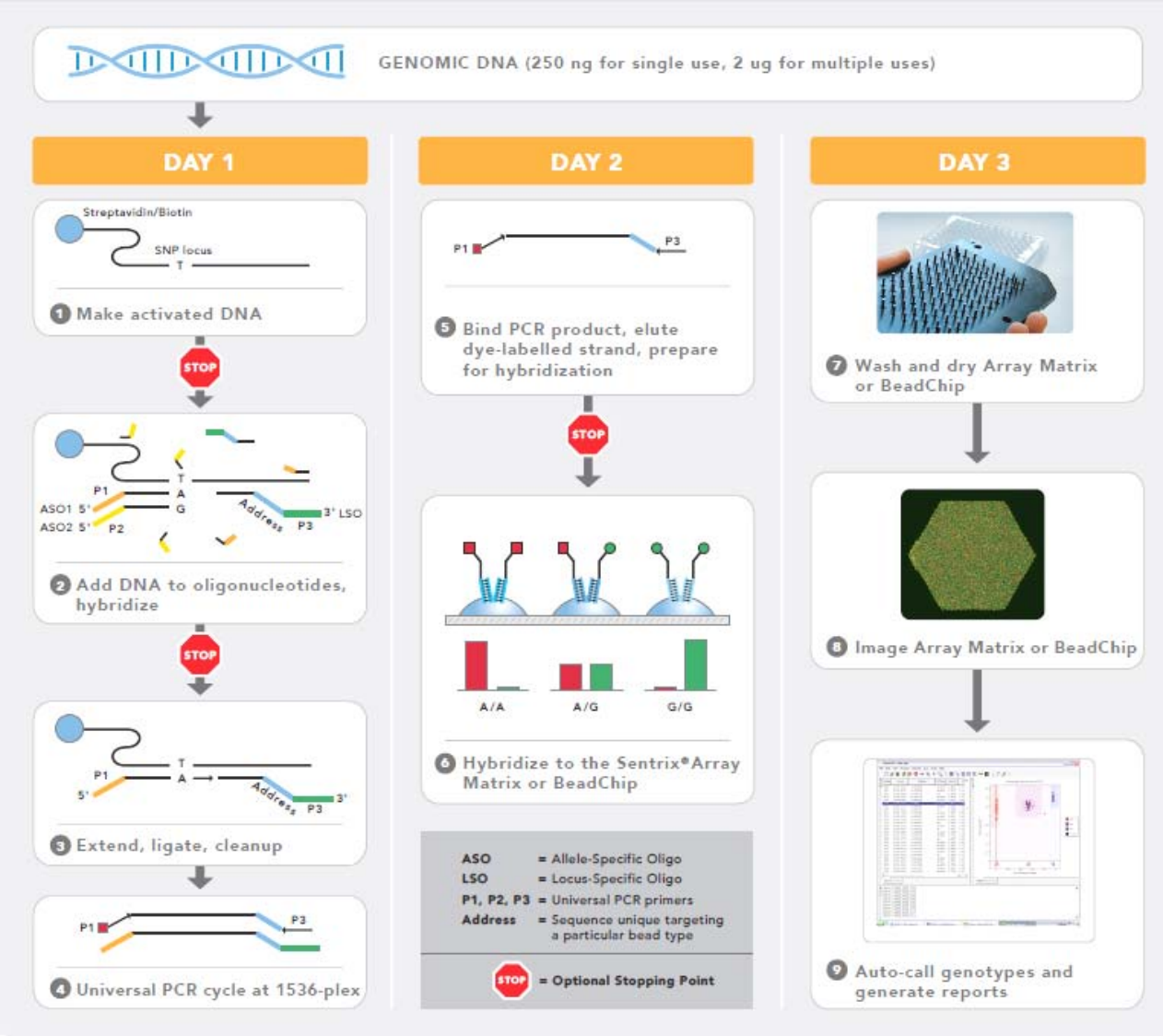
در حال حاضر مناسب‌ترین سنجش ژنتیکی گسترده، سنجش GoldenGate ارائه شده توسط کمپانی Illumina است (Fan et al. 2003). این روش بر اساس دورگه‌سازی آغازگرهای ویژه یک آلل (یا SNP) با DNA ژنومی مستقر بر یک بستر جامد است که مراحل انجام آن بطور اختصار در شکل ۲ بیان شده است. بنابراین، سنجش Goldengate می‌تواند غربال ژنتیکی تعداد ۹۶، ۱۹۲، ۳۸۴، ۷۶۸ و ۱۵۳۶ SNP را با استفاده از یک واکنش و در عرض ۳ روز انجام دهد (Varshney 2010). در بین ژنوم گیاهان، اولین غربال ژنتیکی SNP بر اساس سنجش GoldenGate در جو صورت گرفت. دو روش غربال ژنتیکی گسترده مبتنی بر SNP با استفاده از سنجش اولیگونوکلئوتیدی ترکیبی (Illumina (OPA[2] (که هر یک توانایی غربال ژنتیکی همزمان ۱۵۳۶ مکان SNP در جو را دارد) ارائه شده است (Close et al. 2009). با استفاده از این روش‌ها، غربال ژنتیکی در چندین جمعیت جو انجام شد که منجر به تهیه نقشه ژنی مبتنی بر تعداد ۲۹۴۳ نشانگر SNP (مطابق با ۱۶۵۳ سانتی‌مورگان) شد. بر مبنای نقشه ژنی حاصل و با استفاده از نشانگرهای مشابه در ژنوم‌های توالی‌یابی شده برنج، ذرت، براکی‌پودیوم و سورگوم، می‌توان ژنوتیپ‌های مختلف جو را مورد تجزیه‌وتحلیل قرار داد. در نتیجه، بر اساس تشابه ژنومی می‌توان نمایی کلی از ژن‌های فرضی در ژنوم جو را مشخص کرد. علاوه بر این، Druka et al. (2011) نتایج تجزیه‌وتحلیل ۸۸۱ لاین حاصل از تلاقی برگشتی (که حاوی آلل‌های جهش‌یافته وابسته به صفات نموی و مورفولوژیک بودند) راگزارش کردند. پس از غربال ژنتیکی این لاین‌ها توسط سنجش Illumina-GoldenGate و با استفاده از ۳۰۷۲ توالی EST حاوی SNP در والد دوره‌ای مورد استفاده، امکان مکان‌یابی همزمان ۴۲۶ آلل جهش‌یافته روی قطعات کرموزومی نقشه ژنتیک جو میسر شد. اخیرا، کمپانی Illumina سنجش ژنوتیپی Infinium را معرفی کرده است. این سنجش که SNPCGH نامیده می‌شود، روش بسیار دقیقی در غربال SNPهای موجود در ژنوم است که بر اساس هیبریداسیون ژنوم مقایسه‌ای[3] (Array-CGH) با یک ژنوم مرجع انجام می‌شود و آشکارسازی تنوع و تفاوت‌های آلل‌ها را تسریع می‌نماید. کارایی SNP-CGH توسط دو سنجش ژنوتیپی

[1] Single stranded DNA confirmation polymorphisms

[2] Oligo pool assays

[3] Comparative genomic hybridization

Archive of SID
www.SID.ir



شکل ۲- مراحل اختصاری انجام غربال ژنتیکی گسترده تنها به مدت ۳ روز با استفاده از سنجش Illumina-GoldenGate . در ابتدا لازم است تا آماده سازی DNA ژنومی هر نمونه به میزان ۲۵۰ نانو گرم برای استفاده در یک واکنش منفرد و یا ۲ میکرو گرم در واکنش های چند گانه صورت گیرد. سپس در مرحله اول، زنجیره DNA با استفاده از ماده Streptavidin/Biotin جهت اتصال به یک بستر پارامگنتیک فعال می‌شود. در مرحله دوم، ۲ آغازگر رفت و برگشتی P1 و P2 (ASOs) و یک آغازگر مکان اختصاصی P3 (LSO) برای هر مکان SNP طراحی می‌شود. در مرحله سوم و چهارم، واکنش زنجیره‌ای پلیمراز با استفاده از آغازگرهای طراحی شده همزمان برای ۱۵۳۶ مکان SNP انجام می‌شود. در مرحله پنجم، DNA تک رشته‌ای و بر چسب شده با یک ماده رنگی نظیر Dye برای اتصال به بستر جامد آماده می‌شود. در مرحله ششم، نمونه‌های مرتبط با هر توالی اختصاصی به بستر جامد Array Matrix یا Bead Cheap اتصال می‌یابد. در مرحله هفتم، توالی‌یابی بر اساس خاصیت فلورسنس Sentrix Array Matrix در دستگاه توالی‌یاب انجام می‌شود. در مرحله هشتم، تصویر کلی از نقاط فلورسنس شده مرتبط با توالی هر مکان SNP آشکار می‌شود. در مرحله نهایی نهم، لازم است تا اطلاعات حاصل از توالی‌یابی با استفاده از نرم افزار مرتبط با این روش تجزیه‌وتحلیل شود (اقتباس از: http://www.illumina.com/documents/products/workflows/workflow_goldengate_assay.pdf)

www.SID.ir

Infinium با استفاده از ۱۰۹۰۰۰ و ۳۱۷۰۰۰ چندشکلی تک نوکلئوتیدی (SNP) به منظور شناسایی ناهنجاری‌های کروموزومی موجود در نمونه‌های حاوی ناهنجاری (مانند سلول‌های توموری در انسان) به اثبات رسیده است (Peiffer et al. 2006; Varshney 2010). استفاده از سنجش Infinum هنوز در سیستم‌های گیاهی مرسوم نشده است، ولی به نظر می‌رسد که به زودی با معرفی SNPهای جدید مبتنی بر روش‌های Next Generation Sequencing، روش مذکور نیز در بسیاری از گونه‌های گیاهی قابل اجرا شود.

کاربرد نشانگرهای مبتنی بر EST

کارایی نشانگرهای EST در موارد متعددی به اثبات رسیده است. این نشانگر برای اولین بار در تهیه نقشه ژنی انسان بکار رفت و سپس از آن برای بررسی ژن‌ها و همچنین ارتقاء نقشه‌های ژنتیکی استفاده شد. با افزایش داده‌های ژنومی حاصل از پروژه‌های توالی‌یابی، داده‌های EST به منظور شناسایی SNPهای کاندید، تنوع پیرایش ژن، پیش‌بینی ساختار ژن و همچنین برای تشخیص ژن‌های موثر در بافت‌ها یا بیماری خاص، مورد استفاده قرار گرفت. نشانگرهای EST را می‌توان به منظور تعیین کمیت رونوشت‌های اختصاصی در کتابخانه‌های cDNA و همچنین برای تهیه نمای ژنی، بیان ژن، نقشه ژنی و شناسایی ژن بکار برد (Xu 2010).

نشانگرهای EST حاصل از توالی بخشی از cDNA می‌توانند اطلاعات دقیقی از تفاوت توالی بین آلل‌ها ارائه دهند و در پی آن به عنوان نشانگرهای مولکولی جدید معرفی شوند. لازم به ذکر است که بهبود میزان و کیفیت تولیدات گیاهی به وسیله همسانه‌سازی ژن تسهیل نمی‌شود بلکه اساسا با شناخت مکانیسم ژنتیکی نهفته در تنوع آللی میسر است. اینچنین نشانگرهایی نقش مهمی در شناسایی ژن یا آلل و همچنین پیشرفت برنامه‌های ژنومی گیاهان ایفا می‌کنند (Somers 2004). یکی از جنبه‌های پروژه‌های توالی‌یابی EST، تهیه نقشه EST یا مکان‌یابی ژن‌های کاندید در نقشه‌های ژنی است که بر اساس آن بتوان رابطه بین فنوتیپ و کارکرد ژن را مشخص کردد. SNPها را می‌توان به عنوان نشانگرهای مولکولی در مکان‌یابی دقیق ESTها بر روی نقشه ژنتیکی (بویژه آزمایشات ژنومیکس کارکردی) بکار برد (Somers et al. 2003; Somers 2004).

امروزه نشانگرهای مولکولی به عنوان ابزاری مهم در ژنتیک گیاهی و همچنین بهبود کارایی برنامه‌های اصلاحی مورد استفاده قرار می‌گیرند. ارتقای نشانگرهای مولکولی مبتنی بر روش‌های نوین اصلاحی است به نحوی که، در نهایت منجر به بهبود همزمان ژرم‌پلاسم و انتخاب موثر صفات مهم اقتصادی شود. در حال حاضر روش‌های جدید نشانگری مانند نشانگرهای مبتنی بر EST-SNP به عنوان نشانگرهای اقتصادی و بسیار کارامد برای غربال ژنتیکی در برنامه‌های گزینش به کمک نشانگر (MAS)، از اهمیت ویژه‌ای برخوردارند. در نتیجه استفاده از چنین روش غربالگری، امکان گزینش بر اساس کل ژنوم نه تنها برحسب مکان ژنی مشخص فراهم می‌شود بلکه ارزیابی توام مجموعه‌ای از نوترکیبی‌ها در ژنوم را نیز میسر می‌سازد (Langridge and Chalmers 2004). علاوه بر این، انطباق نقشه‌های ژنتیکی با نقشه‌های فیزیکی و در نهایت مکان‌یابی صفات کمی به روش Association mapping و بدون استفاده از جمعیت‌های ژنتیکی با استفاده از توالی‌های EST حاصل از نشانگرهای SNP تسهیل می‌شود. این نشانگرها سبب افزایش توانایی MAS در برنامه‌های اصلاحی و همچنین شناسایی تنوع آلل‌های وابسته به ژن در مجموعه ژرم‌پلاسم، بانک ژن یا جمعیت‌های جهش‌یافته می‌شود. در نهایت با استفاده از روش‌های مذکور، امکان اجرای غربال ژنتیکی گسترده برای صفاتی با وراثت کمی، به‌ویژه در ژنوم گیاهان زراعی، همراه با صرف هزینه‌های نسبتا کم، تسریع و تسهیل برنامه‌های اصلاحی به کمک نشانگر میسر می‌شود (Varshney 2010).

سپاسگزاری

بدینوسیله از وزارت کشاورزی و موسسه تحقیقات علوم کشاورزی (NIAS) ژاپن (طرح Genomics for Agricultural Innovation, TRC1004) و همچنین موسسه تحقیقات ژنتیک گیاهی (IPK) آلمان (طرح STE1102/2-1) بدلیل حمایت مالی و در اختیار گذاشتن امکانات آزمایشگاهی سپاسگزاری می‌شود.

Archive of SID
www.SID.ir

منابع

Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olde RF, Moreno RF (1991) Complementary DNA sequencing: expressed sequence tags and human genome project Science 252:1651-1656.

Anderson JR, Lübberstedt T (2003) Functional markers in plants. Trends in Plant Science 8:554-560.

Bagge M, Lübberstedt T (2008) Functional makers in wheat: technical and economic aspects. Molecular Breeding 22:319-328.

Barbazuk WB, Emrich SJ, Chen HD, Li L, Schnable PS (2007) SNP discovery via 454 transcriptome sequencing Plant Journal 51:910-918.

Bateman A, Coin L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer ELL (2004) The Pfam protein families database. Nucleic Acids Research 32: D138-41.

Bedell JA, Korf I, Gish W (2000) MaskerAid: a performance enhancement to Repeat Masker. Bioinformatics 16: 1040-1.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. and Rapp BA (2002) GenBank. Nucleic Acids Research 30:17-20.

Bertin I, Zhu JH, Gale MD (2005) SSCP-SNP in pearl millet-a new marker system for comparative genetics. Theoretical and Applied Genetics 110:1467-1472.

Boguski MS, Lowe TM, CM Tolstoshev CM (1993) dbEST: database for "expressed sequence tags". Nature Genetics 4:332-333.

Boguski MS, and Schuler GD (1995) ESTablishing a human transcript map. Nature Genetics 10:369-71.

Boguski MS, Tolstoshev CM, Bassett DE (1994) Gene discovery in dbEST. Science 265:1993-1994.

Castelblanco W, Fregene MF (2006) SSCP-SNP-based conserved ortholog set (COS) markers for comparative genomics in cassava (*Manihot esculenta* Crantz). Plant Molecular Biology Reporter 24:229-236.

Cho YG, Ishii T, Temnykh S, Chen X, Lipovich L, McCouch SR, Park WD, Ayres N, Cartinhour S (2000) Diversity of microsatellites derived from genomic libraries and GeneBank sequences in rice. Theoretical and Applied Genetics 100: 713-722.

Close TJ, Bhat PR, Lonardi S, Wu Y, Rostoks N, Ramsay L, Druka A, Stein N, Svensson JT, Wanamaker S et al. (2009) Development and implementation of high-throughput SNP genotyping in barley. BMC Genomics 10: 582.

Druka A, Franckowiak J, Lundqvist U, Bonar N, Alexander J, Shahinnia F, Radovic S, Vendramin V, Morgante M, Stein N, Waugh R (2011) Genetic dissection of barley morphology and development. Plant Physiology 155: 617-627.

Eujayl I, Sorrels ME, Baum M, Wolters P, Powell W (2002) Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. Theoretical and Applied Genetics 104: 399-407.

Fan JB, Oliphant A, Shen R, Kermani BG, Garcia F et al. (2003) *Highly parallel SNP genotyping*. Cold Spring Harbor Symopsium Quantitaive Biology 68:69-78.

Germano J, Klein AS (1999) Species-specific nuclear and chloroplast single nucleotide polymorphisms to distinguish *Picea glauca*, *P. mariana* and *P. rubens*. Theoretical and Applied Genetics 99:1-2.

Gupta PK, Rustgi S (2004) Molecular markers from the transcribed/expressed region of the genome in higher plants. Functional and Integrated Genomics 4:139-162.

Gupta PK, Rustgi S, Mir PR (2008) Array-based high throughput DNA markers for crop improvement. Heredity 101:5-18.

Gupta PK, Varshney RK (2000) The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica 113: 163-185.

Hotz-Wagenblatt A, Hankeln T, Ernst P, Glatting KH, Schmidt ER, Suhai S (2003) ESTAnnotator: a tool for high throughput EST annotation. Nucleic Acids Research 31: 3716-3719.

Hu J, Vick BA (2003) Target region amplification polymorphism: a novel marker technique for plant genotyping. Plant Molecular Biology Reporter 21: 289-294.

Kanazin V, Talbert H, See D, DeCamp P, Nevo E, Blake T (2002) Discovery and assay of single nucleotide polymorphisms in barley (*Hordeum vulgare*). Plant Molecular Biology 48: 529-537.

Kota R, Rudd S, Facius A, Kolesov G, Theil T, Zhang H, Stein N, Mayer K, Graner A (2003) Snipping polymorphisms from large EST collections in barley (*Hordeum vulgare* L.). Molecular Genetics and Genomics 270: 24 -33.

Langridge P, Chalmers K (2004) The principle: identification and application of molecular markers. In: Lörz H, Wenzel G (eds.) *Molecular Marker Systems*. Biotechnology in Agriculture and Forestry, Vol 55. Springer-Verlag Berlin Heidelberg pp. 3-24.

Lee B, Hong T, Byun SJ, Woo T, Choi YJ (2007) ESTpass: a web-based server for processing and annotating expressed sequence tag (EST) sequences. Nucleic Acids Research 35: W159-W162.

Lee Y, Tsai J, Sunkara S, Karamycheva S, Pertea G, Sultana R, Antonescu V, Chan A, Cheung F, Quackenbush J (2005) The TIGR Gene Indices: clustering and assembling EST and known genes and integration with eukaryotic genomes. Nucleic Acids Research 33:D71-4.

Mao C, Cushman JC, May GD, and Weller JW (2003) ESTAP–an automated system for the analysis of EST data Bioinformatics 19: 1720-1722.

Michaels SD, Amasino RM (1998) A robust method for detecting single-nucleotide changes as polymorphic markers by PCR. Plant Journal 14:381-385.

Archive of SID

www.SID.ir

Nagaraj HS, Gasser RB, Ranganathan S (2006) A hitchhiker's guide to expressed sequence tag (EST) analysis. Briefings in Bioinformatics 8 : 6-21.
Nagaraj SH, Nandan Deshpande N, Robin B, Gasser RB, Ranganathan S (2007) ESTExplorer: an expressed sequence tag (EST) assembly and annotation platform. Nucleic Acids Research 35: W143-W147.
Nam SH, Kim DW, Jung TS, Choi YS, Kim DW, Choi HS, Choi SH, Park HS (2009) PESTAS: a web server for EST analysis and sequence mining. Bioinformatics 25: 1846-1848.
Neff MM, Neff JD, Chory J, Pepper AE (1998) dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: Experimental applications in *Arabidopsis thaliana* genetics. Plant Journal 14:387-392.
Okubo K, Hori N, Matoba R, Niiyama T, Fukushima A, Kojima Y, Matsuba K (1992) Large scale cDNA sequencing for analysis of quantitative and qualitative aspects of gene expression. Nature Genetics 2:173-179.
Pati N, Schowinsky V, Kokanovic O, Magnuson V, Ghosh S (2004) A comparison between SNaPshot, pyrosequencing, biplex invader SNP genotyping methods: Accuracy, cost, and throughput. Journal of Biochemistry and Biophysical Methods 60:1-12.
Peiffer DA, Le JM, Steemers FJ, Chang W, Jenniges T, Garcia F, Haden K, Li J, Shaw CA, Belmont J, Cheung SW, Shen RM, Barker DL, Gunderson KL (2006) High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Research 16:1136-1148.
Picoult-Newberg L, Ideker TE, Pohl MG, Taylor SL, Donaldson MA, Nickerson DA, Boyce- Jacino M (1999) Mining SNPs from EST databases. Genome Research 9: 167-174.
Pontius JU, Wagner L, Schuler GD (2002) Unigene: A unified view of the transcriptome. In: McEntyre J, Ostell J (eds.) *The NCBI Handbook*, National Center for Biotechnology Information USA.
Rao DM, Moler JC, Ozden M, Zhang Y, Liang C, Karro JE (2010) PEACE: Parallel Environment for Assembly and Clustering of Gene Expression. Nucleic Acids Research 38: W737-W742.
Shahinnia F (2007) Mapping of QTLs for important agronomic traits and map marker enrichment in recombinant inbred lines of barley (*Hordeum volgare.* L) PhD dissertation College of Agriculture Isfahan University of Technology. (In Farsi).
Shahinnia F, Druka A, Fronckowiak J, Waugh R, Morgante M, Stein N (2011) High resolution mapping of the gene dense spike (dsp1) to the genetic centromer of barley chromosome 7H. Theoretical and Applied Genetics 124:373-84.
Shahinnia F, and Sayed-Tabatabaei BE (2009) Conversion barley single-nucleotide polymorphism (SNP) into PCR-based marker using dCAPS method. Genetics and Molecular Biology 32: 564-567.
Shahinnia F, SayedTabatabaei BE, Pourkheirandish M, Sato K, Komatsuda T (2009) Mapping of QTL for intermedium spike on barley chromosome 4H using EST-based markers. Breeding Science 59: 383-390.
Somers DJ (2004) Molecular marker systems and their evaluation for cereal genetics. IN: Gupta PK , Varshney RK (eds.), *Cereal Genomics*, Kluwer Academic Publishers, Netherlands pp. 19-34.
Somers DJ, Fedak G, Savard M (2003) Molecular mapping of novel genes controlling Fusarium head blight resistance and deoxynivalenol accumulation in spring wheat. Genome 49:555-564.
Tenaillon MI, Swakins MC, Long AD, Gaut RL, Doebley JF, Gaut BS (2001) Patterns of DNA sequence polymorphism along chromosome 1 of maize (*Zea mays* ssp. Mays L.). Proceedings of National Academy of Sciences 98: 9161-9166.
Thiel T, Michalek W, Varshney RK, 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 106:411-422.
Varshney RK (2010) Gene-Based Marker Systems in Plants: High Throughput Approaches for Marker Discovery and Genotyping. In: Varshney RK, Jain SM, Brar DS (eds.) *Molecular Techniques in Crop Improvement*, Springer Science+Business Media BV pp. 119-142.
Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends in Biotechnology 23: 48-55.
Varshney RK, Mahender T, Aggrawal RK, Borner A (2007) Genic molecular markers in plants: development and applications. In: Varshney RK, Tuberosa R (eds.) *Genomics-Assisted Crop Improvement*, Vol I: Genomics Approaches and Platforms. Springer The Netherlands pp 13-30.
Varshney RK, Nayak SN, May GD, Jackson SA (2009) Next generation sequencing technologies and their implications for crop genetics and breeding. Trends in Biotechnology 27: 522-530.
West MAL, van Leeuwen H, Kozil A, Kliebenstein DJ, Doerge RW, St Clair DA, Michelmore RW (2006) High-density haplotyping with microarray-based expression and single feature polymorphism markers in *Arabidopsis.* Genome Research 16:787-795.
Xu Y (2010) Molecular Plant Breeding. CAB International Printed in UK pp 21-58.
Yu JK, La Rota M, Kantety RV, Sorrells ME (2004) EST-derived SSR markers for comparative mapping in wheat and rice. Moleclar Genetics and Genomics 271: 742-751.
Zdobnov EM, Apweiler A (2001) InterProScan – an integration platform for the signature-recognition methods in InterPro. Bioinformatics 17:847-8.

Archive of SID

www.SID.ir