

آنالیز متابولومیکس با استفاده از درخت تصمیم‌گیری (Decision tree) جهت تمایز بین گیاهان برنج بدون آلودگی و تحت تنش دیازینون

وحیده مهدوی^{۱*}، محمدزمان نوری^۲

۱. بخش تحقیقات آفت‌کش‌ها، موسسه تحقیقات گیاه‌پزشکی کشور، سازمان تحقیقات، آموزش و ترویج کشاورزی، تهران، ایران. ۲. موسسه تحقیقات برنج ایران - شعبه مازندران، سازمان تحقیقات، آموزش و ترویج کشاورزی، آمل، ایران.

تاریخ دریافت: ۱۳۹۹/۴/۰۱

تاریخ پذیرش: ۹۹/۱۰/۲۵

چکیده:

ارائه یک روش طبقه‌بندی موثر در جهت تمایز بین گیاهان بدون آلودگی و نمونه‌های قرار گرفته در معرض آفت‌کش، می‌تواند برای افزایش کنترل کیفیت در محصولات کشاورزی و ایمنی مواد غذایی بسیار مهم و کارآمد باشد. روش پیشنهادی در این مطالعه، الگویی برای تمایز بین نمونه‌های گیاه برنج (*Oryza sativa* L.) بدون آلودگی و تیمار شده با حشره‌کش دیازینون است. بنابراین هدف از این مطالعه، بررسی تغییرات متابولیتی رقم شیروودی برنج (رایج‌ترین رقم در ایران) در پاسخ به حشره‌کش دیازینون و ارزیابی پویا (دینامیکی) تغییرات سطوح غلظتی متابولیت‌ها در طول زمان است. برای این منظور، پس از تیمار گیاه برنج با دیازینون، متابولیت‌های غیر فرار گیاه با مشتق‌سازی به فرم فرار و قابل آنالیز با کروماتوگرافی گازی درآمد و پروفایل متابولیتی حاصل، با استفاده از کروماتوگرافی گازی-طیف‌سنجی جرمی (GC-MS) مورد بررسی قرار گرفت. پیرو شناسایی متابولیت‌های گیاه برنج، دارای بیش‌ترین تغییر در معرض تنش دیازینون؛ برای نخستین بار، با استفاده از تجزیه درخت طبقه‌بندی و رگرسیون (CART)، برای مدل‌سازی و پیش‌بینی تغییرات متابولیت‌های گیاه برنج تحت تنش دیازینون استفاده شد. نتایج آنالیز نشان داد بیش‌ترین تغییرات، مربوط به اسیدهای آمینه، به‌ویژه فنیل‌آلانین و پرولین بوده است. هم‌چنین نتایج حاکی از این است که CART می‌تواند ابزار بسیار قدرتمندی برای تمایز بین گیاهان در معرض آفت‌کش و گیاهان به دور از تنش باشد.

واژه‌های کلیدی: متابولومیکس، دیازینون، برنج، درخت طبقه‌بندی و رگرسیون.

مقدمه:

فن آماری است که از میان تعداد بسیار زیاد متغیرهای (x)، متغیرهای دارای بیشترین تأثیر روی پاسخ (y) را انتخاب می‌کند. این عمل پیرو رشد ساختار درختی انجام می‌شود و داده‌ها را به گروه‌های انحصاری (گروه) تقسیم می‌کند که بخش‌هایی کاملاً یکنواخت از نقطه‌نظر پارامتر خروجی هستند. در واقع چنین درختی از یک ریشه شامل تمامی نمونه‌های موجود، آغاز به رشد کرده و سپس با شاخه‌شاخه شدن دوتایی، به گره‌های دیگر رشد پیدا می‌کند.

بسته به ماهیت داده‌ها، می‌توان این روش را برای رگرسیون یا طبقه‌بندی داده‌ها استفاده کرد. اصطلاح درخت، به دلیل شکل نمایش نتایج حاصل از تجزیه و تحلیل آماری، به شکل درخت باینری اطلاق می‌شود.

هدف از مطالعه حاضر، بررسی تغییرات متابولیتی گیاه برنج (*Oryza sativa* L.) در پاسخ به حشره‌کش دیازینون و ارزیابی دینامیکی تغییرات سطوح غلظتی متابولیت‌ها در طول زمان و با استفاده از روش پروفایل متابولیتی مبتنی بر GC-MS است. در این راستا از الگوریتم CART، به عنوان یک روش طبقه‌بندی قدرتمند استفاده شد.

مواد و روش‌ها:

مواد

استاندارد دیازینون و متوکسی آمین هیدروکلراید^۳ از شرکت Sigma-Aldrich (Fluka, St. Louis, MO, USA) استاندارد سوریتول،^۴ MSTFA^۵، TMCS^۶ و تولوئن^۶ از شرکت Merck (Darmstadt, Germany) تهیه شدند. متانول و آب مورد نیاز در فرایند استخراج با درجه خلوص HPLC، از شرکت Merck (Darmstadt, Germany)

گیاه برنج (*Oryza sativa* L.)، یکی از مهم‌ترین محصولات زراعی در سراسر جهان است که منبع اصلی انرژی، پروتئین و سایر مواد مغذی نیمی از جمعیت جهان را تأمین می‌کند (Kim et al., 2012). با این حال، عملکرد تولید این محصول توسط عوامل مختلفی محدود می‌شود، به‌طور عمده تنش‌های حاصل از عوامل زیستی مانند آفات از مهم‌ترین موانع تولید برنج در جهان به‌شمار می‌روند. یکی از روش‌های سنتی برای مدیریت آفات برنج استفاده از حشره‌کش‌ها است (Zhang et al., 2011). علی‌رغم ممنوعیت تدارک دیازینون از سال ۹۸ و مصرف آن از سال ۱۴۰۰، توسط سازمان حفظ نباتات کشور، این آفت‌کش هم‌چنان در مزارع برنج علیه طغیان کرم ساقه-خوار مصرف می‌شود. از این رو با توجه به اثرات ثابت شده سمیت دیازینون و تأثیر آن به عنوان القاکننده تنش اکسیداتیو، می‌تواند سهم بسزایی در بروز عوارض ناخواسته آفت‌کش‌های مصرفی داشته باشد (Köprücü et al., 2006).

تحقیقات مدرن نظیر متابولومیکس در علوم و صنایع غذایی در حال گذار از روش‌های کلاسیک به سمت استراتژی‌های پیشرفته تجزیه‌ای با کاربرد تکنیک‌های مبتنی بر اسپکترومتری جرمی است (Fiehn, 2002). از روش متابولومیکس برای بررسی تغییرات غیرقابل پیش‌بینی متابولیکی ناشی از تنش‌های محیطی استفاده می‌شود. این زمینه تحقیقاتی، یکی از فن‌آوری‌های بسیار مهم و جذاب به‌ویژه در حوزه سم‌شناسی محیطی است (Kim et al., 2007).

الگوریتم درخت طبقه‌بندی و رگرسیون (CART)، یکی از روش‌های یادگیری نظارتی است که نخستین بار توسط بریمن ارائه شد (Breiman et al., 1986). این روش یک

^۴ N-methyl-N-(trimethylsilyl)trifluoroacetamide

^۵ Trimethylchlorosilane

^۶ Toluene

^۱ L Breiman

^۲ Binary tree

^۳ Methoxyamine hydrochloride

بندی شده و به مخزن نیتروژن مایع منتقل شدند و پس از توقف فرایندهای متابولیکی، تا زمان شروع آزمایش در فریزر دمای ۸۰- درجه سانتیگراد نگهداری شدند.

جهت استخراج متابولوم از گیاه برنج ابتدا همه نمونه در هاون به کمک نیتروژن مایع کاملاً پودر و یکنواخت گردید. ۵۰mg از نمونه خشک شده توزین گردید. مراحل استخراج باید سریع و در دمای پایین انجام گیرد. به این مقدار نمونه ۱۴۰۰µl متانول ۱۰۰٪ خالص در دمای ۴°C اضافه شد. سپس ۵۰µl سوربیتول به عنوان استاندارد داخلی به نمونه اضافه شد. سوربیتول با غلظت ۲mg در ۱ml آب مقطر تهیه شده بود. این مخلوط به مدت ۱۵min در دمای ۷۰°C و رتکس شد. در ادامه به محلول فوق ۱۴۰۰µl آب مقطر ۴°C افزوده و در دور موتور ۴۵۰۰rpm به مدت ۵min سانتریفوژ شد. ۱ml از محلول که مخلوطی از آب و متانول است به میکرولوله سانتریفوژ منتقل شد و در آن خلأ و دمای ۳۰°C به مدت ۶ الی ۱۶ ساعت خشک شد.

برای مشتق سازی متابولیت ها مانده خشک شده از مرحله قبل در ۸۰µl از متوکسی آمین هیدروکلراید حل شد. ۲۰µl از متوکسی آمین هیدروکلراید در ۱ml پیریدین بی آب حل شده بود. علت اضافه کردن این ترکیب حفظ ساختار خطی قندهای کاهنده است. مخلوط به مدت ۹۰min در دمای ۳۰°C قرار داده شد تا واکنش تکمیل گردد. به منظور مشتق سازی متابولیت ها، به هر نمونه ۸۰µl از MSTFA اضافه شد. جهت افزایش بازده فرایند مشتق سازی (Kind et al., 2009)، ۰/۸µl، TMCS معادل ۱٪ از MSTFA اضافه شد و به مدت ۳۰min در دمای ۳۷°C به منظور اطمینان از تکمیل فرایند مشتق سازی نگهداری شد (Roessner et al., 2000).

دستگاه کروماتوگرافی گازی-اسپکترومتری جرمی (GC-MS)

برای مطالعه مشتقات متابولیت ها از کروماتوگرافی گازی CP-3800 شرکت Varian مجهز به ستون DB-5 به طول

Germany) تهیه شدند. حشره کش فرموله شده دیازینون (G 10%) از شرکت گل سم گرگان تهیه شد.

تهیه نمونه های برنج

به منظور آماده سازی نمونه ها برای فراهم ساختن توده مناسب برگی برای استخراج متابولیت ها، جوانه های کشت شده بذور رقم شیرودی به محیط گلدان های خاک منتقل شدند. جوانه ها پس از انتقال به ۲۰ گلدان سفالی، روزانه در دو نوبت صبح و عصر آبیاری شدند. رطوبت حدود ۷۰٪ و دمای نسبی گلخانه در روز ۲۷ درجه سانتیگراد کنترل می شد.

تنش حشره کش دیازینون بر روی گیاه برنج

دز توصیه شده سازمان حفظ نباتات جهت مصرف دیازینون گرانوله ۱۰٪ در مزارع برنج، ۱۵ kg/ha است (Arshad et al., 2006) که با در نظر گرفتن قطر ۲۰cm برای هر گلدان، مساحت هر گلدان ۰/۳۱۴m^۲ می شود که با احتساب ۱۵ Kg به ازای هر ۱۰۰۰۰ m^۲، ۰/۴۷۱gr سم گرانوله دیازینون توزین و در هر یک از گلدان ها ریخته شد. از آنجایی که زمان طغیان آفت کرم ساقه خوار برنج، در مرحله پنجه زنی برنج است و این مرحله از رشد در فاصله تقریبی ۵۰ روز پس از کشت نشاء اتفاق می افتد، زمان سم پاشی گیاهان در گلخانه نیز ۵۰ روز پس از کشت انتخاب شد تا بررسی ها و نتایج به شرایط واقعی که در مزارع اتفاق می افتد نزدیک تر باشد. پس از افزودن ۰/۴۷۱gr گرانول دیازینون در گلدان های حاوی آب، نمونه برداری در زمان های مشخصی انجام گرفت.

نمونه برداری جهت مطالعات متابولومیکس

برای مطالعات متابولومیکس، تا پنج روز پس از افزودن آفت کش، نمونه برداری روزانه با هفت تکرار انجام شد و نمونه برداری به موازات گلدان های تیمار شده، در سه تکرار از گلدان های شاهد انجام گرفت. در مرحله نمونه برداری برای مطالعه متابولومیکس، نمونه ها بلافاصله پس از جدا شدن از خاک، در ورق های آلومینیومی بسته

رفته است، روش هم‌مرکز کردن می‌باشد، هم‌مرکز کردن فرایندی است که برای نمونه‌های متابولومیکس بکار می‌رود و شامل روش‌هایی است که توانایی مقایسه مستقیم همه نمونه‌ها را با یکدیگر می‌دهد. کاربرد رایج این روش در برداشتن و یا کاهش اثرات متغیرهای تصادفی و یا عدم تکرارپذیری دستگاه در اثرات نوسانات برقی و نظایر آن، در هر نمونه است.

این فرایند بر روی ستون‌ها انجام می‌گیرد. هم‌مرکز کردن همه غلظت‌ها را به نوسانات حدود صفر تبدیل می‌نماید. برای هم‌مرکز کردن، مقدار میانگین هر متغیر محاسبه می‌شود و سپس این مقدار از داده‌ها کم می‌شود. بدین ترتیب میانگین هر ستون از ماتریس صفر می‌شود. این عملیات، تفسیر راحت‌تری را در اختیار می‌گذارد. در این بررسی پردازش طیف‌های GC-MS جهت مطالعات متابولومیکس، به روش هم‌مرکز کردن انجام شد.

نرم‌افزارهای مورد استفاده

MATLAB: نرم‌افزار متلب یک زبان سطح بالا می‌باشد، که در ابتدا بر اساس زبان برنامه‌نویسی C توسعه داده شد. واژه متلب هم به معنی محیط محاسبات رقیمی و هم به معنی خود زبان برنامه‌نویسی مربوطه است که از ترکیب دو واژه ماتریس و آزمایشگاه ایجاد شده است. این نام حاکی از رویکرد ماتریس محور برنامه است، که در آن حتی اعداد منفرد هم به عنوان ماتریس در نظر گرفته می‌شوند. اجرای برنامه CART در محیط این برنامه انجام گرفت.

AMDIS²: چالش اصلی در تحلیل داده‌های حاصل از GC-MS به خصوص در محیط‌های پیچیده، تشخیص پیک‌های مربوط به ترکیبات موردنظر از نویز زمینه است. راه‌حل استفاده از الگوریتمی برای استخراج پیک‌های خالص از

۳۰m قطر داخلی ۰/۲۵mm و ضخامت فیلم ۰/۲۵µm متصل به اسپکترومتری جرمی به تله یونی 4000 MS-Ion Trap با قابلیت مطالعه در دامنه جرمی ۴۰-۱۰۰۰ Da استفاده شد. دمای محفظه تزریق ۲۵۰°C و دمای منبع یونی ۲۰۰°C تنظیم شد. برنامه دمایی مورد استفاده به این صورت بود که به مدت ۵ min در دمای ۷۰°C ثابت ماند و سپس با شیب ۵°C/min تا دمای ۳۱۰°C افزایش یافت و به مدت ۱۰ min در این دما ثابت ماند لذا زمان آنالیز ۵۸min بود. گاز هلیوم به عنوان گاز حامل، با فلوی ۱ml/min استفاده شد و اسکن پیک‌ها در محدوده جرمی ۵۰-۶۵۰m/z انجام گرفت. مطالعات تکمیلی با همین شرایط دمایی و ستون با استفاده از دستگاه GC-MS کمپانی Agilent مدل 6890N با آنالیزر جرمی Quadropole انجام گرفت.

برای پردازش داده‌های GC-MS از VARIAN MSD WorkStation به قالب netCDF صادر شد. برای پردازش داده‌ها از نرم‌افزار AMDIS 2.6 (سیستم خودکار شناسایی طیفی انبوه و شناسایی سیستم) استفاده شد. شناسایی قله با استفاده از کتابخانه طیفی جرمی استاندارد (NIST 8.0) و ویلی (۰/۶) انجام شد.

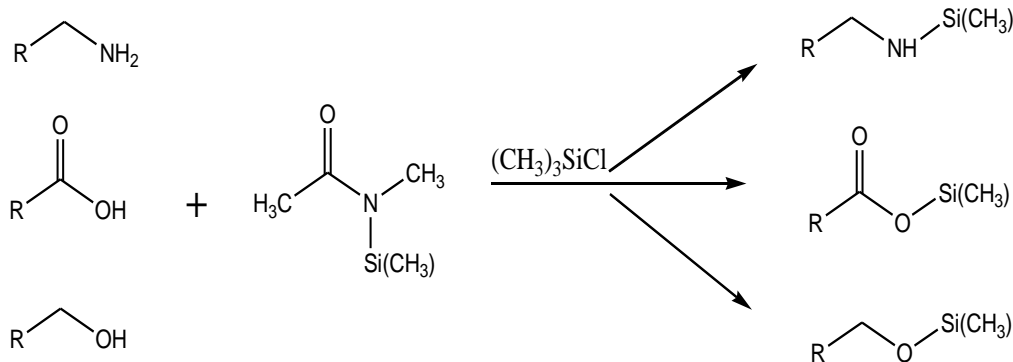
پردازش داده‌ها، مرحله میانی بین طیف خام و تجزیه و تحلیل داده‌ها است. هدف از پردازش داده‌ها تبدیل آن‌ها به فرمی است که بتواند به وسیله روش‌های آماری مورد تجزیه و تحلیل قرار بگیرد. داده‌ها به جداول یا ماتریس‌هایی تبدیل می‌شوند که ردیف آن‌ها مربوط به یک نمونه مجزای تحلیلی (نمونه‌ها در روزهای مختلف پس از تیمار) و هر ستون مربوط به سیگنال (سطح زیر پیک در زمان‌های مختلف بازداري یا شدت شکست‌های جرمی) یک متابولیت می‌باشد. پردازش طیف‌های جرمی داده‌های متابولومیکس به روش‌های مختلفی انجام می‌شود که مهم‌ترین و متداول‌ترین آن‌ها که در این بررسی بکار

² Automated Mass spectral Deconvolution and Identification System

¹ Mean Centering

نتایج:

۱ استخراج متابولیت‌ها و تزریق به دستگاه GC/MS: در این بررسی، هدف پایش متابولیت‌های قطبی گیاه برنج با استفاده از دستگاه GC-MS بود. تنها ترکیبات فراری که دارای فشار بخار بالایی هستند توسط این دستگاه قابل شناسایی هستند. از آنجایی که همه متابولیت‌های قطبی از این ویژگی برخوردار نیستند، این محدودیت را می‌توان با مشتق‌سازی متابولیت‌ها به ترکیبات فرار، برطرف نمود. بدین منظور از بین روش‌های مختلف مشتق‌سازی، روش مشتق‌سازی سیلیل‌دار کردن انتخاب شد. از بین واکنش‌گرها نیز، واکنش‌گر مشتق‌ساز MSTFA، با بیش‌ترین بازده در مشتق‌سازی برگزیده شد (Tsugawa *et al.*, 2014; Frank *et al.*, 2012). شکل (۱) شمایی از فرایند مشتق‌سازی سیلیل‌دار شدن اسیدهای آمینه، اسیدهای آلی و قندها را نشان می‌دهد.



شکل ۱- واکنش‌های مشتق‌سازی سیلیل‌دار کردن اسیدهای آمینه، اسیدهای آلی و قندها.

Fig. 1. Silanization derivative reactions of amino acids, organic acids and sugars.

برنج بدون استاندارد داخلی استفاده شد تا با مقایسه نمونه‌های تیمار شده حاوی استاندارد داخلی، از عدم همپوشانی طیفی و مزاحمت استاندارد داخلی در شناسایی متابولیت‌ها اطمینان حاصل شود.

- استفاده از نمونه تنها شامل استاندارد داخلی با طی کردن

کروماتوگرام مجموعه یون‌ها است. AMDIS نرم‌افزاری است که بر این اساس کار می‌کند و طیف‌های خالص ترکیبات و اطلاعات وابسته به آن مانند شکل پیک و زمان بازداری را از کروماتوگرام پیچیده استخراج کرده و با طیف‌های موجود در کتابخانه مرجع که کتابخانه هدف^۲ نیز نامیده می‌شود، تطبیق می‌دهد. AMDIS اولین بار در سال ۱۹۹۶ برای شناسایی اتوماتیک سلاح‌های شیمیایی و ترکیبات وابسته به آن طراحی شد. اما امروزه برای آنالیز داده‌های طیف‌سنجی جرمی کروماتوگرام مجموعه یون‌ها و شناسایی ترکیبات موردنظر نیز بکار می‌رود.

همه طیف‌های جرمی به وسیله نرم‌افزار AMDIS نسخه ۲/۶۶ از مؤسسه ملی استانداردها و فن‌آوری^۳ (NIST) مورد بررسی قرار گرفتند. پیک‌های با نسبت سیگنال به نویز (S/N) کمتر از ۱۰ حذف شدند.

تهیه ماتریس داده‌ها: قبل از شروع پردازش داده‌ها و سپس آنالیز آن‌ها با الگوریتم‌های مختلف، لازم بود در ابتدا ماتریس داده‌ها جمع‌آوری گردد. به منظور تهیه ماتریس داده، توجه به برخی موارد ضروری است:

- نمونه کنترل بدون استاندارد داخلی: از نمونه‌های کنترل

³ National Institute of Standards and Technology

¹ Total Ion Chromatogram

² Target Library

داخلی همان نمونه تقسیم شد و سپس هم مرکز کردن انجام گرفت. برای هم مرکز کردن همه نمونه‌ها، ابتدا میانگین هر تکرار محاسبه شد. میانگین‌های به دست آمده، به تک تک اعداد تقسیم شدند بدین ترتیب دامنه همه اعداد بین صفر و یک شد و از پراکندگی داده‌ها به شدت کاسته شد.

آنالیز داده‌ها با استفاده از CART

ابتدا در فضای MATLAB، ۷۰٪ داده‌ها به عنوان مجموعه آموزش و ۳۰ درصد باقیمانده به عنوان مجموعه تست مدل استفاده شدند. این الگوریتم به منظور وجود تنوع یکسان در دو مجموعه تست و آموزش بکار گرفته می‌شود. در گام اول، این الگوریتم، دو نمونه که بیشترین فاصله اقلیدسی^۵ را از هم دارند انتخاب می‌کند. سپس دو نمونه دیگر با بیشترین فاصله اقلیدسی در گروه تست قرار داده می‌شوند. این مراحل تا آنجا که تمامی نمونه‌ها در بین دو مجموعه توزیع گردد، ادامه می‌یابد (Snee, 1977).

نمونه‌های مجموعه training با استفاده از CART طبقه بندی شدند. در مدل ایجاد شده، متابولیت‌ها به عنوان توصیف کننده و طبقه یا گروه هر نمونه یعنی کنترل یا تیمار بودن هر نمونه به عنوان پاسخ در نظر گرفته شد، دیاگرام CART در شکل (۲) نشان داده شده است. متغیرهای شماره ۴۹ و ۵۰، به ترتیب فنیل آلانین و پرولین بودند که نحوه شناسایی آن‌ها شرح داده خواهد شد. همان‌طور که در شکل مشاهده می‌شود اگر سطح متابولیت فنیل آلانین کمتر از ۱/۱۱۶۲۴ باشد، نتیجه مربوط به نمونه کنترل است و اگر بالاتر از ۱/۱۱۶۲۴ باشد دو حالت اتفاق می‌افتد، اگر سطح پرولین کمتر از ۰/۴۸۴۹۶۵ باشد مربوط به نمونه کنترل و اگر بیشتر از این مقدار باشد مربوط به نمونه تیمار است.

کامل روش: برای تعیین زمان بازداری tr استاندارد داخلی و اطمینان از مشتق شدن آن انجام شد علاوه بر این که استاندارد داخلی تکرارپذیری روش آماده سازی نمونه، استخراج، تزریق‌های متوالی و عملکرد دستگاه را نشان می‌دهد.

• نمونه‌های کنترل: شامل نمونه‌های برنج جمع آوری شده در روزهای مختلف بود که سم پاشی نشده و در روزهای مختلف همزمان با نمونه‌های تیمار شده جمع آوری شدند، نتایج این مقایسه‌ها برای اطمینان از عدم تأثیر متغیر روز در بین نمونه‌های تیمار استفاده شد.

• تکرار تزریقات: برای اطمینان از تکرارپذیری دستگاه در طی جمع آوری داده‌ها انجام شد که سنجش نویز^۲ را نشان می‌دهد.

• توجه به یون‌های موجود در پیک زمینه: ستون‌های کروماتوگرافی که فاز ساکن آن‌ها بر مبنای سیلیل^۳ می‌باشد مستعد یون‌های زمینه ۷۳، ۱۴۷، ۲۰۷ و غیره می‌باشند که توجه به این یون‌ها برای جلوگیری از خطای مثبت کاذب ضروری می‌باشد.

با لحاظ نمودن نتایج حاصل از بررسی‌های فوق و مقایسه آن‌ها جهت استخراج اطلاعات مرتبط با متابولیت‌ها، می‌توان ماتریس مورد نظر برای مرحله بعد را آماده نمود.

پردازش داده‌ها

همه داده‌های به دست آمده در یک ماتریس جمع آوری شدند، ستون‌های این ماتریس نمونه‌های مختلف کنترل و تیمار در روزهای مختلف و ردیف‌های ماتریس زمان‌های بازداری تشخیص داده شده توسط دستگاه با $S/N=10$ بود و درایه‌ها، سطح زیر پیک در هر یک از زمان‌های بازداری را نشان می‌دادند.

ابتدا سطح زیر پیک هر نمودار به سطح زیر پیک استاندارد

⁴ Elements

⁵ Euclidean Distance

¹ Repeat Injection

² Drift

³ Silyl

جدول ۱- ویژگی توصیف گرهای انتخاب شده توسط CART (مهم ترین مشتقات TMS متابولیت های قطبی در برنج).

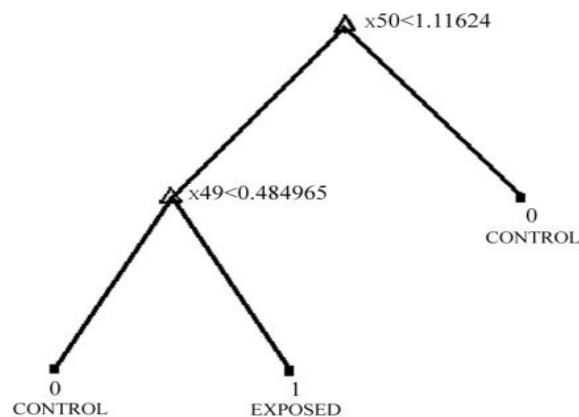
Table1. Specifications of the selected CART descriptors (The most important TMS-derivatives of polar metabolites in rice)

Descriptors	Assignment	RT	RRT	Mass fragment
phenylalanine N,O	C ₉ H ₁₁ NO ₂	20.956	0.971	73,102,103
proline N,O	C ₅ H ₉ NO ₂	20.861	0.967	73,144,218

RT: Retention Time (min)

RRT: Relative Retention Time (retention time of analyte/retention time of sorbitol)

Mass fragment: The first three ions with highest intensities. Ions in bold mean the most intense product ion.



شکل ۲- درخت رگرسیون بدست آمده برای ۳۵ نمونه با دو توصیف گر متغیرهای X₄₉ و X₅₀.

Fig. 2. Regression tree trained on a sample data set consisting of 35 training cases with 2 predictor variables X₅₀ and X₄₉.

متابولیت ها آمده است. در مرحله دوم جداسازی برای ارزیابی مدل ایجاد شده، مجموعه تست توسط مدل طبقه بندی می شوند. مشخصات ماتریس کانفیوژن برای این مدل سازی در جدول (۲) نشان داده شده است.

با توجه به جدول (۱)، متغیرهای فنیل آلانین و پرولین مهم ترین متابولیت هایی هستند که دو دسته کنترل و تیمار را از همدیگر جدا می کند. در این جدول زمان بازداری هر دو متابولیت و نسبت زمان بازداری آن ها نسبت به زمان بازداری استاندارد داخلی سوربیتول و شکست های جرمی شاخص به عنوان فاکتورهای اصلی تعیین و تشخیص نوع

جدول ۲- کانفیوژن ماتریس برای مجموعه آموزش و تست.

Table 2. Confusion matrix for training and test sets.

	Observed	Predicted	
		Exposed class	Control class
Training set	Exposed class	26	0
	Control class	0	9
Test set	Exposed class	8	1
	Control class	0	5

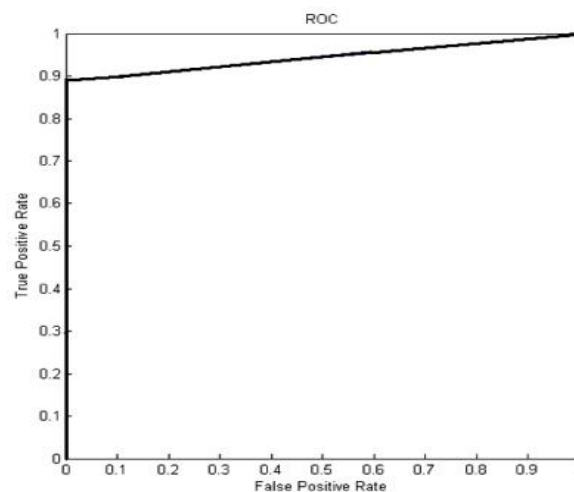
با توجه به جدول (۳) صحت و حساسیت برای مجموعه تست به ترتیب ۹۳ و ۱۰۰ درصد می‌باشند. سایر پارامترهای طبقه‌بندی در این جدول دیده می‌شود. این نتایج نشان می‌دهند که CART می‌تواند شانس بزرگی در تشخیص نمونه‌های بیمار نشده و نمونه‌های تحت تنش قرار گرفته باشد.

جدول ۳- سرعت خطا و سرعت غیر خطا و سایر پارامترهای طبقه‌بندی برای گروه آموزش و تست.

Table 3. Calculated error and non-error rates of the classification index and the classification performances of training and test sets.

	Error rate	Non-error rate	Specificity	Sensitivity	Accuracy
Training set	0	1	1	1	1
Test set	0.07	0.93	0.83	1	0.93

منحنی ROC نیز برای مجموعه آموزش و تست رسم گردید و مساحت زیر منحنی برای هر دو دسته علت دیگری برای بهینه بودن مدل CART در تشخیص نمونه‌های کنترل از نمونه‌های بیمار شده است. منحنی ROC برای گروه تست در شکل (۳) نشان داده شده است.

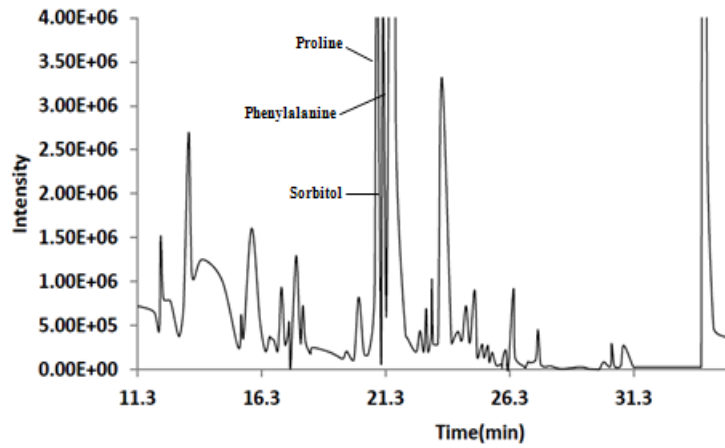


شکل ۳- منحنی ROC حساسیت (مثبت واقعی) بر حسب گزینش پذیری (مثبت کاذب)

Fig. 3. Receiver Operating Characteristic (ROC) curve plotting sensitivity (true positive) against 1 – specificity (false positive).

مساحت زیر منحنی ROC، برابر با ۹۵ درصد دلیل دیگری بر این ادعاست که مدل CART می‌تواند به درستی گروه و طبقه نمونه‌ها را تشخیص دهد.

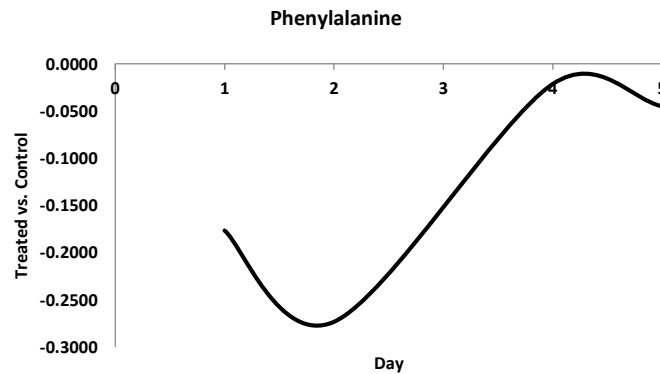
شکل (۴) دو متابولیت را نشان می‌دهد که به عنوان عمده‌ترین تغییرات برنج مطابق با CART شناسایی شدند. محاسبات کمی در رابطه با سطح اوج سطح آنالیت نسبت به استاندارد داخلی (IS) انجام شده است.



شکل ۴- کروماتوگرام یون انتخابی مشتقات TMS استخراج شده از گیاه برنج.

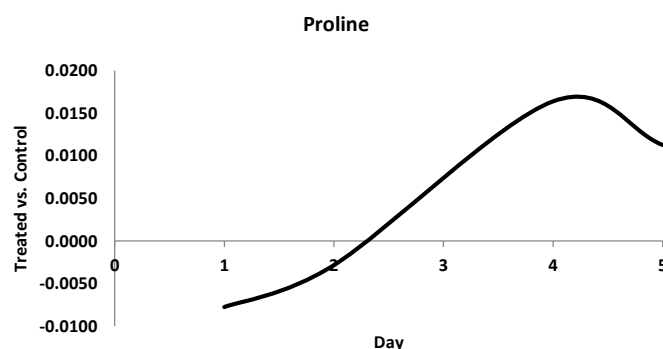
Fig. 4. Selected ion chromatograms of TMS-derivatives of metabolites extracted from rice (*Oryza sativa*).

تابش مربوط به اسیدهای آمینه بویژه فنیل آلانین و پرولین همان طور که از شکل (۵) و (۶) بر اساس نتایج CART بارز بدست آمده ، شدیدترین تغییرات در طول دوره بود.



شکل ۵- تغییرات بر حسب زمان فنیل آلانین (به عنوان اولین متابولیت مهم) تحت تاثیر معنی دار زمان.

Fig. 5. Time-course changes of phenylalanine (as the first most important metabolite) significantly affected by the time factor.



شکل ۶- تغییرات بر حسب زمان پرولین (به عنوان دومین متابولیت مهم) تحت تاثیر معنی دار زمان.

Fig. 6. Time-course changes of proline (as the second most important metabolite) significantly affected by the time factor.

ترنسفرز کنترل می‌شوند و یا سیستم‌های دفاعی آنتی‌اکسیدان غیر آنزیمی مانند α -توکوفرول، اسکوربیک اسید، رتینول، β -کاروتن، سرولوپلاسمین و GSH وارد عمل می‌شوند (Büyükkokuroğlu *et al.*, 2008).

هنگامی که سطح ROS در متابولیسم بیش از ظرفیت سیستم دفاعی آنتی‌اکسیدانی درون‌ریز می‌شود، توازن بحرانی بین اکسیدانت و آنتی‌اکسیدانت، مختل شده و آسیب القا شده توسط ROS افزایش می‌یابد. ROS می‌تواند با ماکرو مولکول‌های بیولوژیکی واکنش دهد و منجر به پراکسیداسیون لیپید، غیر فعال شدن آنزیم یا آسیب DNA شود. مشخصاً استرس اکسیداتیو مانند تنش آفت‌کش‌ها، تأثیر عمیقی بر روی چند مسیر متابولیکی مرکزی، شامل بخش‌های گسترده‌ای از قندها، اسیدهای آمینه و متابولیسم فنیل پروپانویید دارد که می‌تواند منجر به تغییرات متابولیکی وسیع شود.

شواهد بسیار زیادی وجود دارد که بسیاری از آنتی‌اکسیدان‌ها نقش کلیدی در تطابق گیاه با تنش‌های زنده و غیرزنده دارند. گیاهان به طور متداول گروه متنوعی از آنتی‌اکسیدان‌ها به عنوان مکانیسم حفاظتی در برابر ترکیبات اکسیداتیو که در پاسخ به تنش‌های مختلف ایجاد شده‌اند، تولید می‌کنند. بخش قابل توجهی از

بحث:

پاسخ‌های دو متابولیت اصلی با استفاده از اختلاف شدت گروه‌های تیمار با گروه کنترل اندازه‌گیری شد و مشخصات دوره زمانی آنها در طی دوره نمونه‌برداری نشان داده شد.

سموم ارگانوفسفره به عنوان بازدارنده استیل کولین استراز (AChE) شناخته شده‌اند، این آنزیم مسئول هیدرولیز سریع انتقال دهنده‌های عصبی است. در طی فرایند انتقال پیام عصبی در سیناپس‌های cholinergic، استیل کولین به کولین و استات تبدیل می‌شود (Wang *et al.*, 2009). از این رو سیستم‌های اکسیدانی و آنتی‌اکسیدانی (Büyükkokuroğlu *et al.*, 2008)، سطوح زیستی عنصر^۱ (Cemek *et al.*, 2010)، سیستم ایمنی (Neishabouri *et al.*, 2004)، سیستم تناسلی (Joshi *et al.*, 2003) و پارامترهای بیوشیمیایی می‌توانند تحت تأثیر سمیت ترکیبات ارگانوفسفره قرار گیرند. تحت شرایط طبیعی، تشکیل ROSها مانند اکسیژن تک‌آهیدروژن پروکساید، آنیون‌های سوپراکساید و رادیکال‌های هیدروکسیله در متابولیسم به‌وسیله آنزیم‌های درون‌ریز مانند سوپر اکسید دیسموتاز، گلوکاتیون پراکساید، کاتالاز، گلوکاتیون S-

² Single Oxygen

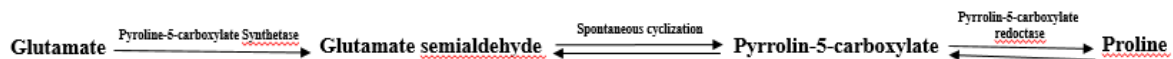
¹ Bio Element Level

به عنوان یک مولکول سیگنالینگ عمل می کند که می تواند عملکرد میتوکندری را تعدیل کرده و بر بیان ژن در هسته تأثیر بگذارد و از این طریق توانایی گیاه را در تحمل استرس افزایش دهد. در پاسخ به تنش های بی حسی در گیاهان، تجمع پرولین به طور معمول در سیتوزول اتفاق می افتد که در تنظیم اسمزی سیتوپلاسمی به طور قابل توجهی کمک می کند.

در گیاهان، بیوسنتز اسید ال گلوتامیک ماده اولیه تولید پرولین است. دو آنزیم پرولین-۵-کربوکسیلات سنتتاز (P5CS) و پرولین-۵-کربوکسیلات ردوکتاز (P5CR) نقش اصلی در مسیر بیوسنتز پرولین دارند (شکل ۷) (Bandurska, 2001).

آنتی اکسیدان هایی که در پاسخ به تنش به وسیله گیاه تولید می شوند، متابولیت های ثانویه هستند که گستره ای از ترکیبات فنلی ساده تا پیچیده را در بر می گیرند که در وهله اول از مسیر فنیل پروپانوئیدی تولید می شوند، ترکیبات فنلی می توانند آنتی اکسیدان های قوی باشند که رادیکال های آزاد را جمع آوری می کنند لذا این ترکیبات به عنوان سیستم های دفاعی گیاه شناخته شده اند. تحت تنش آفت کش دیازینون، گیاه برنج فعالیت بیشتری در مسیر فنیل پروپانوئیدی انجام داده لذا سنتز فنولیک اسیدها افزایش یافته تا فعالیت سمیت زدایی ترکیبات فنلی را تشدید نماید (Baxter et al., 2007).

تجمع پرولین تحت استرس به تثبیت آنزیم ها و ساختارهای غشایی کمک می کند. همچنین تصور می شود که پرولین



شکل ۷- مسیر بیوسنتز پرولین در گیاهان.

Fig. 7. Biosynthetic pathway of proline in plants.

مهم ترین متغیرها را شناسایی و سایر موارد را حذف می نماید. نتایج CART نسبت به تبدیل متغیرهای مستقل آن تغییر نمی کند. به عنوان مثال تبدیل یک یا چند متغیر به مقادیر لگاریتم آن ها یا ریشه دوم آن ها تأثیری در نتایج CART ندارد فقط مقادیر هر شکاف را تغییر می دهد. به آسانی می تواند داده های پرت را دسته بندی نماید و بدون هیچ فرضی محاسبات را سریع انجام می دهد.

بنابراین، تجمع پرولین در گیاهان تحت استرس غیرزنده به خوبی ثبت شده است.

روش CART دارای مزایای برجسته ای است که استفاده از آن برای تشخیص سریع مهم ترین تفاوت ها بین دو گروه مورد مقایسه بخصوص تشخیص سریع بیومارکرها، توصیه می شود. روش CART نیاز به پارامتر و فرایندهای بهینه سازی ندارد. الگوریتم CART به خودی خود

References:

- Arshad, U., Aliakbar, A., Sadeghi, M., Jamalzad, F. and Chubian, F. 2006. Pesticide (diazinon and butachlor) monitoring in waters of the Shahid Beheshti Sturgeon Hatchery, Rasht, Iran. *J. Appl. Ichthyol.* (22): 231–233.
- Aydin, R. and Köprücü, K. 2005. Acute toxicity of diazinon on the common carp (*Cyprinus carpio* L.) embryos and larvae. *Pestic. Biochem. Physiol.* (82): 220–225.
- Bandurska, H. 2001. Does proline accumulated in leaves of water deficit stressed barley plants confine cell membrane injuries? II. Proline accumulation during hardening and its involvement in reducing membrane injuries in leaves subjected to severe osmotic stress. *Acta physiologiae plantarum.* 23(4): 483-490.
- Baxter, C. J., Redestig, H., Schauer, N., Repsilber, D., Patil, K.R., Nielsen, J., Selbig, J., Liu, J., Fernie, A. R., Sweetlove, L. J. 2007. The metabolic response of heterotrophic Arabidopsis cells to oxidative stress. *Plant Physiol.* (143): 312–325.
- Buyukokuroglu, M. E., Cemek, M., Yurumez, Y., Yavuz, Y. and Aslan, A. 2008. Antioxidative role of melatonin in organophosphate toxicity in rats. *Cell Biol. Toxicol.* (24): 151–158.
- Breiman, L. 1986. Classification and regression trees: CRC press.
- Cemek, M., Büyükben, A., Büyükokuroğlu, M. E., Aymelek, F. and Tür, L. 2010a. Protective roles of vitamin E (α -tocopherol), selenium and vitamin E plus selenium in organophosphate toxicity in vivo: A comparative study. *Pestic. Biochem. Physiol.* (96): 113–118.
- Cemek, M., Emin Büyükokuroğlu, M., Yürümez, Y., Yavuz, Y., Aslan, A., Büyükben, A., Aymelek, F. 2010b. Tissue trace and major element levels in organophosphate insecticide fenthion toxicity in rats: Prophylactic and therapeutic effect of exogenous melatonin. *Ecotoxicol. Environ. Saf.* (73): 206–212.
- Fiehn, O. 2002. Metabolomics—the link between genotypes and phenotypes. *Plant molecular biology.* 48(1-2): 155-171.
- Frank, T., Rohlig, R., Davies, H. V, Barros, E. and Engel, K. H. 2012. Metabolite Profiling of Maize Kernel - Genetic Modification versus Environmental Influence. *J. Agric. Food Chem.* (60): 3005–3012.
- Isik, I. and Celik, I. 2008. Acute effects of methyl parathion and diazinon as inducers for oxidative stress on certain biomarkers in various tissues of rainbowtrout (*Oncorhynchus mykiss*). *Pestic. Biochem. Physiol.* (92): 38–42.
- Joshi, S.C., Mathur, R., Gajraj, A. and Sharma, T. 2003. Influence of methyl parathion on reproductive parameters in male rats. *Environ. Toxicol. Pharmacol.* (14): 91–98.
- Kim, J. K., Bamba, T., Harada, K., Fukusaki, E., & Kobayashi, A. 2007. Time-course metabolic profiling in Arabidopsis thaliana cell cultures after salt stress treatment. *Journal of Experimental Botany.* 58(3): 415-424.
- Kim, J. K., Park, S. Y., Lim, S. H., Yeo, Y., Cho, H. S. and Ha, S. H. 2012. Comparative metabolic profiling of pigmented rice (*Oryza sativa* L.) cultivars reveals primary metabolites are correlated with secondary metabolites. *Journal of Cereal Science.* 57:14–20.
- Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S. and Fiehn, O. 2009. FiehnLib: Mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal. Chem.* (81): 10038–10048.
- Köprücü, S. Ş., Köprücü, K., Ural, M. Ş., Ispir, Ü. and Pala, M. 2006. Acute toxicity of organophosphorous pesticide diazinon and its effects on behavior and some hematological parameters of fingerling European catfish (*Silurus glanis* L.). *Pestic. Biochem. Physiol.* (86): 99–105.
- Mallard, W. G. and Reed, J. 1997. Automated Mass Spectral Deconvolution & Identification System AMDIS - USER GUIDE. Program.
- Neishabouri, E. Z., Hassan, Z. M., Azizi, E. and Ostad, S. N. 2004. Evaluation of immunotoxicity induced by diazinon in C57bl/6 mice. *Toxicology.* (196): 173–179.
- Ralston-Hooper, K. J., Adamec, J., Jannash, A., Mollenhauer, R., Ochoa-Acuña, H. and Sepúlveda, M. S. 2011. Use of GC×GC/TOF-MS and LC/TOF-MS for metabolomic analysis of *Hyalella azteca* chronically exposed to atrazine and its primary metabolite, desethylatrazine. *Journal of Applied Toxicology.* 31(5): 399-410.
- Regoli, F. and Principato, G. 1995. Glutathione, glutathione-dependent and antioxidant enzymes in mussel, *Mytilus galloprovincialis* exposed to metals under field and laboratory conditions: Implications for the use of biochemical biomarkers. *Aquat. Toxicol.* (31): 143–164.

- Roessner, U., Wagner, C., Kopka, J., Trethewey, R. N. and Willmitzer, L. 2000.** Simultaneous analysis of metabolites in potato tuber by gas chromatography–mass spectrometry. *The Plant Journal*. 23(1): 131-142.
- Snee, R. 1977.** Validation of regression models: methods and examples. *Technometrics*. (19): 415–428.
- Tsugawa, H., Tsujimoto, Y., Sugitate, K., Sakui, N., Nishiumi, S., Bamba, T. and Fukusaki, E. 2014.** Highly sensitive and selective analysis of widely targeted metabolomics using gas chromatography/triple-quadrupole mass spectrometry. *J. Biosci. Bioeng.* (117): 122–128.
- Wang, L., Ye, W., Zhou, S., Lin, K., Zhao, M., Liu, W. 2009.** Acute and chronic toxicity of organophosphate monocrotophos to *Daphnia magna*. *J. Environ. Sci. Health. B.* (44): 38–43.
- Zhang, Y., Hou, Y., Chen, F., Xiao, Z., Zhang, J. and Hu, X. 2011.** The degradation of chlorpyrifos and diazinon in aqueous solution by ultrasonic irradiation: Effect of parameters and degradation pathway. *Chemosphere*. 82(8): 1109-1115.

Metabolomics Analysis Using Decision Tree to Differentiate between Rice Plants without Contamination and the Ones under Diazinon Stress

Mahdavi, V.*¹ and Nori, M. Z.²

1. Department of Pesticide Research, Iranian Research Institute of Plant Protection, Agricultural Research, Education and Extension Organization (AREEO), Tehran, Iran. 2. Rice Research Institute of Iran-Mazandaran Branch, Agricultural Research, Education and Extension Organization (AREEO), Amol, Iran.

Received: Jun, 21, 2020

Accepted: Jan, 4, 2021

Abstract:

Presenting an effective classification procedure that discriminates between unexposed and exposed plant samples with pesticide, can be very important and efficient to increase quality control in agricultural products and food safety. The proposed method shows a pattern for discriminating between rice plant without contamination and those treated with diazinon. The aim of this study was to investigate the metabolic changes of shiroodi cultivar rice plant (the most common cultivar in Iran) in response to diazinon and to evaluate the dynamic changes in the concentration levels of metabolites over time. For this purpose, after treating the rice plant with diazinon, un-volatile metabolites were derivatized to make volatile and analyzable compounds with gas chromatography, and the metabolite profile was investigated using GC-MS. Using Classification and Regression Tree (CART), the metabolites that showed the most changes under diazinon stress were identified. For the first time, CART was used to model and predict the metabolic changes of control and under diazinon stress rice plant. The results of the analysis showed that, the most intensive changes were related to amino acids particularly phenylalanine and proline. The results obtained, also, show that CART can be a very powerful tool for discriminating between exposed and un-exposed plants.

Keywords: Metabolomics, Diazinon, Rice, Classification and regression tree.

* Corresponding author: Vahideh Mahdavi, Email: yahideh.mahdavi@yahoo.com