

آشنایی با نرم افزار GMOseek به منظور شناسایی و بررسی موجودات تراریخته

مریم زکوی^۱، مسعود توحیدفر^{۲*}

^۱دانش آموخته دکتری بیوتکنولوژی، دانشکده مهندسی انرژی و فناوری های نوین، دانشگاه شهید بهشتی، تهران، ایران

^{۲*}دانشیار دانشکده فناوری های نوین، دانشکده مهندسی انرژی و فناوری های نوین، دانشگاه شهید بهشتی، تهران، ایران

M_Tohidfar@sbu.ac.ir

چکیده

با توجه به سرعت فزاینده تایید و تجاری سازی محصولات تراریخته در سراسر جهان، آزمایش های مربوط به شناسایی این محصولات در غذا، علوفه و نمونه های بذری با توجه به مقررات مربوط، بسیار مشکل و هزینه بر خواهد بود. بسیاری از روش های تشخیصی در حال حاضر با توجه به بودجه محدود متناسب با منابع در دسترس، برای افزایش بهره وری از نرم افزارهای مبتنی بر جدول ماتریسی از داده ها "رویکرد ماتریسی" استفاده می نمایند. در بسیاری از مواقع رویکرد ماتریسی با استفاده از ابزار محاسباتی خاص و اطلاعات بسیار اندک، در جهت استفاده کارآمد از داده های موجود قابل اجرا است. نرم افزار توسعه یافته GMOseek برای پشتیبانی از تصمیم گیری های انجام شده در تمامی مراحل آزمایش های تشخیص موجودات تراریخته، از جمله تفسیر نتایج خام آزمایشگاهی طراحی شده است. این نرم افزار از جدول ماتریسی رخداد (event) و عناصر ژنتیکی مربوط به آن ها و اطلاعات موجود در مورد نمونه استفاده می کند. این ابزار با استفاده از استراتژی بهینه سازی، روش های غربالگری مناسب تری برای نمونه پیشنهاد می دهد. در ادامه، نرم افزار GMOseek اجازه می دهد تا کاربر به دلخواه به جستجوی ترکیب مقرون به صرفه ای از تست های غربالگری بر روی نمونه خاص تراریخته بپردازد. بعلاوه این نرم افزار به کاربر امکان انتخاب تجزیه و تحلیل مناسب برای تعیین حضور یک رخداد به خصوص در نمونه مورد بررسی را می دهد.

کلمات کلیدی: موجودات تراریخته، رویکرد ماتریسی، بهینه سازی، هزینه یک روش کارآمد

مقدمه

بیوتکنولوژی گیاهی، اخذ تاییدیه و کنترل تطبیقی

محصولات تراریخته از شرایط لازم مقررات ایمنی زیستی در بسیاری از کشورها است [۱۱].

موجود تراریخته ارگانیسمی است که ژنوم آن با معرفی ساختار ژنتیکی خارجی (یک ژن خارجی)

محصولات تراریخته از ابتدای تجاری سازی در سال ۱۹۹۶ تا کنون در مقیاس جهانی سهم قابل توجهی در کشاورزی و زنجیره غذایی [۱،۲] داشته اند. برای پاسخ دادن به ملاحظات در مورد استفاده از محصولات

الگوریتم نرم افزار GMOseek

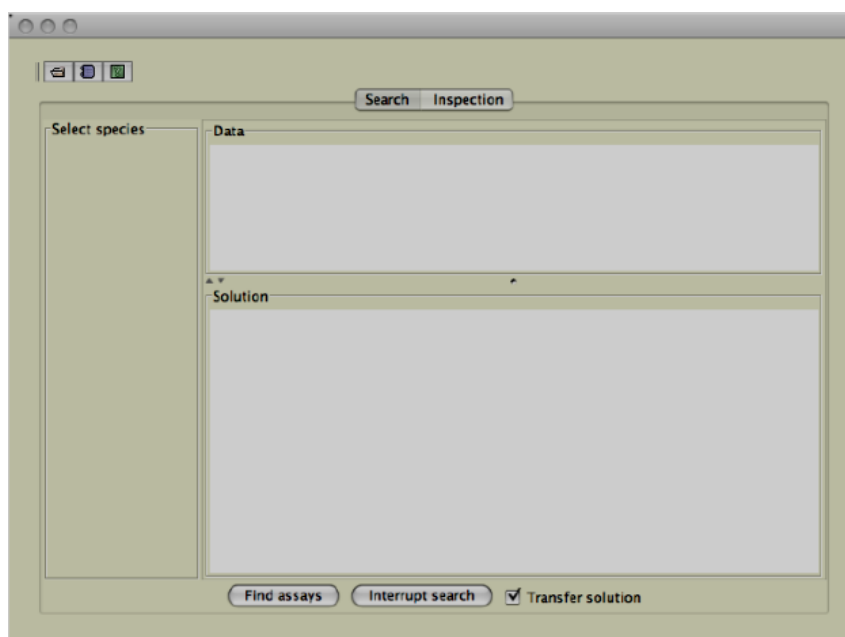
الگوریتم GMOseek استفاده از استراتژی جستجوی هوشمند برای پیدا کردن روش بهینه شناسایی محصولات تراریخته می‌باشد. نرم‌افزار این امکان را برای کاربر فراهم می‌کند که با ارزیابی روش‌هایی که بیشترین طیف محصولات تراریخته را از مجموعه داده اولیه تحت پوشش قرار می‌دهند، در کوتاه‌ترین زمان ممکن تست ترکیبی مناسبی را با توجه به هزینه‌های مورد انتظار اتخاذ نماید. محاسبه هزینه بسیار پیچیده است، اما به زبان ساده می‌توان گفت که این نرم‌افزار بعد از یافتن مطلوب‌ترین روش، ترکیبات دیگر را با توجه به هزینه مورد انتظار بهینه‌ترین روش محاسبه می‌کند و اگر هزینه این روش‌ها خیلی بیشتر از مطلوب‌ترین روش باشد، آن‌ها را حذف می‌کند [۱۱].

این نرم‌افزار به زبان جاوا نگارش شده است و بر روی هر سیستمی که دارای جاوا با ورژن ۱/۵ یا بالاتر باشد اجرا می‌شود. نرم‌افزار GMOseek به صورت یک فایل به نام GMO.jar ارائه شده است. این برنامه را می‌تواند به شکل رایگان از وب سایت <http://www.gmoseek.com/gmoseek> دریافت و آن را در هر مکان دلخواهی در سیستم مورد استفاده کاربر ذخیره نمود [۱۱]. بلافاصله بعد از اجرای برنامه پنجره زیر قابل مشاهده است (شکل ۱). نوار ابزار شامل سه آیکون برای فراخوانی فایل جدید، فراخوانی فایل‌هایی که اخیراً باز شده‌اند و راهنمای نرم‌افزار می‌باشد [۱۱].

متشکل از چندین جز ژنتیکی (ژن‌های مورد نظر، توالی‌های تنظیمی برای فعالیت عملکرد ژن اصلی در ارگانیسم میزبان و غیره) اصلاح یافته است. بنابراین، مناسب‌ترین روش برای تشخیص این قبیل موجودات، بررسی حضور توالی DNA تراریخته در یک رخداد خاص (event) خواهد بود. تا به امروز، تکنیک‌هایی مرجع شامل واکنش زنجیره‌ای پلیمرز (PCR)، سادرن‌بلات و ریل‌تایم پی‌سی‌آر (Real-time PCR) بوده است [۳، ۴، ۵ و ۶]. سنجش PCR به طور معمول به منظور شناسایی و غربالگری اجزای ژنتیکی (یا گروهی از اجزای ژنتیکی) موجودات تراریخته استفاده می‌شود [۴]. همچنین می‌توان از آن در ردیابی یک محصول تراریخته خاص (event-specific tests) بهره گرفت [۷].

هدف اصلی این مطالعه، ارائه و ارزیابی کاربردی نرم‌افزار GMOseek است که می‌توان آن را از جهات بسیاری، نوع ارتقا یافته نرم‌افزار GMOTrack دانست. این نرم‌افزار بر اساس داده‌هایی که به فرمت ماتریس فراخوانی شده‌اند، عملیات تجزیه و تحلیل را انجام می‌دهد [۸]. رویکرد ماتریس برای ردیابی محصولات تراریخته، متکی بر ماتریسی از محصولات تراریخته تولید شده (به منظور تهیه غذا، خوراک دام) و اجزای ژنتیکی آن‌ها است. مزایا و محدودیت‌های رویکرد مبتنی بر ماتریس، در مقالات متعددی مورد بحث قرار گرفته است و از جمله آن‌ها می‌توان به حساسیت سنجش و اختصاصیت این روش اشاره کرد که می‌تواند مسئولیت هرگونه اشتباهی را در تفسیر نتایج بر عهده بگیرد [۸، ۹ و ۱۰].

"زکوی و توحیدفر، آشنایی با نرم افزار GMOseek به منظور شناسایی و بررسی موجودات تراریخته"



شکل ۱- پنجره ورودی نرم افزار

تهیه فایل ورودی نرم افزار GMOseek

نرم افزار GMOseek از یک ماتریس داده به عنوان پایگاه دانشی برای آنالیز بهینه سازی و پشتیبانی تصمیم گیری استفاده می کند. اجزای اصلی ماتریس داده ها، شامل اطلاعات مربوط به محصولات تراریخته و روش های غربالگری برای شناسایی آنها می باشد. این اطلاعات به وسیله داده های مربوط به گونه گیاه تراریخته و احتمال حضور یک رخداد خاص در یک

نمونه تکمیل شده است. مجموعه داده لازم برای شروع کار با نرم افزار GMOseek را می تواند با ویرایش در ابزارهایی مانند آفیس اپن سورس کالک (OpenOffice Calc)، اکسل (Excel) و نوت پد (Notepad)، متناسب با دسترسی کاربر به این برنامه ها تهیه و در نهایت در یک فایل با پسوند تب (*.tab) ذخیره نمود [۱۱]. فرمت ماتریس داده های ورودی در جدول ۱ شرح داده شده است.

جدول ۱- مثالی از داده های ورودی مناسب برای نرم افزار GMOseek

GMOname	species	probability	P-35S	P-TA29	P-nos	CP4-epsps	bar	Barstar	T-nos	T-35S	P-35S::bar
d	d	c	d	d	d	d	d	d	d	d	d
m	m	m									
RRS	soybean	0.54	1	0	0	1	0	0	1	0	0
Bt176	maize	0.11	1	0	0	0	1	0	0	1	1
MS1	oilseed rape	0.05	0	1	1	0	1	0	1	0	1
RF1	oilseed rape	0.01	0	1	1	0	1	1	1	0	1
RF2	oilseed rape	0.01	0	1	1	0	1	1	1	0	0
HCN92	oilseed rape	0.01	1	0	1	0	0	0	0	1	0

همان‌طور که اشاره شد، پسوند فایل ورودی این نرم افزار تب (*.tab) می‌باشد، برای تهیه ماتریس داده‌ها باید شروط زیر رعایت شود:

- در خط یا ردیف اول اسم هر ستون نوشته می‌شود. در سه ستون اول نام محصول تراریخته، گونه و احتمال حضور آن در نمونه مورد بررسی آورده می‌شود. ستون چهارم به بعد شامل روش‌های شناسایی محصولات تراریخته است.

- ستون چهارم تا آخر به ترتیب مربوط به پروموتر 35S (P-35S)، پروموتر TA29 (P-TA29)، پروموتر nos (P-nos)، آنزیم 3-enolpyruvylshikimate-phosphate synthase (CP4-epsps) از آگروباکتریوم استرین CP4، نشانگر انتخابی مقاومت به آفت‌کش bialaphos (bar)، ژن مسئول برطرف‌سازی نرعیمی (Barstar)، ترمیناتور nos (nos)، ترمیناتور 35S (35S)، وجود ژن مقاومت به آفت‌کش و پروموتر 35S به‌طور همزمان (P-35S::bar) می‌باشد.

- در ستون اول از سطر چهارم تا انتها ارقام یا رخدادهای تراریخته ذکر شده است. به این‌صورت که RRS یک رقم سویای تراریخته مقاوم به علف‌کش، Bt176 یک رقم ذرت تراریخته مقاوم به آفت‌کش، MS1 کلزای مقاوم به آفت‌کش، RF1 رقم کلزای حاوی ژن‌های مقاومت به آفت‌کش و ژن

برطرف‌سازی نرعیمی، RF2 رقم کلزای حاوی ژن‌های مقاومت به آفت‌کش، ژن برطرف‌سازی نرعیمی، ژن مقاومت به آفت‌کش و پروموتر 35S، و HCN92 رقم کلزای مقاوم به علف‌کش گلیکوفوزات می‌باشد.

- خط دوم نشان دهنده‌ی نوع داده است، d و c به‌ترتیب برای داده‌های اسمی و عددی استفاده می‌شود.

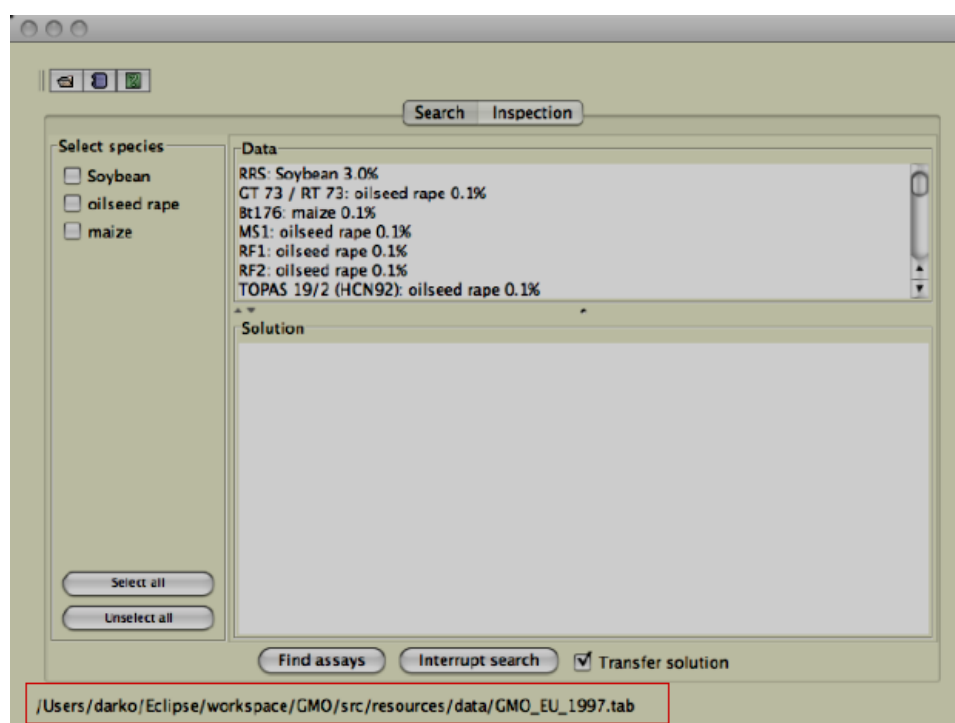
- خط سوم نشان دهنده‌ی نوع ستون می‌باشد، m در سه ستون اول برای نام محصول تراریخته، گونه و احتمال حضور آن در نمونه مورد بررسی استفاده می‌شود. در مورد ستون‌های مربوط به روش‌های شناسایی خط سوم خالی باقی گذاشته می‌شود.

- در ستون‌های مربوط به روش‌های تشخیص محصول تراریخته از ۰ و ۱ استفاده می‌شود. ۱ به معنی پاسخ مثبت به روش و صفر پاسخ منفی به روش شناسایی است.

بررسی داده‌ها و تفسیر نتایج در نرم‌افزار GMOseek

هنگامی که فایل ورودی بارگذاری شود، پنجره‌ای مشابه به شکل ۲ مشاهده می‌شود. همان‌طور که در شکل ۲ دیده می‌شود، در قسمت پایین پنجره (شکل دوم کادر قرمز رنگ)، نام فایل بارگذاری شده نشان داده می‌شود.

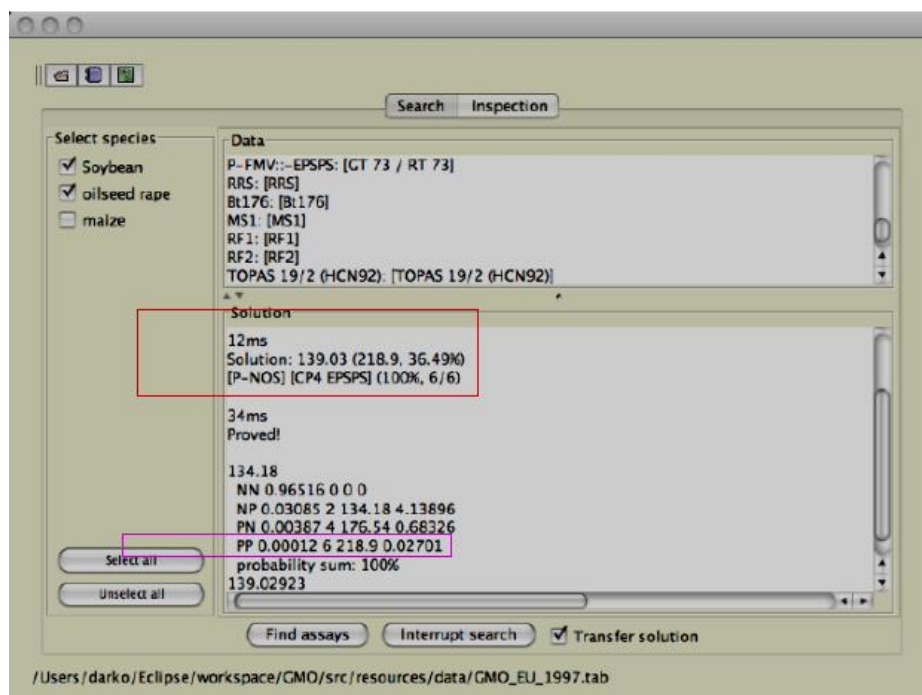
"زکوی و توحیدفر، آشنایی با نرم افزار GMOseek به منظور شناسایی و بررسی موجودات تراریخته"



شکل ۲- پنجره‌ی ورودی نرم افزار پس از بارگذاری فایل

[RRS] یعنی پروموتور P35S برای غربالگری هر سه محصول تراریخته‌ی Bt176، TOPAS 19/2 (HCN92) و RRS در نمونه مورد بررسی استفاده می‌شود. بعد از بارگذاری داده‌ها و انتخاب گونه‌هایی مورد نظر از قسمت Select species (در این مثال Soybean و Oil seed rape انتخاب شده است)، برای بررسی و آنالیز آن‌ها گزینه Find assays انتخاب می‌شود. مطابق شکل ۳ در قسمت Solution نتایج و یا خروجی نرم افزار قابل مشاهده است.

بعد از بارگذاری فایل داده، در سمت چپ نرم افزار در قسمت Select species، اسم گونه‌ها و در قسمت Data، اطلاعات ورودی شامل نام محصول، اسم گیاه و درصد احتمال حضور محصول تراریخته در نمونه مورد بررسی قابل مشاهده است. برای مثال RRS: Soybean 3.0% به این معنی است که احتمال حضور محصول تراریخته RRS در نمونه سویای مورد بررسی ۳ درصد می‌باشد. همچنین در مورد روش تشخیص P35S: [Bt176] [TOPAS 19/2 (HCN92)] مثلاً



شکل ۳- خروجی نرم افزار

فایل ورودی را دارد. روش اولی که در قسمت Solution می آید، همان روش نهایی و اصلی است و اگر نرم افزار با صرف مدت زمان بیشتر، روش بهتری پیدا کند. در زیر روش اول نتایج روش های بعدی آورده می شود. در انتها مدت زمان لازم برای اثبات بهینه بودن روش مشاهده می شود. برای مثال در اینجا ۳۴ میلیونیم ثانیه، مدت زمان لازم برای اثبات بهینه بودن روش اول است. آخرین قسمت در خروجی نتایج، نشان دهنده ی جزئیات هزینه روش اصلی است. در ابتدای این قسمت هزینه ثابت روش های غربالگری نوشته شده است (۱۳۴/۱۸)، سپس تمام نتایج ممکن برای روش و همچنین هزینه کل (۱۳۹/۰۲۹۲۳) آورده می شود که همان هزینه محاسبه شده در آخرین روش است. P و N در این قسمت به ترتیب به معنی منفی (Negative) و مثبت (Positive)

در سه خط اول نتایج (در قسمت Solution) نشان داده می شود که اولین روش در ۱۲ میلیونیم ثانیه (12 ms) پیدا شده است و علاوه بر این، هزینه این روش معادل ۱۳۹/۰۳ واحد است (شکل سوم کادر قرمز رنگ) که ۳۶/۴۹ درصد از هزینه روش های غربالگری اختصاصی رخداد را در برمی گیرد (هزینه تست رخداد خاص معادل ۲۱۸/۹ واحد می باشد). در نهایت، روش انتخابی برای غربالگری نشان داده می شود، مثلاً در این جا P-NOS و CP4 EPSPS انتخاب شده است. به دنبال آن درصد پوشش محصولات تراریخته در این روش مشخص می شود که نشان می دهد چه تعداد از محصولات تراریخته در فایل ورودی با این روش غربالگری قابل شناسایی هستند. مثلاً ۶/۶ بدین معنی است که روش ذکر شده قابلیت شناسایی هر شش محصولات تراریخته در

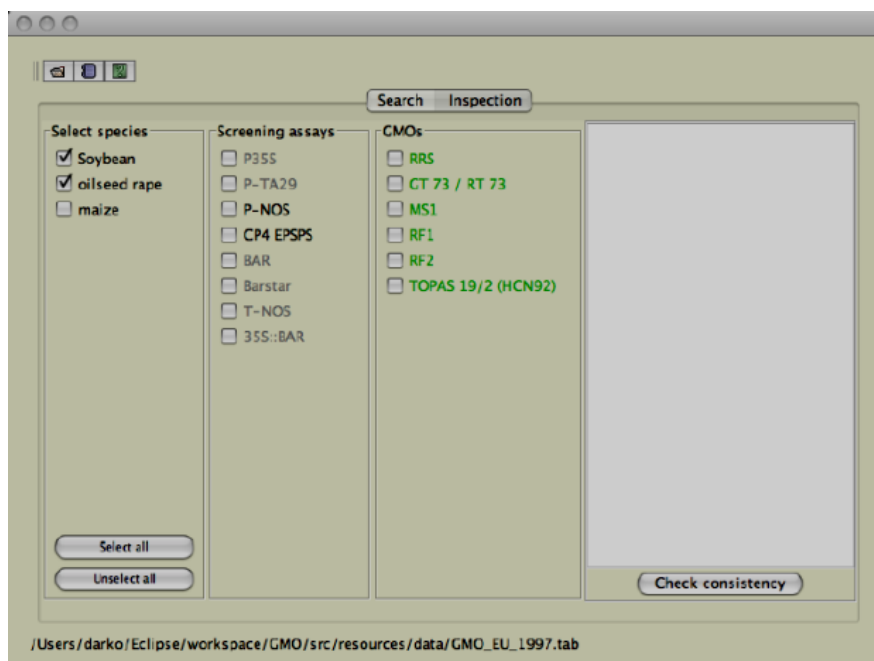
"زکوی و توحیدفر، آشنایی با نرم افزار GMOseek به منظور شناسایی و بررسی موجودات تراریخته"

توسط کاربر اصلاح شود. زمانی که ترکیب تست‌های غربالگری واقعی توسط کاربر انتخاب شد، GMOseek تمام آزمایشات اختصاصی رخداد مورد نیاز برای شناسایی حضور تمام رخدادهای ممکن در نمونه را نشان می‌دهد.

همان‌طور که در شکل ۴ مشاهده می‌شود، سمت چپ این پنل گونه‌ها، در قسمت میانی تمامی روش‌های غربالگری و در سمت راست محصولات تراریخته قرار دارند. روش‌های غربالگری حاصل از پنل جستجو نرم‌افزار به رنگ سیاه و سفید و روش‌های دیگر خاکستری رنگ دیده می‌شوند. وضعیت هر یک از روش‌ها را می‌توان با کلیک کردن بر روی دکمه سمت راست ماوس و تیک زدن آن‌ها تغییر داد. این پنل یک آزمون واقعی را بر اساس روش انتخاب شده نشان خواهد داد که نتایج این بررسی در قسمت Check consistency قابل مشاهده است.

هستند. در هر یک از خطوطی که ترکیبی از N و P نوشته شده است، چهار داده مشاهده می‌شود که به ترتیب مربوط به احتمال رخداد یک ترکیب (احتمال رخداد NN، PP، PN و NP)، تعداد آزمون‌های مربوط به یک رخداد خاص که باید تست شود، هزینه چنین آزمونی برای یک رخداد خاص و هزینه نرمال شده‌ی این آزمون توسط نرم افزار می‌باشد. در خط بعدی probability sum: 100% نشان دهنده‌ی این است که تمام ترکیبات در نظر گرفته شده است.

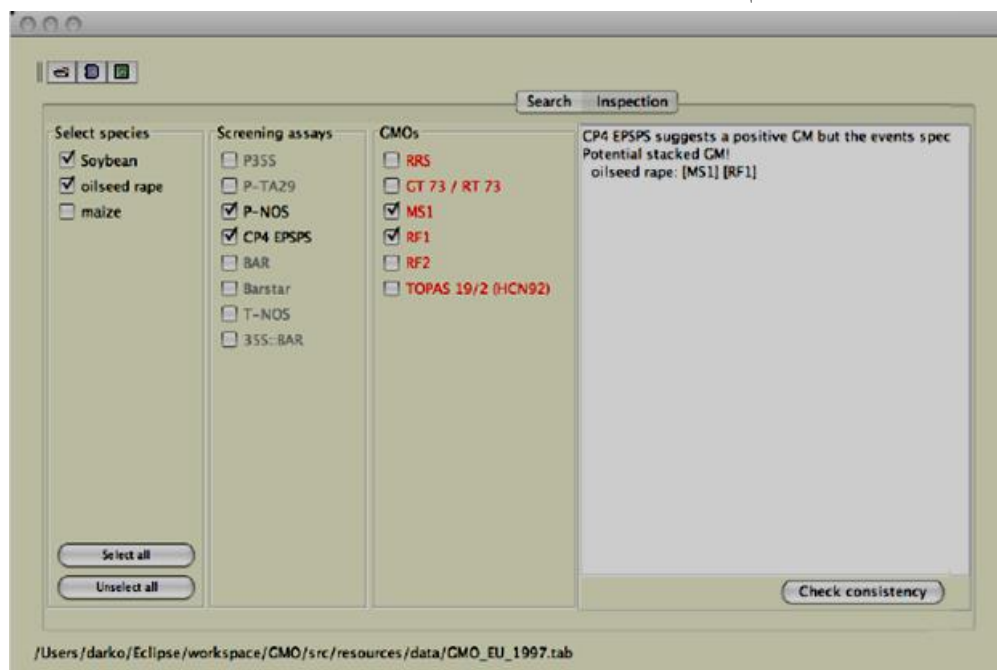
قسمت بعدی نرم‌افزار، پنل Inspection (به معنی بازرسی و معاینه) می‌باشد که در نوار ابزار بالا در کنار پنل Search قابل مشاهده است. این قسمت برای بازرسی داده‌های ورودی و آزمایش سازگاری و یا عدم سازگاری روش انتخاب شده برای تشخیص گروهی چندین رخداد می‌باشد. در پانل بازرسی، کاربر می‌تواند ترکیبی از روش به دست آمده در مرحله جستجو را انتخاب نماید. این ترکیب می‌تواند



شکل ۴- پنل Inspection

رنگ) بنابراین P-NOS و CP4 EPSPS که اجزای این روش غربالگری هستند، باید بازرسی شوند.

برای مثال برای روشی که درنتایج پنل Search به صورت PP نمایش داده شده است، احتمال رخداد ترکیب ۰/۰۰۰۱۲ است (شکل سوم کادر صورتی



شکل ۵- بازرسی به وسیله پنل Inspection

نتیجه‌گیری

برای آنالیز محصولات تراریخته عمدتاً محققان به دنبال روشی هستند که علاوه بر راحتی انجام آن، دقت و صحت نتایج را نیز تضمین کند. نرم‌افزار GMOseek با استفاده از راهبرد جستجوی پیشرفته، در عرض مدت زمان بسیار کوتاهی، مطلوب‌ترین استراتژی مقرون به صرفه را برای یک مجموعه داده بزرگ، به روش مبتنی بر ماتریس، با استفاده از یک کامپیوتر معمولی محاسبه می‌کند. این نرم‌افزار با استفاده از سیستم تصمیم‌گیری پشتیبان، امکان انتخاب روش غربالگری، اتخاذ تصمیم برای تست‌های مربوط به یک رخداد و تفسیر نهایی نتایج را به کاربر می‌دهد. با بررسی‌های انجام شده بر روی ابزار GMOseek در

مطابق شکل ۵ بعد از انتخاب P-NOS و CP4 EPSPS تمام محصولات تراریخته‌ای که به رنگ قرمز نشان داده شده‌اند، باید با آزمون اختصاصی رخداد بازرسی شوند. برای مثال در این جا MS1 و RF1 انتخاب و در قسمت Check consistency (به معنی بررسی سازگاری)، هشدارهای مربوط به تست بازرسی نشان داده شده است. اولین هشدار تناقض بین نتایج آزمون غربالگری و آزمون اختصاصی رخداد و دومین هشدار پتانسیل گونه گیاهی برای داشتن محصول تراریخته را نشان می‌دهد. مثلاً MS1 و RF1 هر دو محصولات تراریخته‌ی مربوط به یک گونه (کلزا) هستند.

"زکوی و توحیدفر، آشنایی با نرم افزار GMOseek به منظور شناسایی و بررسی موجودات تراریخته"

آزمایشگاه مرجع (TestLab) مشخص شد که این
نرم افزار ظرفیت بالایی را برای کاهش هزینه ها و
سهولت آنالیز محصولات تراریخته دارد.

References

فهرست منابع

- 1- James, C., 2012. *Global Status of Commercialized Biotech/GM Crops: 2012: Executive Summary*. International Service for the Acquisition of Agri-biotech Applications (ISAAA).
- 2- Stein, A.J. and Rodríguez-Cerezo, E., 2010. International trade and the global pipeline of new GM crops. *Nature Biotechnology*, 28(1), p.23.
- 3- Bonfini, L., 2002. *Report on GMO detection identification and quantification methods submitted to collaborative studies*. Institute for Health and Consumer Protection, Food Products and Consumer Goods Unit.
- 4- Holst-Jensen, A., Rønning, S.B., Løvseth, A. and Berdal, K.G., 2003. PCR technology for screening and quantification of genetically modified organisms (GMOs). *Analytical and Bioanalytical Chemistry*, 375(8), pp.985-993.
- 5- Holst-Jensen, A., 2009. Testing for genetically modified organisms (GMOs): Past, present and future perspectives. *Biotechnology advances*, 27(6), pp.1071-1082.
- 6- Žel, J., Milavec, M., Morisset, D., Plan, D., Van den Eede, G. and Gruden, K., 2012. How to reliably test for GMOs. In *How to Reliably Test for GMOs* (pp. 1-95). Springer, Boston, MA.
- 7- Querci, M., Kleter, G., Malingreau, J.P., Broll, H. and Van den Eede, G., 2008. Scientific and technical contribution to the development of an overall health strategy in the area of GMOs. *Luxembourg: Office for Official Publications of the European Communities, EUR*, 23542.
- 8- Block, A., Debode, F., Grohmann, L., Hulin, J., Taverniers, I., Kluga, L., Barbau-Piednoir, E., Broeders, S., Huber, I., Van den Bulcke, M. and Heinze, P., 2013. The GMOseek matrix: a decision support tool for optimizing the detection of genetically modified plants. *BMC bioinformatics*, 14(1), p.256.
- 9- Holst-Jensen, A., Bertheau, Y., de Loose, M., Grohmann, L., Hamels, S., Hougs, L., Morisset, D., Pecoraro, S., Pla, M., Van den Bulcke, M. and Wulff, D., 2012. Detecting un-authorized genetically modified organisms (GMOs) and derived materials. *Biotechnology advances*, 30(6), pp.1318-1335.
- 10- Holst-Jensen, A., Bertheau, Y., Allnutt, T., Broll, H., De Loose, M., Grohmann, L., Henry, C., Hougs, L., Moens, W., Morisset, D. and Ovesna, J., 2011. Overview on the detection, interpretation and reporting on the presence of unauthorised genetically modified materials. *EUR 25008 EN*.
- 11- Morisset, D., Novak, P.K., Zupanič, D., Gruden, K., Lavrač, N. and Žel, J., 2014. GMOseek: a user friendly tool for optimized GMO testing. *BMC bioinformatics*, 15(1), p.258.

Study of GMOseek software in order to detection of genetically modified organisms

Maryam Zakavi¹, Masoud Tohidfar^{*2}

¹PhD graduated, Department of Biotechnology, Faculty of Engineering and New Technologies, Shahid Beheshti University, Tehran, Iran

^{2*}Associate Professor, Department of Biotechnology, Faculty of Engineering and New Technologies, Shahid Beheshti University, Tehran, Iran

M_Tohidfar@sbu.ac.ir

Abstract

Due to the increasing speed of approval and commercialization of GM crops around the world, matching tests for food, feed and seed samples according to relevant regulations will be very difficult and very expensive. Many of them have already adopted the so called "matrix approach" to rationalize the resources and efforts used to increase their efficiency within a limited budget. Most of the time, the "matrix approach" is implemented using limited information and some proprietary computational tool to efficiently use the available data. The tool makes use of a tabulated matrix of GM events and their genetic elements, of the laboratory analysis history and the available information about the sample at hand. The tool uses an optimization approach to suggest the most suited screening assays for the given sample. The practical GMOseek user interface allows the user to customize the search for a cost-efficient combination of screening assays to be employed on a given sample. It further guides the user to select appropriate analyses to determine the presence of individual GM events in the analyzed sample.

Keywords: Genetically Modified Organism, Matrix approach, Constraint optimization, Cost efficient GMO testing