

آنالیز غنی‌سازی مجموعه ژنی در غده‌های سیب‌زمینی (*Solanum tuberosum* L.) در جریان مراحل نمو

مریم شیرانی بیدآبادی¹، فرهاد نظریان فیروزآبادی^{2*}، کریم سرخه³، احمد اسماعیلی²

1. دانشجوی دکتری مهندسی ژنتیک و ژنتیک مولکولی، گروه زراعت و اصلاح نباتات، دانشکده کشاورزی، دانشگاه لرستان، خرم‌آباد، ایران.

2. استاد بیوتکنولوژی، گروه زراعت و اصلاح نباتات، دانشکده کشاورزی، دانشگاه لرستان، خرم‌آباد، ایران

3. استادیار، گروه مهندسی تولید و ژنتیک گیاهی، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز، ایران

(تاریخ دریافت: 1399/1/22 - تاریخ پذیرش: 1399/3/23)

Gene set enrichment analysis in potato tubers (*Solanum tuberosum* L.) during developmental stages

Maryam Shirani-Bidabadi¹, Farhad Nazarian-Firouzabadi^{2*}, Karim Sorkheh³, Ahmad Ismaili²

1. Ph.D.Candidate, Agronomy and Plant Breeding Department, Faculty of Agriculture, Lorestan University, Khorramabad, Iran.

2. Professor, Agronomy and Plant Breeding Department, Faculty of Agriculture, Lorestan University, Khorramabad, Iran.

3. Assistant Professor, Department of Production Engineering and Plant Genetics, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

(Received: Apr. 10, 2020 - Accepted: Jun, 11, 2020)

Abstract

Potato is the fourth most valuable plant for human nutrition. In addition to carbohydrates, potato tuber contains important vitamins and micronutrients for human health. During developmental stages of Potato; morphological, physiological and molecular changes occur. From the genetic point of view, studying these events is crucial for breeding potatoes with nutritional values and a higher yield. The next-generation sequencing methods generate abundant and useful genetic data for molecular breeding of crop plants. In this study, two different potato tuber developmental stages (S1 and S2) were chosen and sampled. Following RNA extraction and cDNA synthesis, Illumina RNA sequencing was performed in two replications for each stage. Bioinformatics analysis, gene ontology, and gene set enrichment analysis were performed. Finally, out of 1829 differential expression genes, 1186 genes were identified and validated by bioinformatics resources. The result of gene ontology comparison showed that 393, 483 and 669 genes were involved in biological processes, cellular components, and molecular functions, respectively. Most genes were present in the phenylpropanoid biosynthesis, biosynthesis of secondary metabolites and metabolic pathways. Peroxidase and membrane transporters coding genes were the most important genes at the beginning of tuber onset to tuberization. These results suggested that development in potato tuber activates metabolic pathways, which not only promotes growth and development but also activates the pathways involved in stress responses and synthesis of different compounds. Ten key genes were identified involved in starch biosynthesis, most of which showed a significant up-regulation, suggesting that the starch biosynthesis pathway starts from the onset of stolon development to tuber maturation.

Keywords: Development, Gene ontology analysis, Gene set enrichment analysis, Potato, RNA-seq

چکیده

سیب‌زمینی، چهارمین گیاه ارزشمند در تغذیه بشر است. غده سیب‌زمینی علاوه بر کربوهیدرات‌ها، دارای ویتامین‌ها و ریزمغذی‌های مهمی برای تأمین سلامت بشر است. طی مراحل نمو غده سیب‌زمینی، تغییرات متعدد مورفولوژیکی، فیزیولوژیکی و مولکولی رخ می‌دهند. مطالعه این اتفاق‌ها از منظر مولکولی، برای به‌نژادی ارزش غذایی و بهبود عملکرد سیب‌زمینی بسیار مهم است. روش‌های توالی‌یابی نوین، داده‌های ژنتیکی فراوان و سودمندی برای به‌نژادی مولکولی محصولات زراعی تولید می‌کنند. در این مطالعه، از غده در حال نمو سیب‌زمینی در دو دوره نمو نمونه‌برداری صورت گرفت. توالی‌یابی cDNA و ساخت RNA پس از استخراج با دو تکرار برای هر Illumina انجام شد. آنالیزهای بیوانفورماتیکی، هستی‌شناسی ژن و غنی‌سازی گروه‌های ژنی صورت گرفت. در نهایت، از 1829 ژن افتراقی، 1186 ژن توسط پایگاه‌های معتبر شناسایی شدند. نتایج مقایسه‌های هستی‌شناسی ژن نشان داد که به ترتیب 393، 483 و 669 ژن در فرایندهای بیولوژیکی، اجزای سلولی و کارکردهای مولکولی نقش داشتند. بیشترین تعداد ژن در مسیرهای بیوسنتز فنیل پروپانوئیدها، بیوسنتز متابولیت‌های ثانویه و مسیرهای متابولیکی حضور داشتند. ژن‌های کدکننده پراکسیدازها و ناقلین غشایی مهم‌ترین ژن‌ها در ابتدای دوره نمو تا شروع غده‌دهی بودند. نمو غده در سیب‌زمینی مسیرهای متابولیکی بسیاری را فعال می‌سازد، که نه تنها موجب رشد و نمو می‌گردند، بلکه باعث فعال شدن مسیرهای مقابله با تنش‌ها و همچنین مسیرهای ساخت مواد مورد نیاز می‌گردند. در مجموع تعداد ده ژن کلیدی برای بیوسنتز نشاسته در جریان نمو، شناسایی شدند که بیشتر آن‌ها افزایش بیان معنی‌داری نشان دادند. این موضوع نشان می‌دهد که فرایند بیوسنتز نشاسته از شروع نمو استولون تا بلوغ غده پیوسته ادامه دارد.

واژه‌های کلیدی: آنالیز غنی‌سازی گروه‌های ژنی، آنالیز هستی‌شناسی

ژن، سیب‌زمینی، نمو، RNA-seq.

مقدمه

سیب‌زمینی (*Solanum tuberosum* L.)، یکی از گیاهان مهم زراعی برای تأمین غذای بشر است. این گیاه نه تنها یکی از گیاهانی است که نقش مهمی در رفع گرسنگی در جهان دارد؛ بلکه یک گیاه مهم اقتصادی است که باعث ایجاد فرصت‌های شغلی بسیاری می‌شود (Bradshaw, 2010). بر اساس آمار سازمان خواربار جهانی، سیب‌زمینی با میانگین عملکرد 20946 Kg/ha چهارمین محصول مهم پس از برنج، ذرت و گندم می‌باشد. در ضمن این گیاه طی ده سال اخیر با افزایش عملکردی معادل kg/ha 2932 مواجه بوده است (Faostat, 2018) که نشان می‌دهد تا چه اندازه بشر به این محصول وابسته است. با توجه به اهمیت این محصول، شناسایی هرچه بیشتر مکانیسم‌های ژنتیکی آن حائز اهمیت است تا بتوان میزان عملکرد این گیاه را در واحد سطح به کمک آگاهی از پتانسیل‌های ژنتیکی آن و استفاده از ارقام مقاوم به تنش‌های زیستی و غیرزیستی افزایش داد (Dahal et al., 2019).

در جریان رشد و نمو غده سیب‌زمینی، تغییرات فیزیولوژیکی و مورفولوژیکی مهمی رخ می‌دهند که در جریان آن‌ها، مسیرهای بیوسنتزی فراوانی فعال می‌شوند. چگونگی نمو غده سیب‌زمینی یکی از مباحث جالب در رشد و نمو گیاهان، به ویژه اندام‌های زیرزمینی است. جوانه‌های جانبی سیب‌زمینی می‌توانند موجب تشکیل استولون شوند. تحت شرایط مطلوب رشد، غده‌ها روی این استولون‌ها شکل می‌گیرند. به طور کلی، مراحل نموی مهمی مانند: تحریک و تشکیل استولون، توقف رشد طولی استولون، تحریک و شروع غده‌دهی در تشکیل غده‌های گیاه سیب‌زمینی دیده می‌شوند. مراحل فوق‌الزاماً به ترتیب زمانی رخ نمی‌دهند. تحریک استولون‌دهی تا شروع غده‌دهی به یکدیگر بسیار نزدیک هستند و مجزا ساختن این دو بسیار مشکل است، ولی از لحاظ فیزیولوژی متفاوت می‌باشند

(Vreugdenhil and Struik, 1989). زمان و محل تقسیم سلولی و توسعه سلولی در نواحی مختلف استولون و غده در حال نمو متفاوت است. تقسیم سلولی ابتدا در نواحی رأسی استولون رخ می‌دهد. سلول‌ها به‌طور عرضی تقسیم شده و سپس طویل می‌شوند (Xu et al., 1998). در جریان رشد غده، نشاسته، پروتئین‌های ذخیره‌ای و طیف گسترده‌ای از متابولیت‌ها در غده ذخیره می‌شوند. اندازه نهایی غده، مورفولوژی و محتوای غده، جنبه‌های متفاوتی از کیفیت غده را رقم می‌زند که به‌طور عمده به ژنوتیپ و شرایط محیطی وابسته است (Kloosterman, 2006).

در دسترس بودن ژنوم سیب‌زمینی و همین‌طور ترانسکریپتوم بسیاری از اندام‌ها و مراحل رشد آن، باعث تسهیل روند مطالعات روی این گیاه شده است و محققان را قادر به توسعه ارقام ارزشمندی از سیب‌زمینی کرده است (Massa et al., 2011). آنالیزهای ترانسکریپتوم، داده‌های بسیار ارزشمندی از چگونگی بیان ژن‌ها در شرایط زمانی، بافت و اندام خاص و تحت شرایط آزمایشی مختلف فراهم می‌کنند. توالی‌یابی RNA (RNA-seq)¹ از جمله روش‌های با توان عملکردی بالاست که حجم وسیعی از داده‌ها را ایجاد می‌کند و محدودیت‌های روش‌های مولکولی گذشته را ندارد (Massa et al., 2011). تکنولوژی توالی‌یابی cDNA با توان بالا² آنالیزهای کامل و کارآمدی از ترانسکریپتوم در سلول را ممکن ساخته است. به علت تفاوت در عمق توالی و ترکیب RNAها تجزیه و تحلیل در روش RNA-seq یک مشکل می‌باشد. برای رفع این مسئله روش‌های مختلف برای مقایسه بیان ژن بین نمونه‌ها، روش‌های نرمال‌سازی متنوعی ایجاد شده است و تنوعی از روش‌ها برای آزمون بیان افتراقی هر ژن بر

1. RNA-Sequencing

2. High-throughput sequencing

فوتویی تمایز پیچیده، حالت‌های بیماری و پاسخ به تنش‌های محیطی نیستند (Hung *et al.*, 2012; Subramanian *et al.*, 2005). GSEA باعث تسهیل در تفسیر آزمایش در مقیاس وسیع می‌گردد. این امر توسط شناسایی مسیرها و فرایندها انجام می‌شود (Subramanian *et al.*, 2005). در آنالیز روش‌های با عملکرد بالا لیست بزرگی از ژن‌ها (شامل صدها الی هزاران ژن) به دست می‌آید که تفسیر آن بسیار مهم است و نیاز به کمک بسته‌های نرم‌افزاری بیوانفورماتیک دارد. در مطالعه حاضر از منبع بیوانفورماتیک DAVID (<https://david.ncifcrf.gov/>) استفاده شد. این پایگاه داده قادر به رسیدگی به تمام لیست‌های ژنی فارغ از این‌که از چه بسته نرم‌افزاری به دست آمده‌اند، می‌باشد (Sherman and Lempicki, 2009).

به طور کلی، هدف از انجام این پژوهش بررسی ژن‌های بیان شده افتراقی و شناسایی مسیرهای فعال در دوره نموی در غده سیب‌زمینی و همچنین تفسیر و تحلیل حجم وسیعی از داده‌های RNA-Seq به کمک روش GSEA می‌باشد تا با درک بهتر گروه‌های ژنی بتوان به شناخت بهتر مسیرها و ژن‌های دخیل در دوره‌های نموی غده سیب‌زمینی پی برد که خود در ایجاد ارقام با پتانسیل بالاتر کمک شایانی خواهد کرد. در مطالعه حاضر برای اولین بار حجم وسیعی از اطلاعات حاصل از آنالیز ترانسکریپتوم در دوره نموی غده سیب‌زمینی مورد بررسی قرار گرفت. به نظر می‌رسد که آنالیز ترانسکریپتوم در مراحل نموی غده در سیب زمینی به عنوان یک گیاه مدل، بتواند به تفهیم بهتر بیان ژن‌ها و کارکردهای مربوط به آن‌ها کمک کند.

مواد و روش‌ها

مواد گیاهی

غده‌های گیاه سیب‌زمینی رقم مارفونا از جهاد کشاورزی و منابع طبیعی اصفهان تهیه شد. پس از شستشو و

پایه مدل‌های احتمالی گسسته، توسعه یافته است. برای تفسیر نتایج آنالیز بیان افتراقی، هستی‌شناسی ژن (GO) و یا سایر گروه‌بندی ژنی که کارکردهای مشترک را به اشتراک می‌گذارند، با نام آنالیز GO استفاده می‌شود. روش مفید دیگر آنالیز غنی‌سازی گروه ژنی^۲ (GSEA) است. برخلاف آنالیز هستی‌شناسی، GSEA از آستانه‌ای برای شناسایی ژن‌هایی با بیان افتراقی استفاده نمی‌کند، بلکه آماره کلموگرو-اسمیرنوف را به کار می‌گیرد که آیا سهم هر ژن در فوتیپ، غنی شدن هر گروه ژنی آزمون شده است. بدین وسیله GSEA قادر است تغییرات ظریف اما هماهنگ در یک گروه ژنی را ثبت کند. روش GSEA به طور معمول برای یافتن مسیرها و کارکردهای مهم در بیماری‌های گوناگون و شرایط مختلف سلولی استفاده می‌شود (Yoon *et al.*, 2016). از طرفی بیان‌های افتراقی نسبی بوده و فهرست بزرگی از ژن‌هایی را ارائه می‌دهد که از لحاظ آماری معنی‌دار هستند و موضوع بیولوژیکی واحدی نداشته و باعث سردرگمی و اتلاف وقت زیست‌شناسان می‌شوند. برای غلبه بر چالش‌های ذکر شده، روش آنالیز غنی‌سازی گروه ژنی توسعه یافت است که داده‌ها را در گروه‌های ژنی ارزیابی می‌کند. هدف از GSEA تعیین تمایل اعضای یک گروه ژنی در بالا یا پایین یک مجموعه ژنی است؛ که این گروه ژنی با یک طبقه فوتویی ارتباط دارد (Subramanian *et al.*, 2005).

روش‌های سنتی برای آنالیز بیان ژنی روی شناسایی تک ژن‌ها بین دو مرحله یا شرایط مختلف، تمرکز دارد. با وجود اطلاعات ارزشمند این گونه مطالعات، این روش‌ها قادر به آشکارسازی فرایندهای بیولوژیکی مثل مسیرهای متابولیکی، درک مراحل

1. Gene Ontology
2. Gene Set Enrichment Analysis (GSEA)

(RIN) بالا، ساخت cDNA توسط کیت (Illumina TruSeqTM RNA) انجام شد. سپس نمونه‌ها با پلتفرم HiSeq2500 توالی‌یابی شدند (شرکت نواژن، چین). توالی‌یابی با طول 150 جفت باز و از هر دو انتها انجام شد.

آنالیزهای بیوانفورماتیکی

کیفیت توالی‌های خام با نرم‌افزار FastQC ارزیابی شد و پیرایش بازهای ابتدائی و توالی‌های نامناسب به‌دست‌آمده از نرم‌افزار FastQC با استفاده از Trimmomatic 3-6 انجام شد (Andrews, 2010; Bolger *et al.*, 2014). توالی‌های پیرایش‌شده با ژنوم سیب‌زمینی

(ftp://ftp.ensemblgenomes.org/pub/release-

34/plants/fasta/solanum_tuberosum/dna/

توسط نرم‌افزار Bowtie2.v2.3.3.1 هم‌ردیف گردیدند (Langmead and Salzberg, 2012). در نهایت برای شمارش خوانش‌ها از RSEM 3.3.1 استفاده شد (Li and Dewey, 2011). خروجی این نرم‌افزار حجم وسیعی از ژن‌ها به‌همراه میزان بیان آن‌ها بود که به‌منظور نرمال‌سازی داده‌های حاصل و مقایسه دوره‌های نمو، نرم‌افزار Deseq2 به‌کار گرفته شد (Love *et al.*, 2014). برای جداسازی ژن‌هایی که در مقایسه دوره‌های نمو بیان افتراقی داشتند از مقدار قدر مطلق لگاریتم نسبت دو دوره (FC)² اعداد استفاده شد و ژن‌هایی با $1 \leq \log FC \leq -1$ و نرخ خطای $FDR \leq 0.05$ برای مراحل بعدی آنالیز انتخاب شدند. در آنالیزهای بیوانفورماتیک با در نظر گرفتن و محاسبه میزان اشتباه نوع اول (FDR)³ ژن‌هایی با درصد اطمینان بالاتر برای مراحل بعدی تجزیه و تحلیل، انتخاب می‌شوند.

استریل کردن با هیپوکلریت سدیم 4٪، در گلدان حاوی خاک شنی استریل شده با فرمالین با عمق 30 سانتی‌متر کشت شد و در دمای 24 تا 26 درجه سانتی‌گراد گلخانه دانشکده کشاورزی دانشگاه لرستان قرار گرفت. اولین نمونه‌برداری شش هفته پس از کاشت و از اندام زیرزمینی در مرحله رشدی نوک قلاب مانند استولون با دو تکرار بیولوژیک صورت گرفت (شکل 1). نمونه‌های دوره نمو دوم از مرحله شروع غده‌دهی با دو تکرار بیولوژیک و دو هفته پس از نمونه‌برداری دوره اول، به‌منظور استخراج RNA از گیاه جدا شدند (شکل 1-الف و ب). نمونه‌ها بلافاصله در ازت مایع منجمد گردیدند و تا زمان استخراج RNA در فریزر -80°C نگهداری شدند.



شکل 1. مراحل نمو مورد نمونه‌برداری. (الف) نوک قلاب مانند؛ (ب) شروع غده‌دهی

استخراج RNA، سنتز cDNA و توالی‌یابی

ابتدا نمونه‌ها در هاون چینی به کمک ازت مایع آسیاب شدند و استخراج RNA با استفاده از روش لیتیم کلراید در دو تکرار برای هر دوره نمو صورت گرفت (Chang *et al.*, 1993). کیفیت RNA استخراج شده با الکتروفورز ژل آگاروز 1٪، و کمیت آن با استفاده از روش اسپکتروفتومتری سنجیده شد. نمونه‌های با کیفیت مناسب که نسبت جذب 260/280 در آن‌ها تقریباً برابر عدد 2 بود؛ جهت توالی‌یابی (شرکت نواژن، چین) ارسال گردید. برای نمونه‌هایی با عدد سالم بودن¹

2. Log Fold Change
3. False Discovery Rate

1. RNA Integrity Number (RIN)

این سه گروه فرایندهای بیولوژیکی (BP)^۱، اجزای سلولی (CC)^۲ و کارکردهای مولکولی (MF)^۳ بودند که هر گروه دارای زیر گروه‌های مختلفی بود. به دلیل بزرگی ابعاد تصویر از رسم دسته‌های مربوط به هر زیرگروه صرف‌نظر شده است. با تبدیل ID ژن‌ها به ID قابل شناسایی برای پایگاه DAVID، داده‌های مقایسه‌ای بین دو دوره نمودی غده سیب زمینی، در این پایگاه قرار گرفت. در مورد داده‌های افتراقی بین این دو دوره، تعداد 1186 ژن شناسایی شدند. در قسمت مستندسازی کاربردی (FA)^۴، بخش هستی‌شناسی ژن؛ 393 مورد در فرایندهای بیولوژیکی، 483 مورد در اجزای سلولی و 669 مورد در کارکردهای مولکولی نقش داشتند.

فرایندهای بیولوژیکی

همان‌طور که در جدول 1 مشخص است پایگاه DAVID زیر گروه‌های مختلفی برای قسمت فرایندهای بیولوژیکی ارائه می‌دهد که در این میان به‌ترتیب فرایندهای متابولیسم و پاسخ به تنش اکسیداتیو و سپس فرایند کاتابولیسمی هیدروژن پراکسید بیشترین تعداد ژن درگیر را داشتند. با تجزیه و تحلیل داده‌ها در پایگاه ebi.ac.uk، به فرایندهای متابولیسمی شماره هستی‌شناسی GO:0008152 تعلق گرفت. این شماره هستی‌شناسی به مسیرهای متابولیسمی و واکنش‌های شیمیایی شامل آنابولیسم و کاتابولیسم مرتبط است که به کمک آن‌ها موجودات زنده مواد شیمیایی را انتقال می‌دهند. این مورد شامل فرایندهای متابولیسمی مولکول‌های کوچک و همچنین فرایندهای ماکرومولکول‌ها مثل؛ همانندسازی و اصلاح DNA و سنتز و تجزیه پروتئین‌ها است.

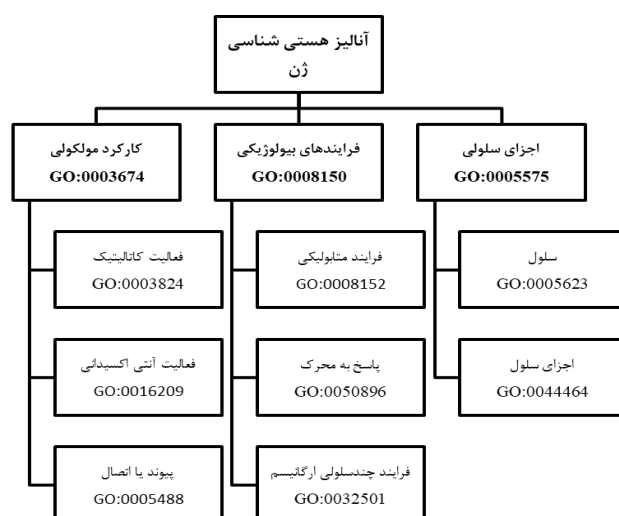
آنالیز هستی‌شناسی و غنی‌سازی گروه‌های ژنی در آنالیز اولیه، کدهای PGSC Protein ID حاصل از توالی ژنوم سیب‌زمینی، در سایت Agrigo.v2 قرار گرفت و از قسمت SEA و با آزمون فیشر تجزیه و تحلیل شد (Tian *et al.*, 2017). برای مستندسازی و تفسیر داده‌های با بیان افتراقی از پایگاه داده‌ای DAVID استفاده شد (Dennis *et al.*, 2003). برای استفاده از پایگاه، ابتدا کد ژن‌های حاصل توسط پایگاه Plant Ensemble به شماره رهگیری NCBI تبدیل شدند (Kinsella *et al.*, 2011). شماره رهگیری‌ها در پایگاه DAVID قرار گرفتند و در مرحله بعدی لیست ژنی موجود در این پایگاه در ارتباط با گیاه سیب‌زمینی به‌دست آمد. گروه‌های مختلف ژنی، مسیرهای مختلف و خوشه‌بندی براساس مستندسازی عملکردی حاصل شد.

نتایج و بحث

چگونگی نمو غده سیب‌زمینی به عنوان یک گیاه مدل، با توجه به اهمیت غذایی آن می‌تواند به محققان در پیشبرد اهداف به‌نژادی کمک شایانی نماید. اگرچه تحقیقات فراوانی در مورد مراحل نمو غده سیب‌زمینی از منظر مورفولوژیکی (Cenzano *et al.*, 2003; Schwall *et al.*, 2000)؛ فیزیولوژیکی (Ewing and Struik, 1992; Struik and Ewing, 1995) و تأثیر عوامل محیطی (Jackson *et al.*, 1998; Papathanasiou *et al.*, 1999) صورت گرفته است، هنوز مطالعه جامعی در خصوص عوامل ژنتیکی کنترل‌کننده این مراحل با توجه به ترکیبات موجود در غده صورت نگرفته است. هدف از انجام آنالیز هستی‌شناسی ژن اتحاد ژن و فراورده ژن در بین همه گونه‌ها است (Chen *et al.*, 2020a).

در نتایج حاصل ژن‌ها در سه گروه اصلی طبقه‌بندی شدند. نتایج حاصل از هستی‌شناسی ژن‌های حاصل در Agrigo.v2 در شکل 2 نشان داده شدند.

1. Biological Process
2. Cellular Component
3. Molecular Function
4. Functional Annotation



شکل 2. آنالیز هستی‌شناسی ژن‌های افتراقی در سه گروه کارکرد مولکولی، فرایندهای بیولوژیکی و اجزای سلولی

جدول 1. هستی‌شناسی ژن‌های افتراقی درگیر در مراحل نمو در سه گروه فرایند بیولوژیکی، اجزای سلولی و کارکرد مولکولی

	شماره هستی‌شناسی ژن	تفسیر	تعداد ژن درگیر	P-value	Benjamini
فرایند بیولوژیکی	GO:0042744	فرایند کاتابولیک پراکسید هیدروژن	20	1.4E-8	3.2E-6
	GO:0006979	پاسخ به تنش اکسیداتیو	23	5.5E-7	6.3E-5
	GO:0009691	فرایند بیوسنتز سیتوکینین	6	3E-5	0.002
	GO:0006869	انتقال لیپید	9	0.002	0.12
	GO:0046274	فرایند کاتابولیک لیگنین	6	0.003	0.15
	GO:0006817	انتقال یون فسفات	4	0.005	0.19
	GO:0008152	فرایند متابولیکی	34	0.005	0.17
	GO:0006857	انتقال الیگوپپتید	5	0.01	0.39
	GO:0030001	انتقال یون فلز	9	0.02	0.42
	GO:0009734	مسیر سیگنال دهی فعال شده اکسین	8	0.02	0.39
	GO:0009958	رشد افقی مثبت	3	0.02	0.42
	GO:0008643	انتقال کربوهیدرات	6	0.03	0.5
	GO:0045893	تنظیم مثبت نسخه برداری	5	0.05	0.62
	GO:0005985	فرایند متابولیکی ساکاروز	3	0.08	0.74
	GO:0010311	تشکیل ریشه جانبی	3	0.08	0.74
	GO:0016049	رشد سلولی	3	0.08	0.74
	اجزای سلولی	GO:0016021	اجزای کامل غشا	295	6E-9
GO:0005576		نواحی خارج سلولی	30	1E-5	2.6E-4
GO:0005886		غشای پلاسما	29	0.01	0.16
GO:0048046		آپوپلاست	11	0.02	0.28
GO:0031225		اجزای اتصال یافته ی غشا	3	0.06	0.5
GO:0009505		دیواره سلول گیاهی	5	0.09	0.56
کارکرد مولکولی	GO:0020037	اتصال هم	56	3.2E-10	7E-8
	GO:0004601	فعالیت پراکسیداز	23	9.9E-8	1E-5
	GO:0005215	فعالیت ناقل	25	3E-5	0.002
	GO:0016799	فعالیت هیدرولازی	6	5.1E-5	0.002
	GO:0016705	فعالیت اکسیدوردوکتاز	31	3.1E-4	0.01
	GO:0005506	اتصال یون آهن	35	4.2E-4	0.01
	GO:0022857	فعالیت ناقل از طریق غشا	9	9.7E-4	0.02
	GO:0004497	فعالیت مونو اکسیژناز	26	0.001	0.03
	GO:0043565	اتصال اختصاصی به توالی DNA	30	0.001	0.03
	GO:0046872	اتصال به یون فلز	68	0.002	0.04
	GO:0008289	اتصال لیپید	12	0.003	0.06
	GO:0005315	ناقل فسفات غیرآرگانیک از طریق غشا	5	0.005	0.08
	GO:0052716	فعالیت اکسیدوردوکتاز اکسیژن	6	0.005	0.09
	GO:0003700	فاکتور نسخه برداری	42	0.01	0.24
	GO:0015238	فعالیت ناقل دارویی از طریق غشا	7	0.05	0.52
	GO:0015297	فعالیت آنتی پورتر	7	0.05	0.52
	GO:0008271	ناقل سولفات فعال ثانویه از طریق غشا	4	0.06	0.56
	GO:0016157	فعالیت سنتز ساکاروز	3	0.07	0.61
	GO:0016491	فعالیت اکسیدوردوکتاز	30	0.07	0.61
	GO:0008171	فعالیت اکسیژن-متیل ترانسفراز	5	0.09	0.68

فرایندهای متابولیکی از جمله موارد مهم است که حتی در تنش‌های غیرزیستی گیاه سبب‌زمینی نیز بروز می‌کند (Chen *et al.*, 2020b). هستی‌شناسی GO:0006979 مربوط به پاسخ به تنش‌های اکسیداتیو است که نشان‌دهنده هر فرایندی است که منجر به تغییر حالت یا فعالیت یک سلول یا ارگانسیم (از جمله حرکت، ترشح، تولید آنزیم، بیان ژن و غیره) در نتیجه استرس اکسیداتیو می‌گردد. تنش‌ها و یا بیمارگرها از جمله مواردی هستند که باعث تغییر شده و ژن‌هایی که در گروه تنش‌های اکسیداتیو هستند را فعال می‌کنند (Sanchez *et al.*, 2020). این مطلب نشان‌دهنده آمادگی کامل گیاه در برابر هر تغییری است و مهم‌ترین تغییر یعنی نمو را نیز دربر می‌گیرد. در نهایت فرایند کاتابولیکی هیدروژن پراکسید با شماره هستی‌شناسی GO:0042744 در فرایند شکستن H₂O₂ دخیل است. یکی از موارد مهم دیگر طی نمو سبب‌زمینی، متابولیسم ساکاروز و رشد سلولی است. متابولیسم ساکاروز (GO:0005985) مسیری شامل ساکاروز می‌باشد. وجود این شماره هستی‌شناسی بر پویایی سنتز ساکاروز و پیرو آن ساخت نشاسته طی دوره نمو غده دلالت دارد. همان‌طور که در جدول یک مشخص است، در این مسیر سه ژن نقش داشتند. GO:0016049 در رشد سلولی دخالت دارد که دارای فرایندی برای افزایش اندازه سلول طی زمان است.

اجزای سلولی

در فرایند مهم اجزای سلولی همان‌طور که در جدول 1 مشخص است، گروه‌های مختلف اجزای غشا، نواحی خارج سلولی، غشای پلاسمایی، آپوپلاست و غیره درگیر هستند که بیشترین تعداد ژن به ترتیب در سه مورد اول وجود داشت. در قسمت اول اجزای غشا شامل محصول ژن و کمپلکس پروتئینی هستند که حداقل برخی قسمت‌های پپتیدشان را در قسمت آب‌گریز غشا تعبیه می‌کنند. شماره هستی‌شناسی GO:0016021 در این رابطه تعیین شد. نواحی خارج سلولی با شماره

کارکردهای مولکولی

کارکردهای مولکولی در مرتبه بعدی قرار دارند (جدول 1). این بخش شامل زیرگروه‌های متنوعی از جمله فعالیت و انتقال پراکسیداز، اتصال یون آهن، اتصال لیپید و غیره است، اما بیشترین تعداد ژن به ترتیب مربوط به metal ion binding (GO:0046872) با 68 ژن، heme binding (GO:0020037) با 56 ژن و DNA-binding transcription factor activity (GO:0003700) با 42 ژن شناخته شده است. به‌منظور بررسی ژن‌های مربوط در داده‌های RNA-Seq، گروه‌های هستی‌شناسی که P-Value‌هایی با معنی‌داری بیشتر داشتند انتخاب شدند و از اطلاعات حاصل، جدول 2 ترسیم گردید (Subramanian *et al.*, 2005). پس از به‌دست‌آمدن جدول 2، با تبدیل شماره رهگیری-های NCBI به کد Ensemble ژن‌های مربوطه در داده‌ها جستجو شد. در این بین برخی ژن‌ها بین هر سه BP، CC و MF مشترک بودند. برخی ژن‌ها در CC و MF و در سایر موارد، ژن‌ها بین مراحل BP و MF مشترک بود. ژن مشترکی بین دو گروه CC و BP وجود نداشت.

جدول 2. مهم‌ترین زیرگروه‌های آنالیز هستی‌شناسی ژن از نظر معنی‌داری

Category	عنوان	تعداد ژن‌های درگیر در فرایند	P-value	تعداد کل ژن‌ها در هر گروه
BP ^a	GO:0042744~hydrogen peroxide catabolic process	20	1.41E-08	393
BP	GO:0006979~response to oxidative stress	23	5.56E-07	393
BP	GO:0009691~cytokinin biosynthetic process	6	3.05E-05	393
CC ^b	GO:0016021~integral component of membrane	295	6.27E-09	483
CC	GO:0005576~extracellular region	30	1.01E-05	483
MF ^c	GO:0020037~heme binding	56	3.29E-10	669
MF	GO:0004601~peroxidase activity	23	9.93E-08	669
MF	GO:0005215~transporter activity	25	3.09E-05	669
MF	GO:0016799~hydrolase activity	6	5.12E-05	669

a: BP فرایندهای بیولوژیکی؛ b: CC اجزای سلولی؛ c: MF کارکرد مولکولی.

شوری نقش دارند و به‌عنوان نشانگر بیوشیمیایی برای تعیین سن فیزیولوژیک غده به‌کار می‌رود (Rahnama and Ebrahimzadeh, 2005; Sandhu *et al.*, 2007). (Sandhu *et al.*, 2007) پراکسیداز را به‌عنوان یک نشانگر بیوشیمیایی در برگ معرفی کردند که فعالیت آن در اواسط رشد بیشترین بیان را دارد و پس از آن تا انتهای رشد مقدار پراکسیداز تا اواخر رشد روبه کاهش است. در زیر گروه فرایند بیوسنتز سیتوکینین از گروه BP و زیر گروه فعالیت هیدرولاز در گروه MF غالباً کربوکسی‌لیاز موجود بودند که در همه موارد کاهش بیان داشتند ولی مطالعه‌ای با ارتباط نزدیک برای این ژن‌ها یافت نشد.

ژن‌های موجود در گروه اجزای سلولی

تعداد 297 ژن در زیرگروه اجزای سلولی وجود داشت که همگی از اجزای غشا بودند؛ در این بین، تعداد 75 ژن افزایش بیان داشتند که بیشترین بیان مربوط به ناقلین نیترات و اجزای انتقال نیترات، ناقلین و مبدلین اکسالات، بی‌کربنات و سولفات، ناقل فسفات غیر ارگانیک و PT2 بود. تعداد 222 ژن طی دو دوره نمودی غده سیب‌زمینی کاهش بیان داشتند. ژن‌هایی که بیشترین کاهش بیان را داشتند، عبارت بودند از: ناقلین سولفات، پروتئین خانواده پاسخ به فسفات، CYP72A54، پروتئین انگشت روی، نودولین، پروتئین فیبر Fb34، کالمودولین، پروتئین خانواده سنتز بتا-کتواسیل-کوآ، پروتئین غشایی

آنالیز ژن‌های مربوط به هر GO در داده‌های حاصل از RNA-Seq

ژن‌های موجود در گروه فرایندهای بیولوژیکی

در بین ژن‌های حاصل، انواع مختلف پراکسیداز در فرایندهای بیولوژیکی؛ در دو زیر گروه پاسخ به تنش اکسیداتیو و فرایند کاتابولیک پراکسیداز هیدروژن پراکسید نقش داشتند. لازم به ذکر است برخی از این ژن‌های پراکسیداز از جمله پراکسیدازهای 12، 66، 17، 72 و آسکوربات پراکسیداز از مرحله اول به سوم کاهش بیان داشتند؛ در صورتی که پراکسیدازهای 55، 15، پراکسیداز کاتیونی، پراکسیداز آسکوربات و غیره افزایش بیان داشتند. در زیر گروه نواحی خارج سلولی از اجزای سلولی نیز 30 ژن به دست آمد؛ که غالب آن‌را پراکسیدازها تشکیل می‌دهند. آلفا دی-اکسیژناز با دو ژن و پریکارپ پراکسیداز 3 با یک ژن در زیر گروه پاسخ به تنش اکسیداتیو وجود داشتند. در حالیکه پریکارپ پراکسیداز 3 کاهش بیان داشت، آلفا دی‌اکسیژناز در یک ژن کاهش و در دیگری افزایش بیان داشت. پراکسیدازها در سرتاسر گیاه بیان می‌شوند و به دلیل ارتباط با کارکرد کاتالیتیکی فراوان بسیار مورد توجه هستند. پراکسیدازها در فعالیت‌های مختلف از جمله هیدروکسیل کردن پرولین، بهبود زخم، مقابله با تنش آبی، مقاومت به بیماری‌ها، تحمل به عنصر روی در لگوم‌ها و ریزوم و غیره نقش دارند. در گیاه سیب‌زمینی (*Solanum tuberosum L.*)، پراکسیدازها با نگهداری کیفیت و مقاومت به پوسیدگی طی انبارداری، در تحمل به

افزایش و کاهش هم‌زمان ژن‌های مشابه در یک زیرگروه در بافت‌های متفاوت و یا بافت‌های یکسان احتمالاً حاکی از مدیریت فرایندهای مختلف در یک زمان است که نیازمند بررسی‌های بیشتر و ردیابی هر ژن به صورت خاص می‌باشد.

آنالیز مسیرها

یکی دیگر از موارد مورد بحث در قسمت مستندسازی کاربردی، آنالیز مسیرهاست. به طور کلی 208 شماره رهگیری با مسیرهایی که در آن دخیل هستند شناسایی شدند. در این قسمت پایگاه DAVID با اتصال مستقیم به پایگاه KEGG (Kanehisa and Goto, 2000) ژن‌ها را در مسیرهای متابولیکی مربوط شناسایی می‌کند. هشت مسیر کلی در این مطالعه شناسایی شدند (جدول 3) که عبارتند از: بیوسنتز فنیل پروپانویید (32 ژن)، بیوسنتز متابولیت‌های ثانویه (81 ژن)، مسیرهای متابولیکی (109 ژن)، انتقال سیگنال هورمون گیاهی (26 ژن)، متابولیسم گلاکسیلات و دی‌کربوکسیلات (9 ژن)، متابولیسم گلوکاتایون (10 ژن)، تجزیه اسیدهای چرب (6 ژن) و متابولیسم ساکاروز - نشاسته (10 ژن). به دلیل اهمیت نشاسته و غنی بودن گیاه سبب‌زمینی از این نظر بررسی بیشتری روی این مسیر صورت گرفت. در این مسیر ده ژن شناسایی شدند که اطلاعات آن در جدول 4 ارائه شده است. لازم به ذکر است که الزاماً یک ژن تنها متعلق به یک مسیر خاص نمی‌باشد. در خوشه‌بندی مستندسازی کاربردی با اعمال سختگیرانه‌ترین پارامترها (طبق حالت پیش فرض پایگاه با اعمال آزمون بنجامینی) 41 خوشه وجود داشت. هر خوشه از گروهی از ژن‌های مشابه تشکیل شده است که دارای مفهوم بیولوژیکی یکسان می‌باشند. در این مرحله اولین خوشه فرایندهای کاتابولیکی پراکسید هیدروژن شامل 20 ژن می‌باشد که پراکسیدازهای گیاهی را شامل می‌شود.

UPF0497، پروتئین تنظیمی رسیدگی DDTFR18، پروتئین WAX2، ناقل پرولین، پروتئین COBRA، UPA16، پروتئین نودولین.

ژن‌های موجود در گروه کارکرد مولکولی

در گروه هم‌بایندینگ از گروه کارکرد مولکولی، سیتوکروم‌ها با ژن‌های مختلف وجود دارد که در برخی افزایش و در سایرین کاهش بیان مشاهده شد. گروه مهم دیگر در قسمت کارکرد مولکولی، فعالیت ناقلین است. در این میان ژن‌های ناقلین پورین با دو ژن (PGSC0003DMG400002133)، الیگوپپتید، (PGSC0003DMG400009706)، پپتید/نیترات و آکوئاپورین با سه ژن (PGSC0003DMG400004438)، PGSC0003DMG400026705 و (PGSC0003DMG400027819) کاهش بیان و ژن‌های ناقل پورین (PGSC0003DMG400013659) و آکوئاپورین با دو ژن دیگر (PGSC0003DMG400028182) و (PGSC0003DMG400020742) افزایش بیان نشان دادند.

ژن و آنزیم کلیدی در نقل و انتقالات و متابولیسم نیتروژن، ناقلین نیترات هستند. در مطالعه Tiwari و همکاران (2020) ناقلین نیترات، آکوئاپورین، سولفات/بی‌کربنات/اکسالات و ناقل sat-1، کمک کننده ناقل پرولین/سدیم بیان متفاوتی از کاهش و افزایش بین سه بافت ریشه، ساقه و استولون نشان دادند. در این مطالعه که در شرایط استرس کمبود نیتروژن در مقابل زیادبود نیتروژن اتفاق افتاد، ناقلین یون‌ها (ناقل اوره فعال و سدیم/پرولین) در استولون افزایش بیان نشان دادند، در صورتی که ناقلین نیترات کاهش بیان داشتند. به علاوه سیتوکروم P450 هیدورکسیلاز در بافت ریشه و استولون هم افزایش و هم کاهش بیان داشت (Tiwari et al., 2020).

جدول 3. نتایج آنالیز ژن‌های درگیر در برخی مسیرهای متابولیکی بین دو دوره نموی

شماره مسیر در KEGG	عنوان مسیر	تعداد ژن درگیر	P-value	Benjamini
sot00940	بیوسنتز فنیل پروپانویید	32	1.26E-10	1.1E-8
sot01110	بیوسنتز متابولیت‌های ثانویه	81	9.8E-8	4.6E-6
sot01100	مسیرهای متابولیک	109	1.1E-4	0.003
sot04075	انتقال سیگنال هورمون گیاهی	26	1.7E-4	0.004
sot00630	متابولیسم گلایکوسیلات و دی‌کربوکسیلات	9	0.01	0.18
sot00480	متابولیسم گلوکتانین	10	0.02	0.29
sot00071	تجزیه اسید چرب	6	0.06	0.58
sot00500	متابولیسم نشاسته و ساکاروز	10	0.06	0.5

جدول 4. ژن‌های شناسایی شده در مسیر متابولیسم ساکاروز - نشاسته

شماره رهگیری	نام ژن
102577594	Sucrose Synthase(SS16)
102577726	UDP-Glucose Pyrophosphorylase Precursor
102601570	Beta-Glucosidase 13-Like
102577816	Fructokinase-Like
102583115	Granule-Bound Starch Synthase 2
102586120	Alpha,Alpha-Trehalose-Phosphate Synthase
102581151	Sucrose-Phosphate Synthase 2
102591031	Trehalose-Phosphate Phosphatase
102586510	Sucrose Synthase-Like
102588295	Vicianin Hydrolase-Like

شدن، غده‌ها بزرگترین منبع گیاه سیب‌زمینی برای ذخیره کربوهیدرات (عمدتاً نشاسته) و پروتئین هستند. در حدود 2٪ از وزن غده تازه سیب‌زمینی را پروتئین و بین 15-25٪ آن را نشاسته تشکیل می‌دهد. تحقیقات نشان می‌دهند که نشاسته برای توسعه غده ضروری نیست. برای مثال، در گیاهان تراریخته که بیان ژن کدکننده ADP-پیروفسفوریلاز آن‌ها کاهش یافته بود، تنها تغییر، افزایش در تعداد غده و کاهش اندازه غده بود که احتمالاً نشان‌دهنده رقابت بین منابع مختلف است. Castro و همکاران (2020) نشان دادند که فنیل آلانین آمونیا-لیاز (PAL) به‌عنوان یک رابط بین متابولیسم اولیه و ثانویه سنتز عمل می‌کند. این آنزیم پیش‌ماده‌هایی برای انواع مختلف متابولیت‌های ثانویه را توسط کنترل جریان کربن در مسیر فنیل پروپانویید فراهم می‌کند. ترکیبات مشتقات فنیل پروپانویید نقش مهمی در رشد و نمو گیاهان بازی می‌کنند. این فرایندها شامل پشتیبانی مکانیکی، مولکول‌های سیگنال‌دهنده

خوشه دوم با 23 ژن در فرایندهای هم‌پراکسیداز و پاسخ به تنش نقش دارند و خوشه سوم که تعداد بسیار زیادی از ژن‌ها (295 ژن) در آن درگیر هستند، اجزای غشا می‌باشند. خوشه‌های بعدی تعداد ژن‌های کمتری دارند. این امر نشان‌دهنده اهمیت غشا و نقل و انتقالات آن طی مراحل رشد است. با وجود این‌که در این مطالعه تنشی اعمال نشده، حضور ژن‌های پاسخ‌دهنده به تنش‌ها پویایی این مکانیسم در برابر هر تغییر کوچک محیط گیاه را نشان می‌دهد. شروع و رشد از قسمت زیر زمینی به نام استولون آغاز و القا می‌شود و به دنبال آن غده‌ها که اندام‌های ذخیره‌سازی هستند تشکیل می‌شوند. هر دو فرایند مورفولوژیکی و بیوشیمیایی توسط الگوی بیان ژن‌های خاصی تنظیم می‌گردد (Sarkar, 2008). علاوه بر تغییرات در مورفولوژی و تقسیم سلولی که در ابتدای فاز استولون - غده اتفاق می‌افتد، آغاز غده‌دهی و بزرگ شدن به همراه تغییرات گسترده‌ای در فیزیولوژی و متابولیسم همراه است. در طی بزرگ

مسیر بیوستتز فنیل‌پروپانوئید مهم‌تر هستند که حتی بدون وجود تنش‌های زیستی و یا غیرزیستی گیاه را در حالت آماده‌باش قرار می‌دهد. با مطالعه انجام شده می‌توان ادعان کرد، تنها یک ژن برای عملی خاص در گیاه وجود ندارد، بلکه گروهی از ژن‌ها؛ مثلاً گروهی از پراکسیدازها یا گروهی از کربوکسی-لیازها، دخیل هستند که الزاماً از روند ثابت افزایش یا کاهش بیان خاصی پیروی نمی‌کنند. به‌همین دلیل با وجود مطالعات بی‌شمار روی گیاه سیب‌زمینی هنوز ابهامات زیادی وجود دارد که نیاز به بررسی‌های بیشتر برای درک ارتباط این ژن‌ها دارد، تا بتوان مراحل رشد را موشکافانه‌تر زیر نظر قرار داد.

سپاسگزاری

از آقای مهندس مرتضی هادی‌زاده به دلیل کمک در آنالیز برخی از داده‌ها سپاسگزاری می‌شود. همچنین از مرکز پردازش‌های سریع دانشگاه شهید چمران اهواز، تشکر و قدردانی می‌گردد.

REFERENCES

- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30 (15):2114-2120.
- Bradshaw JE (2010) Root and tuber crops, vol 7. Springer Science & Business Media.
- Castro AHF, Coimbra MC, da Fonseca STD, de Melo Souza AA (2020) Phenylalanine ammonia-lyase in higher plants: A key enzyme for plant development. *Advances In Chemistry Research*.
- Cenzano A, Vigliocco A, Kraus T, Abdala G (2003) Exogenously applied jasmonic acid induces changes in apical meristem morphology of potato stolons. *Annals of botany* 91 (7):915-919.
- Chang S, Puryear J, Cairney J (1993) A simple and efficient method for isolating RNA from pine trees. *Plant molecular biology reporter* 11 (2):113-116.
- Chen L, Lu Y, Hu Y, Xue X (2020a) RNA-Seq reveals that sucrose-free medium improves the growth of potato (*Solanum tuberosum* L.) plantlets cultured in vitro. *Plant Cell, Tissue and Organ Culture (PCTOC)* 140 (3):505-521.
- Chen Y, Li C, Yi J, Yang Y, Lei C, Gong M (2020b) Transcriptome Response to Drought, Rehydration and Re-Dehydration in Potato. *International Journal of Molecular Sciences* 21 (1):159.

که موجب نگهداشتن تقابل گیاه- میکروب می‌شود، سیستم‌های آنتی‌اکسیدان و یا ترکیبات جاذب UV، رنگدانه‌های جاذب گرده‌افشان‌ها و مقاومت به بیماری‌ها است. فنیل‌پروپانوئید شامل گروهی از متابولیت‌های ثانویه می‌باشد که دارای گروه فنیل هستند و به زنجیره کناری سه کربن در ساختار شیمیایی به نام فنیل پروپان می‌چسبند. فلاونوئید، لیگنین، کومارین، تانین و لیگنان از اجزای کلی در مسیر فنیل‌پروپانوئید هستند. بیشتر ژن‌های فعال در این مطالعه پراکسیدازها می‌باشند که به طور مشترک در آنالیز خوشه‌بندی و آنالیز مسیرها وجود داشتند.

به‌طورکلی، در جریان مراحل نموی غده سیب‌زمینی، فرایندهای مولکولی بسیاری در جریان است. در مراحل اولیه رشد و نمو غده متابولیسم سنتز نشاسته تنها یکی از مسیرهای متابولیکی درگیر طی این فرایند است. با وجود اهمیت نشاسته، وجود نشاسته برای رشد و نمو غده ضروری نمی‌باشد و در مراحل اولیه غده‌دهی مسیرهای دیگری از جمله

- Dahal K, Li X-Q, Tai H, Creelman A, Bizimungu B (2019) Improving potato stress tolerance and tuber yield under a climate change scenario—a current overview. *Frontiers in plant science* 10.
- Dennis G, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA (2003) DAVID: database for annotation, visualization, and integrated discovery. *Genome biology* 4 (9): R60.
- Ewing E, Struik P (1992) Tuber formation in potato: induction, initiation, and growth. *Horticultural reviews* 14 (89): 197.
- Faostat F (2018) Crops. Cow peas, dry Disponível em.
- Hung J-H, Yang T-H, Hu Z, Weng Z, DeLisi C (2012) Gene set enrichment analysis: performance evaluation and usage guidelines. *Briefings in bioinformatics* 13 (3):281-291.
- Jackson SD, James P, Prat S, Thomas B (1998) Phytochrome B affects the levels of a graft-transmissible signal involved in tuberization. *Plant Physiology* 117 (1):29-32.
- Kanehisa M, Goto S (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* 28 (1): 27-30.
- Kinsella RJ, Kähäri A, Haider S, Zamora J, Proctor G, Spudich G, Almeida-King J, Staines D, Derwent P, Kerhornou A (2011) Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database 2011.
- Kloosterman B (2006) Transcriptomic analysis of potato tuber development and tuber quality traits using microarray technology: in search of candidate genes. Univ.
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nature methods* 9 (4):357.
- Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC bioinformatics* 12 (1): 323.
- Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology* 15 (12):550.
- Massa AN, Childs KL, Lin H, Bryan GJ, Giuliano G, Buell CR (2011) The transcriptome of the reference potato genome *Solanum tuberosum* Group Phureja clone DM1-3 516R44. *Plos one* 6 (10).
- Papathanasiou F, Mitchell SH, Watson S, Harvey BMR (1999) Effect of environmental stress during tuber development on accumulation of glycoalkaloids in potato (*Solanum tuberosum* L.). *Journal of the Science of Food and Agriculture* 79 (9):1183-1189.
- Rahnama H, Ebrahimzadeh H (2005) The effect of NaCl on antioxidant enzyme activities in potato seedlings. *Biologia plantarum* 49 (1):997-3.
- Sanchez PAG, Babujee L, Mesa HJ, Arcibal E, Gannon M, Halterman D, Jahn M, Jiang J, Rakotondrafara AM (2020) Overexpression of a modified eIF4E regulates potato virus Y resistance at the transcriptional level in potato. *BMC genomics* 21 (1):18.
- Sandhu S, Marwaha R, Pandey S (2007) Peroxidase as a biochemical marker of maturity levels in potato (*Solanum tuberosum*) cultivars grown under short days. *New Zealand journal of crop and horticultural science* 35 (1):171-175.
- Sarkar D (2008) The signal transduction pathways controlling in planta tuberization in potato: an emerging synthesis. *Plant Cell Reports* 27 (1):1-8.
- Schwall GP, Safford R, Westcott RJ, Jeffcoat R, Tayal A, Shi Y-C, Gidley MJ, Jobling SA (2000) Production of very-high-amylose potato starch by inhibition of SBE A and B. *Nature biotechnology* 18 (5):551-554.
- Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* 4 (1):44.

- Struik P, Ewing E (1995) Crop physiology of potato (*Solanum tuberosum*): responses to photoperiod and temperature relevant to crop modelling. In: Potato ecology and modelling of crops under conditions limiting growth. Springer, pp 19-40.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* 102 (43):15545-15550.
- Tang H, Zhang X, Gong B, Yan Y, Shi Q (2020) Proteomics and metabolomics analysis of tomato fruit at different maturity stages and under salt treatment. *Food Chemistry* 311:126009.
- Tian T, Liu Y, Yan H, You Q, Yi X, Du Z, Xu W, Su Z (2017) agriGO v2. 0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic acids research* 45 (W1):W122-W129.
- Tiwari JK, Buckseth T, Zinta R, Saraswati A, Singh RK, Rawat S, Dua VK, Chakrabarti SK (2020) Transcriptome analysis of potato shoots, roots and stolons under nitrogen stress. *Scientific Reports* 10 (1):1-18.
- Vreugdenhil D, Struik PC (1989) An integrated view of the hormonal regulation of tuber formation in potato (*Solanum tuberosum*). *Physiologia plantarum* 75 (4):525-531.
- Xu X, Vreugdenhil D, Lammeren AAV (1998) Cell division and cell enlargement during potato tuber formation. *Journal of Experimental Botany* 49 (320):573-582.
- Yoon S, Kim S-Y, Nam D (2016) Improving gene-set enrichment analysis of RNA-Seq data with small replicates. *PloS one* 11 (11).