

توالی‌یابی *de novo* ترانسکرپتوم و آنالیز عملکردی ژن‌های با بیان متفاوت بذرهای بزرگ و کوچک گیاه توق خاردار (*Xanthium strumarium* L.) در دوره نمو

ایمان نعمتی¹، محمد صدقی²، قاسم حسینی سالکده³، رضا توکل افشاری⁴، محمدرضا تقوی⁵

1- دانشجوی دوره دکتری علوم و تکنولوژی بذر دانشگاه محقق اردبیلی

2- استاد دانشکده کشاورزی دانشگاه محقق اردبیلی

3- استاد پژوهشکده بیوتکنولوژی کشاورزی ایران

4- استاد گروه آگروتکنولوژی، دانشکده کشاورزی دانشگاه فردوسی مشهد

5- استاد پردیس کشاورزی و منابع طبیعی کرج دانشگاه تهران

(تاریخ دریافت: 1397/04/30؛ تاریخ پذیرش: 1397/10/12)

چکیده

گیاه توق خاردار (*Xanthium strumarium* L.) گیاهی هرز با ارزش دارویی بالا است که به دلیل وجود خواب عمیق فقط در یکی از دو بذر درون یک کپسول این گیاه، برای مطالعات خواب بذر بسیار مورد علاقه پژوهشگران است. با این وجود، نبود اطلاعات ژنومیک در این گیاه، سبب شده است که دانسته‌ها در مورد آن در سطح بسیار پایینی باشد. در این تحقیق، ترانسکرپتوم بذرهای کوچک و بزرگ این گیاه در طول دوره نمو با استفاده از پلتفرم Illumina، با هدف شناسایی و بررسی عملکردی رونوشت‌های با بیان متفاوت میان دو بذر، توالی‌یابی شد. توالی‌های شناسایی شده بین دو بذر با یکدیگر مقایسه شدند و رونوشت‌های با بیان متفاوت بین دو بذر مورد تفسیر عملکردی قرار گرفتند. در این پژوهش تعداد 191192 توالی ژن با میانگین طول 989/69 جفت‌باز شناسایی شد. آنالیزهای تشابه توالی و تفسیر عملکردی این توالی‌ها بر علیه پایگاه داده nr، آنالیز GO و KEGG انجام شد. ژن‌های با بیان متفاوت، از نظر top hits، بیشتر تشابه را با آفتابگردان زراعی (83/41 درصد) داشتند. آنالیز GO سبب شناسایی 615 تفسیر عملکردی در 36 گروه شد. فراوان‌ترین GO یافت شده در دسته فرایند بیولوژیک، فرایندهای بیوسنتزی بود. نتایج این تحقیق فرایندهای بیوسنتزی و متابولیک بیشتر را در بذر بزرگ نسبت به بذر کوچک‌تر و نقش تنظیمی کلیدی رونویسی را در دوره نمو بذر نشان داد.

کلمات کلیدی: توق خاردار، ترانسکرپتوم، بذرهای بزرگ و کوچک، نمو، توالی‌یابی

De novo transcriptome sequencing and functional annotation of differentially expressed genes in large and small seeds of common cocklebur (*Xanthium strumarium* L.) during seed development

I. Nemati¹, M. Sedghi^{2*}, Gh. Hoseini Salekdeh³, R. Tavakol Afshari⁴, M.R. Naghavi⁵

1- PhD student, Seed Sciences and Technology

2- Professor, Faculty of Agriculture, University of Mohaghegh Ardabili (corresponding author)

3- Professor, Agricultural Biotechnology Institute of Iran

4- Professor, Department of Agrotechnology, Faculty of Agriculture, Ferdowsi University of Mashhad

5- Professor, Karaj Agriculture and Natural Resources, University of Tehran

(Received: Jul. 21, 2018 – Accepted: Jan. 02, 2019)

Abstract

Common cocklebur (*Xanthium strumarium* L.) is a widespread weed with high medical value and interested for researchers because of deep dormancy in one of its two seeds in one bur. However, lack of genomic data has led to low information about it. Transcriptome large and small seeds were sequenced using Illumina platform to identify and functional analysis of differentially expressed transcripts in two seeds. Identified sequences in each seeds were compared and differentially expressed genes were functionally annotated. In this research 191192 sequence with a mean of 989.69 bp were detected. Sequence similarity analysis and functional analysis was carried out against nr, GO and KEGG databases. Differentially expressed genes had the most similarity with sunflower (83.41 per cent) in terms of top hits. GO analysis led to identify 615 functional annotation distributed in 36 categories. The most abundant GO in biological process was biosynthesis. Results of our research shows the higher biosynthetic and metabolic processes in large seed of certain bur rather than the small one and also key regulatory role of transcription during seed development.

Keywords: common cocklebur, transcriptome, large and small seeds, development, sequencing

* Email: m_sedghi@uma.ac.ir

مشابهی نیز که نشان‌دهنده الگوهای دقیق جداسازی مکانی mRNA در بذر هستند، در آزمایشاتی که از Affymetrix soybean GeneChip برای آنالیز زیرنمونه‌های چهار مرحله نموی بذر استفاده شد، به دست آمد و اختصاصی بودن برخی از رونوشت‌ها را برای بذر، و نه به صورت مشترک با اندام‌های هوایی، نشان داد (Le et al. 2007).

تکنیک توالی‌یابی NGS¹ مزایای زیادی نسبت به ریزآرایه² دارد که شامل توانایی یافتن رونوشت‌های فراوانی اندک، تعیین ایزوفرم‌های جدید mRNA، نیاز کمتر به دانش *priori* در ارتباط با موجود، حساسیت بالا در کشف ژن‌ها با بیان افتراقی، نتایج با تکرارپذیری بیشتر، توانایی مقایسه الگوی بیان بین موجودات جداگانه مرتبط با یکدیگر، می‌باشند. به عنوان مثال، توالی‌یابی RNA، مطالعه تجمع روغن را در چهار دانه روغنی غیر مدل (و یا در حال مدل شدن)، شامل کرچک³، کلزا⁴، شمشاد بالدار⁵ و لادن⁶ (Troncoso-Ponce et al. 2011) تسهیل کرد. این گونه‌ها در محل تجمع روغن، ترکیب و مقدار تری‌گلیسرول، با یکدیگر تفاوت دارند. آنالیز نتایج موجب شناسایی مسیر آنزیمی درگیر در تولید تری‌گلیسرول شد که الگوی بیان آن از نظر زمانی در تمام گونه‌ها مشابه بود و نشان‌دهنده ارتباط تکاملی در تولید روغن‌های بذر می‌باشد. دو تحقیق عمده نیز در رابطه با سویا در این زمینه وجود دارد، یکی آنالیز کل بذر سویا در هفت مرحله و در بازه زمانی 10 تا 42 روز پس از باروری (Severin et al. 2010) و مطالعه دیگر، آنالیز کل بذر سویا در بازه 15 تا 65 روز پس از باروری (Chen et al. 2012) بود. این مطالعات نشان دادند که 49151 رونوشت در دوره نمو بذر یافت شدند که تقریباً 12000 mRNA، بیش از 37500 رونوشت به دست آمده از آرایه

مقدمه

یکی از پیچیدگی‌های مطالعه ژنومیک بذراین است که برخلاف سایر بافت‌های گیاه، بذر دارای سه بخش است که از نظر ژنتیکی کاملاً متفاوت از یکدیگر هستند: جنین (2n) با 50 درصد منشا مادری و 50 درصد منشا پدری، اندوسپرم (3n) با دو سوم منشا مادری و یک‌سوم منشا پدری، پوسته (2n) با منشا کاملاً مادری. این موضوع زمانی پیچیده‌تر می‌شود که بدانیم در دوره نمو بذر، بیان ژن‌های مختلف در بافت‌های مختلف بذر به طور ناشناخته و کنترل‌شده‌ای در نوسان است (Belmonte et al. 2013) و برای شناسایی روند سطح بیان، باید مقیاس‌ها برای جداسازی زمانی و مکانی نمونه‌های برداشت شده را به شدت کوچک‌تر انتخاب کرد. اوج پیچیدگی نیز زمانی ظاهر می‌شود که بدانیم، ژنوم مادری و پدری، علاوه بر توزیع نامتقارن مکانی در بافت‌های مختلف بذر، بیان متفاوت زمانی در دوره نمو بذر دارند (Autran et al. 2011). به طور کلی گذار از پدر و مادر به تخم بارور شده فعال، پیچیدگی‌هایی منحصر به فرد دارد که در سایر اندام‌های بذر وجود ندارد.

یکی از جامع‌ترین پژوهش‌های نمو بذر، با هدف شناسایی تمام ژن‌ها و معرفی شبکه‌های تنظیم ژنی مسئول نمو بذر توسط Belmonte و همکاران (2013) به چاپ رسیده است (Belmonte et al. 2013). 36 زیربخش بذر در طی 5 مرحله نموی *Arabidopsis*، الگوهای غالب پیچیده‌ای را از فعالیت ژن‌ها، هم در فضا و هم در زمان، نشان داد. ترکیب LMD و ATH1 GeneChip، موجب شناسایی 17594 mRNA مجزا گردید که در مراحل نمو بذر قابل یافت می‌باشند و از بین این mRNA ها حداقل 1316 عدد، به طور ویژه در بذر *Arabidopsis*، و نه در اندام‌های رویشی و زایشی، بیان می‌شوند (Belmonte et al. 2013). آن‌ها همپوشانی غیرقابل انتظاری را نیز در ویژگی‌های عملکردی زیربخش‌های مختلف بذر شناسایی کردند. نتایج

¹ Next-Generation Sequencing

² Microarray

³ *Ricinus communis*

⁴ *Brassica napus*

⁵ *Euonymus alatus*

⁶ *Tropaeolum majus*

موجود در یک غلاف از نظر ابعاد و جوانه‌زنی با یکدیگر تفاوت دارند، انتظار بر این بود که تفاوت در الگوی بیان ژن‌های دو بذر در دوره نمو، عامل بروز این تفاوت‌ها باشد. بنابراین برای شناسایی چگونگی بیان ژن‌های موثر بر نمو بذر و بررسی بیان افتراقی این ژن‌ها در دو بذر درون یک میوه، این آزمایش اجرا شد.

مواد و روش‌ها

بذرهای گیاه توق خاردار (*Xanthium strumarium* L.) در پاییز 95 از مزرعه پژوهشی پردیس کشاورزی و منابع طبیعی کرج، واقع در 35/80 درجه و 92 دقیقه و 21 ثانیه شمالی و 50/96 درجه و 70 دقیقه و 44 ثانیه شرقی در ارتفاع 1300 متر از سطح دریا، جمع‌آوری شده و پس از یک نسل کشت و بذرگیری، بذرهای فاقد خواب نسل بعد برای کشت در آزمایش اصلی مورد استفاده قرار گرفتند. به منظور کشت بذرهای برداشت شده از کشت اول، بذرهای ضد عفونی شده، در پتری‌های شیشه‌ای با قطر 10 سانتی‌متری قرار داده شدند که کف آن‌ها دو لایه کاغذ واتمن قرار داشت و به مدت 14 روز در دمای 23 درجه سلسیوس در دستگاه ژرمیناتور قرار گرفتند تا گیاهچه‌ها ظاهر شوند. سپس گیاهچه‌ها به گلدان‌های 4 لیتری حاوی 50 درصد پیت و 50 درصد خاک مزرعه منتقل شدند. گلدان‌ها در اتاقک‌های رشد و در دمای 16-18 (شب-روز) درجه سلسیوس و طول دوره روشنایی 8 ساعت با شدت نور 9000 لوکس، به مدت 4 ماه نگهداری شدند و پس از آن به گلخانه با دمای 21-25 (شب-روز) درجه سلسیوس و طول دوره روشنایی 14 ساعت منتقل شدند تا وارد فاز زایشی شوند. پس از ورود به فاز زایشی، در دوره نمو بذر به ترتیب در زمان‌های 3، 10، 20، 30 روز پس از گرده‌افشانی و یک مرحله نیز از بذر بالغ، نمونه‌برداری انجام شد. کپسول‌ها به دقت توسط تیغ استریل از گیاه مادری جدا و بلافاصله در ازت مایع قرار داده شدند. سپس بذرهای کوچک و بزرگ موجود در هر میوه، در

جاری Affymetrix سویا بود. علاوه بر آن، رونوشت‌های دارای بیان متفاوت نسبت به مراحل ابتدایی نمو بذر، بین 14058-9930 (Severin et al. 2010) تا 16255-11592 (Chen et al. 2012) گزارش شده‌اند.

با آغاز دوره ژنومیک فعلی، از بسیاری از مسیرهای تنظیمی و نموی مسئول ساخت بذر، پرده‌برداری شده است. با این حال هنوز تا شناخت کامل مکانیسم‌های هماهنگ‌کننده فعالیت ژن‌ها، طی نمو بخش‌ها و زیربخش‌های مختلف بذر، راه زیادی در پیش داریم. بسیاری از مکانیسم‌های تنظیم‌کننده مربوط به متابولیسم اولیه و ثانویه، تنظیم هورمونی، نقش‌پذیری¹ ژن‌ها و تنظیمات رونویسی، ترجمه‌ای و پس‌ترجمه‌ای، همگی به صورت هماهنگ با یکدیگر، برای راه‌اندازی فرایندهای پیچیده طی نمو بذر، فعالیت می‌کنند. این فعالیت‌ها توسط صدها و یا هزاران ژن کنترل می‌شوند که اغلب به دلیل پیچیدگی‌های ژنتیکی، مبهم و ناشناخته هستند و شناسایی آنها از طریق روش‌های غربال ژنتیکی رو به جلو، بسیار سخت است (Curtin et al. 2011). بهترین راه برای مطالعه فرایندهای هماهنگی نظیر سرنوشت سلولی، تمایز و ریخت‌زایی در بذر در حال نمو، مشاهده بیان تعداد زیادی از مجموعه‌های ژنی با تکنیک‌های مختلف و استراتژی‌های توالی‌یابی است. سپس آنالیزهای بیوانفورماتیک می‌تواند برای شناسایی شبکه‌ها و تنظیم‌کننده‌های کلیدی نمو بذر، مورد استفاده قرار گیرد. هدف از انجام این تحقیق، بررسی رونوشت‌های مربوط به بذرهای گیاه توق خاردار (*Xanthium strumarium* L.) در دوره نمو و مقایسه این رونوشت‌ها با یکدیگر بود. این گیاه دارای کپسول‌هایی است که در هر کدام، دو بذر قرار دارند؛ یک بذر بزرگ و بدون خواب و یک بذر کوچک‌تر و دارای خواب عمیق. این موضوع سبب حذف تاثیر متفاوت عوامل محیطی بر بذرهای مورد مطالعه می‌شود و می‌توان مقایسه دقیق‌تری بین دو بذر از نظر الگوی ترانسکرپتوم انجام داد. با توجه به اینکه دو بذر

¹ Imprinting

Trinity صورت De novo با استفاده از نرم افزار Trinity (Grabherr et al. 2011) انجام شد. در این نرم افزار، همگذاری و همردیف کردن و اتصال توالی های کوتاه برای بازسازی رونوشت و تشکیل کانتینگ ها مورد استفاده قرار می گیرد. برای شناسایی تک ژن هایی که بیان متفاوت داشتند از نرم افزار EdgeR (Robinson, McCarthy, and Smyth 2010) و بسته نرم افزاری RSEM (Li and Dewey 2011) استفاده شد. نرم افزارهای عملکردی ژن های دارای بیان متفاوت مورد استفاده قرار گرفتند.

تفسیر عملکردی با مقایسه توالی های همترازهای دسته بندی شده با پایگاه های داده عمومی انجام شد. جستجوی تشابه توالی های یونی ترانسکرپت ها نیز با ابزار BLASTX صورت گرفت. همولوگ تک ژن های همتراز شده نیز در NCBI non-redundant protein (Nr) با استفاده از پارامترهای پیش فرض جستجو شدند. نتایج BLASTX به دست آمده بر علیه Nr در مجموعه Blast2GO وارد شدند تا عملیات mapping، بازیابی Gene Ontology (GO) و unique enzyme code (EC) بر تک ژن های همتراز شده صورت گیرد. GO های به دست آمده، به توالی های مورد بررسی اختصاص داده شدند و ژن های حاضر در ترانسکرپتوم، در سه دسته سلولی، عملکرد ملکولی و فرایند بیولوژیک دسته بندی شدند.

رسم شکل های مربوط به تفسیر نتایج کلی و مسیرهای بیوسنتزی، در محیط Blast2Go انجام شد.

نتایج و بحث

توالی یابی Paired End کتابخانه cDNA و همترازی De novo رونوشت ها
نمونه های RNA استفاده شده در این آزمایش دارای کیفیت بالایی بودند. میانگین عدد RIN مربوط به نمونه ها، 8/4 بود که کیفیت بسیار عالی آنها را نشان می دهد.

آزمایشگاه، توسط تیغ استریل جدا شده و تا زمان استخراج RNA در فریزر 80- درجه سلسیوس نگهداری شدند.

استخراج RNA

استخراج RNA با استفاده از محلول TRIzol (Invitrogen) و بر اساس دستورالعمل شرکت سازنده انجام شد. با استفاده از دستگاه نانودراپ و ژل الکتروفورز، کمیت و کیفیت RNA به دست آمده ارزیابی شد. همچنین عدد RIN نمونه ها نیز در انستیتو ژنومیکس پکن (BGI) توسط Agilent 2100 Bioanalyzer برآورد شد.

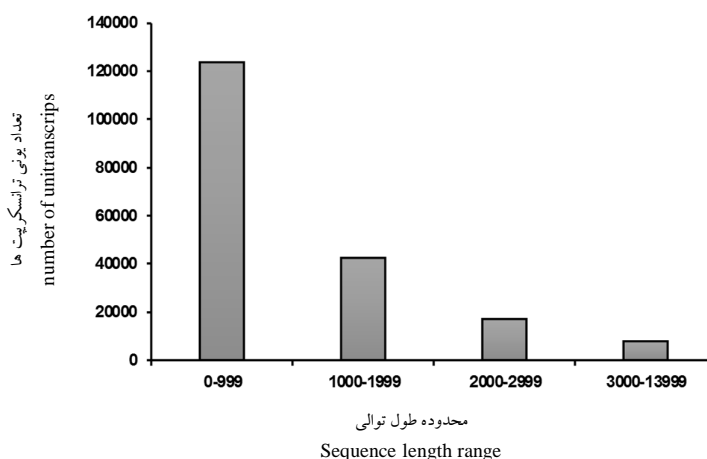
ساخت کتابخانه cDNA

RNA استخراج شده، تحت تیمار DNase قرار گرفت؛ سپس mRNA های دارای دم پلی A، با استفاده از Oligo dT جداسازی شدند. به دنبال آن، mRNA ها به قطعات کوچکتر تبدیل شده و با استفاده از آنزیم ترانسکرپتاز معکوس به cDNA تبدیل شدند. رشته دوم با استفاده از DNA پلیمراز سنتز شد و زنجیره پلی A به آن افزوده شده و آداپتور ویژه توالی یابی نیز به آن متصل شد. توالی یابی با استفاده از کیت Illumina و بر اساس دستور شرکت سازنده انجام شد. تمام مراحل ساخت cDNA و توالی یابی، در انستیتو ژنومیکس پکن (BGI) و با استفاده از پلتفرم Illumina HiSeq2000 انجام شد.

آنالیز نتایج توالی یابی

نتایج توالی یابی توسط شرکت BGI در قالب فایل های fastq ارسال شدند. ارزیابی و کنترل کیفیت اولیه، بر اساس اطلاعات کیفیت توالی یابی برای هر نوکلئوتید صورت گرفت. آداپتورها و خوانش های دارای نوکلئوتیدهای ناشناخته، پیش از آنالیزهای اصلی، توسط نرم افزار Fastx Toolkit حذف شدند و شاخص کیفیت >20 Q برای غربال خوانش های با کیفیت و استفاده در آنالیزهای اصلی، به کار رفت. بازسازی رونوشت ها به

کلاستر شده در این همگذاری، در گروه توالی‌هایی دارای طول کمتر از 500 جفت‌باز (81362 توالی) بود و پس از آن 42398 توالی دارای طول بین 500 تا 999 جفت‌باز، 24955 توالی دارای طول 1000 تا 1499 جفت‌باز، 17636 توالی با طول 1500 تا 1999 جفت‌باز، 10940 توالی با طول 2000 تا 2499 جفت‌باز و 6085 توالی دارای طول 2500 تا 2999 جفت‌باز بودند. توزیع تعداد رونوشت‌ها بر اساس طول، در شکل 1 نشان داده شده است.



شکل 1- توزیع تعداد توالی‌های شناسایی شده در بذره‌های کوچک و بزرگ تونق خاردار بر اساس طول توالی

Figure 1- Identified sequences distribution in small and large seeds of common cocklebur according to length

وجود داشت و درصد اندکی (0/16 درصد) از کل توالی‌های شناسایی شده، میان دو بذر تفاوت معنی‌دار نشان دادند (شکل 2B).

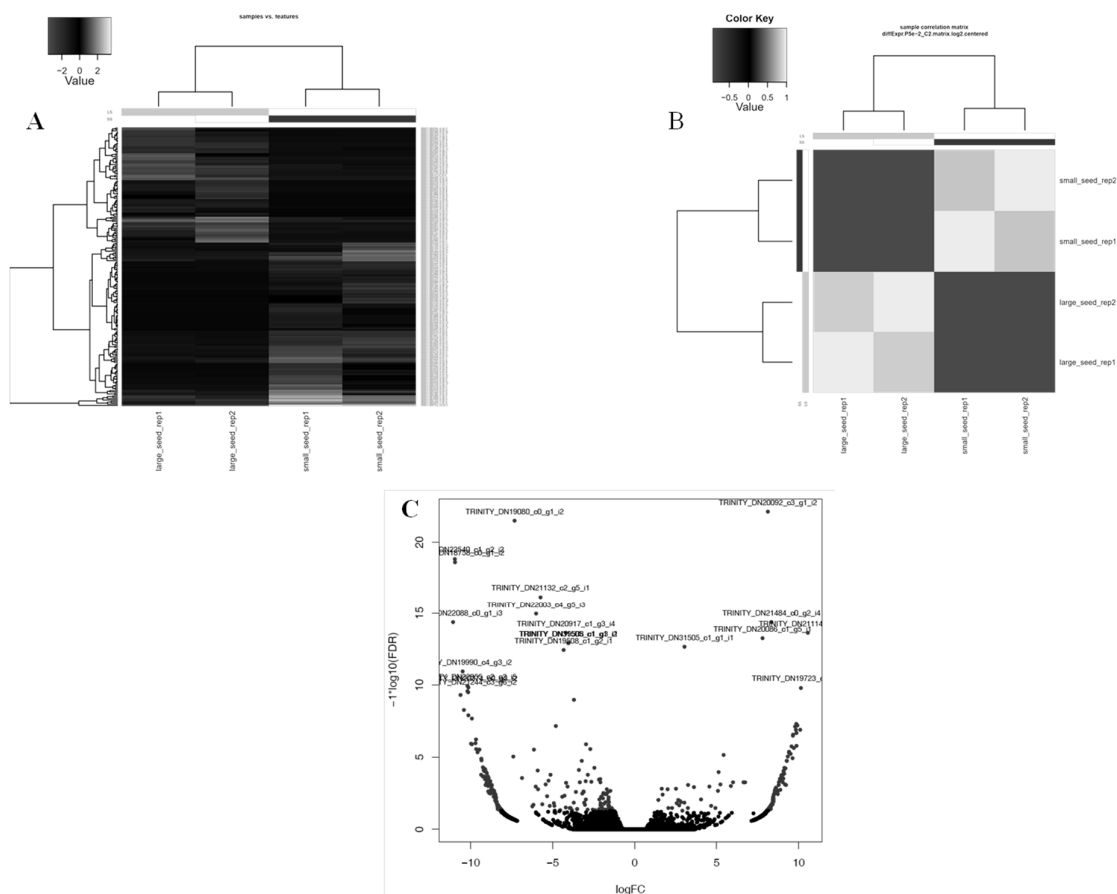
بر اساس نمودار حرارتیبه دست آمده، تعداد ژن‌هایی که افزایش بیان داشتند، در بذر کوچک مطالعه شده، بیش از بذر بزرگ بودند (شکل 2A). نمودار آتشفشانی نیز توزیع کلی تک‌ژن‌های با بیان متفاوت را در دو نوع بذر مورد مطالعه نشان می‌دهد (شکل 2C).

تفسیر عملکردی ترانسکرپتوم بذره‌های در حال نمو تونق خاردار

تفسیر تک‌ژن‌های تراز شده دارای بیان متفاوت با استفاده از ابزار BLASTX بر علیه پایگاه داده Nr با E -value $< 1e^{-6}$ انجام شد.

توالی‌یابی Paired End کتابخانه cDNA با استفاده از HiSeqTM 2000 میلیون‌ها خوانش با طول 150bp تولید کرد. پس از حذف توالی‌های آداپتور و خوانش‌های دارای کیفیت پایین با $\text{phred score} < 20$ ، طول کل خوانش‌های با کیفیت بالا 189220182 بود. با استفاده از همگذار Trinity تعداد 191192 توالی ژن ایجاد شد که طول آنها در محدوده 201 تا 13863 با میانگین 989/69 و N50 برابر 1620 جفت‌باز بود. بیشترین تعداد رونوشت‌های

الگوی بیان ژن‌های با بیان متفاوت، در شکل 2 نشان داده شده است. مقایسه توالی‌های شناسایی شده در دو بذر و مقایسه آن‌ها با یکدیگر در این آزمایش نشان داد که رونوشت‌های مربوط به 324 ژن از نظر فراوانی، متفاوت بودند. از میان این ژن‌ها، 188 ژن، کاهش بیان داشتند به طوری که میزان کاهش بیان، بین 2/04 تا 11/06 برابر و با میانگین 7/20- بود. رونوشت‌های مربوط به 136 ژن نیز با میانگین log FC برابر 7/80، افزایش بیان نشان دادند. بر اساس این نتایج، تعداد ژن‌هایی که در بذر بزرگ‌تر افزایش بیان داشتند، بیش از ژن‌هایی بود که کاهش بیان نشان دادند (شکل 2A). با این وجود، با توجه به تعداد کل توالی‌های به دست آمده در این آزمایش، همپوشانی بسیار بالایی میان دو بذر از نظر بیان ژن‌ها در دوره نمو بذر



شکل 2- الگوی بیان متفاوت ژن‌ها در دو بذر بزرگ و کوچک درون کپسول بذر توق خاردار.

A: منحنی حرارتی؛ B: ماتریکس همبستگی میان نمونه‌ها؛ C: منحنی آتشفشانی

Figure 2- Identified sequences distribution according to length.

A: thermal plots; B: sample correlation matrix; C: volcano plot

زراعی (*Helianthus annuus* L.) به میزان 13/30 درصد و پس از آن نیز *Cynara cardunculus* var. *scolymus* با 4/92 درصد تشابه بود (شکل 5B). از نظر Top hits نیز بیشترین تشابه میان تک‌ژن‌های توق خاردار با آفتابگردان زراعی و *Cynara cardunculus* var. *scolymus* بود که به ترتیب 83/41 درصد و 6/63 درصد تشابه بودند (شکل 5A).

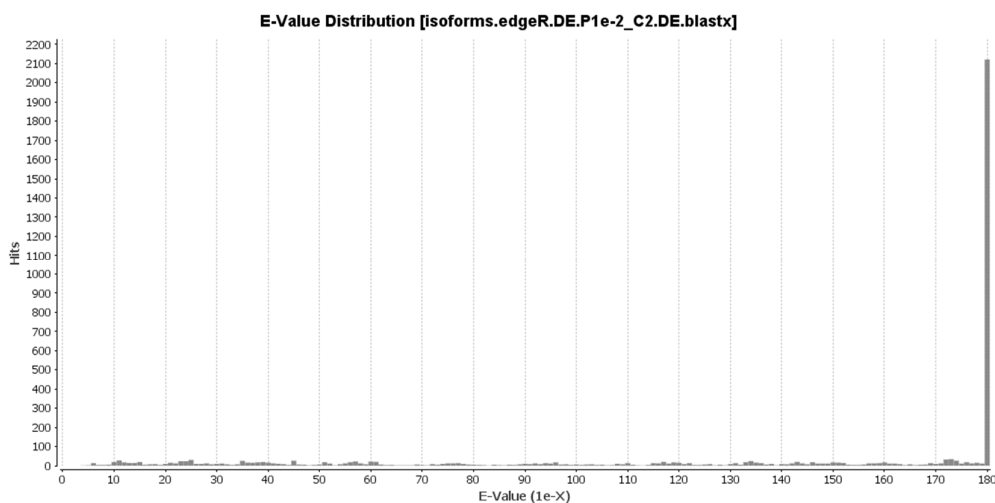
رونوشت‌ها با بیان متفاوت در بذرهای در حال نمو توق خاردار، بر اساس هومولوگی توالی، به GOهای مختلف اختصاص داده شدند. در مجموع بر اساس نتایج حاصل از BLAST+، 615 تفسیر به دست آمد (شکل 6A). این GOها در 36 گروه عملکردی توزیع شده که

آنالیز توزیع E -value بر اساس نتایج تفسیر پایگاه داده Nr نشان داد که به ترتیب 50/20 و 71/53 درصد از توالی‌های تطبیق‌یافته، دارای E -value تشابهی بین 10^{-6} تا 10^{-15} و E -value $< 10^{-20}$ بودند و 39/22 درصد از توالی‌های تطبیق‌یافته، از نظر E -value تشابهی، بین 10^{-15} تا 10^{-34} بودند (شکل 3). آنالیز توزیع تشابه BLAST hits نشان داد که 56/30 درصد از کل توالی‌ها تشابه بیش از 80 درصد را نشان دادند (شکل 4).

بر اساس آنالیز توزیع گونه‌ای، تشابه تک‌ژن‌های دارای بیان متفاوت توق خاردار در چندین گونه گیاهی دیگر یافت شده‌اند (شکل‌های B و 5A). بیشترین شباهت با آفتابگردان

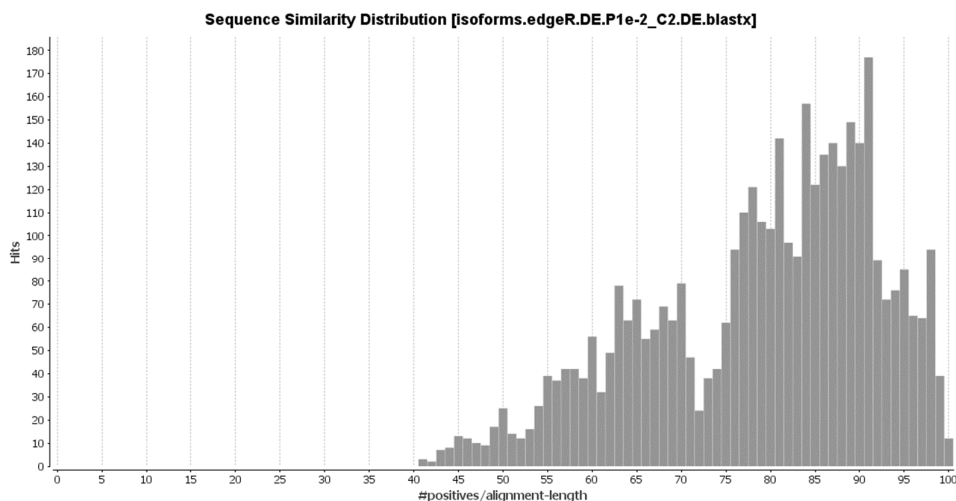
بیوسنتزی، فرایند متابولیک ترکیبات نیتروژنی سلولی و فرایند اصلاحات پروتئینی سلولی با اختصاص 9، 9 و 13 توالی، فراوان‌ترین GOها بودند (شکل 6C). در گروه عملکرد مولکولی نیز، فراوان‌ترین GOهای مشاهده شده شامل اتصال به یون (33 توالی)، اتصال به پروتئین (15 توالی) و اتصال به DNA (15 توالی) بودند (شکل 6D).

همگی در سه دسته بخش سلولی، عملکرد مولکولی و فرایند بیولوژیک توزیع شدند. در دسته بخش سلولی، اجزای سلولی، هسته و ترکیبات حاوی پروتئین به ترتیب با 12 و 17 و 24 توالی، فراوان‌ترین GOهای مشاهده شده بودند (شکل 6B). بیشترین گروه‌های مشاهده شده در دسته فرایند بیولوژیک پس از تطبیق توالی‌ها، فرایند



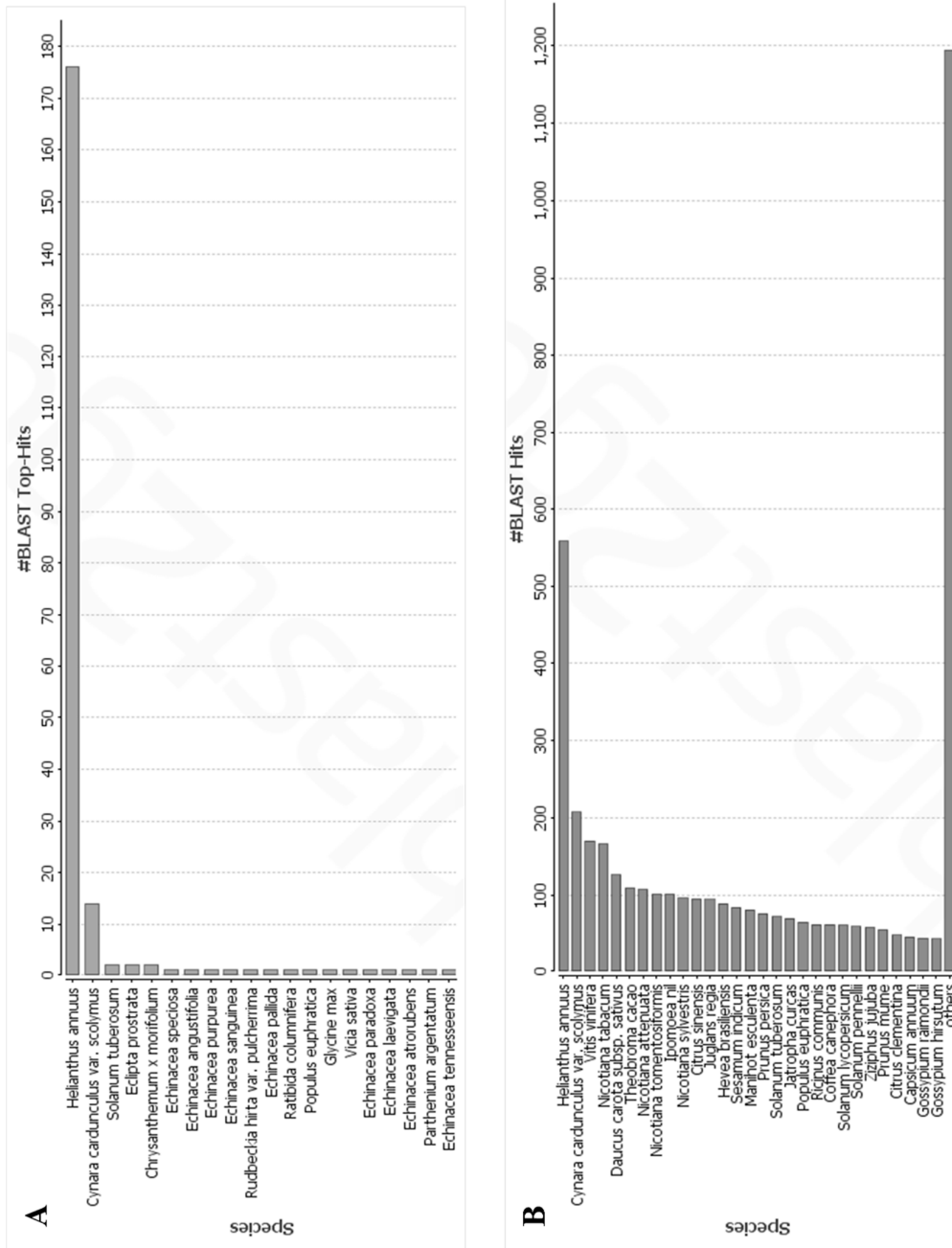
شکل 3- توزیع E-value تک‌ژن‌های تراز شده با بیان متفاوت در دو بذر بزرگ و کوچک درون کپسول تون خاردار

Figure 3- E-value distribution of assembled differentially expressed unigenes in large and small seeds in common cocklebur bur

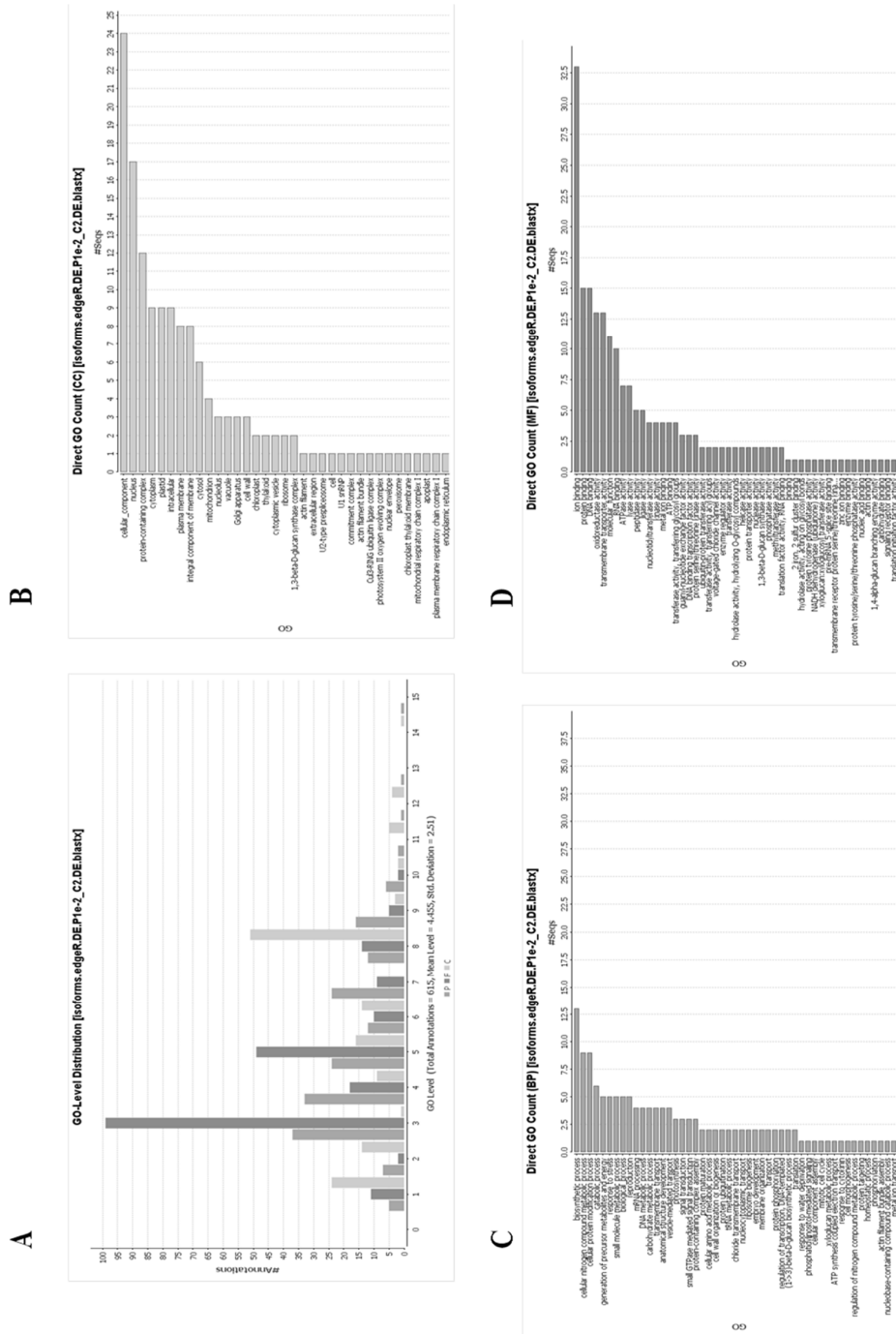


شکل 4- نمودار توزیع تشابه Blast hits مربوط به تک‌ژن‌های تراز شده با بیان متفاوت در دو بذر بزرگ و کوچک درون کپسول تون خاردار

Figure 4- Blast hits similarity of assembled differentially expressed unigenes in large and small seeds in common cocklebur bur



شکل 5- تشابه تک‌ژن‌های شناسایی‌شده با بیان متفاوت در بذرهای بزرگ و کوچک توت خاردار با سایر گونه‌های گیاهی
 Figure 5- Similarity of assembled differentially expressed unigenes in large and small seeds in common cocklebur bur to other plant species. A: Top hits; B: blast hits

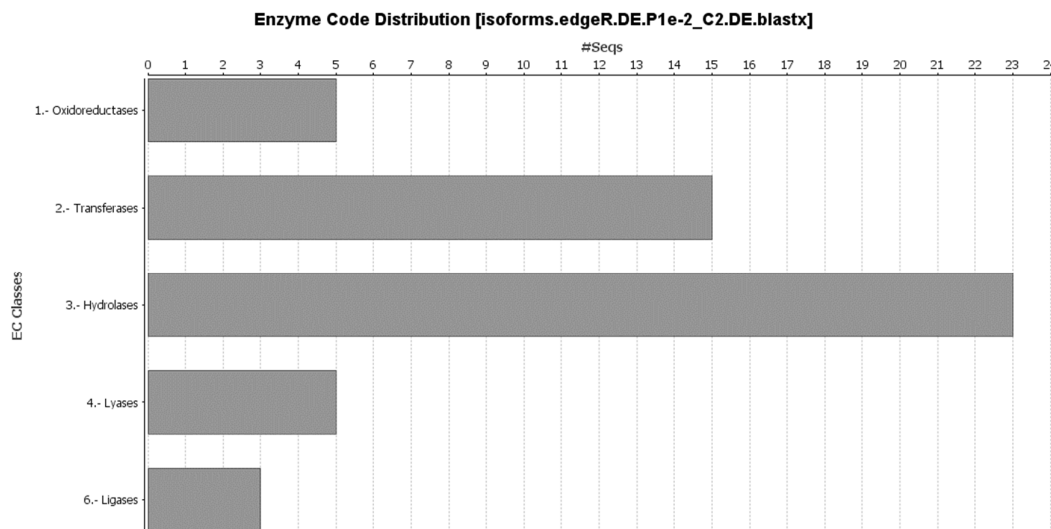


شکل 6- توزیع توالی های تراژ شده با بیان متفاوت در بذر بزرگ و کوچک درون کپسول توق خاردار بر اساس GOهای شناسایی شده. توزیع GO: B; GOهای مربوط به بخش سلولی; C: GOهای مربوط به فرایند بیولوژیک; D: GOهای درون کپسول توق خاردار بر اساس GOهای شناسایی شده.

Figure 6- GO analysis of assembled differentially expressed sequences in large and small seeds in common cocklebur. A: GO distribution; B: GOs in cellular component; C: GOs in biological process; D: GOs in molecular function

درصد)، ترانسفرازها (29/41 درصد)، هیدرولازها (45/10 درصد)، لیازها (9/80 درصد) و لیگازها (5/88 درصد) بودند (شکل 7).

جستجو بر علیه پایگاه داده نشان داد که در مجموع، 51 تک‌ژن با بیان متفاوت، با کدهای آنزیمی مختلف تفسیر شدند. کدهای آنزیمی تفسیر شده در شش گروه دسته‌بندی شدند که شامل اکسیدوردوکتازها (9/80



شکل 7- گروه‌بندی تک‌ژن‌های شناسایی شده با بیان متفاوت بر اساس انطباق با کدهای آنزیمی

Figure 7- Grouping of assembled differentially expressed according to enzyme codes

ذخیره شده در هیدرات کربن‌ها، چربی‌ها و پروتئین‌ها به شکل ATP انجام می‌شود. نتایج به دست آمده از این آزمایش، تغییر در میزان بیان ژن‌های بیان‌کننده دو آنزیم درگیر در این چرخه را نشان داد. این آنزیم‌ها شامل دی‌هیدرولیپویل دهیدروژناز¹ (EC: 1.8.1.4) و اکونیتاز هیدراتاز² (EC: 4.2.1.3) بودند. دی‌هیدرولیپویل دهیدروژناز میتوکندریایی (mtLPD) یک فلاوان آنزیم هومودایمر از خانواده آنزیمی اکسیدوردوکتاز پیریدین نوکلئوتید-دی‌سولفید است که ساختاری حفاظت‌شده داشته و در تمام ارگانیسم‌های هوازی وجود دارد. LPD به همراه پیرووات دهیدروژناز و دی‌هیدرولیپویل ترانس استیلاز، یک بخش ضروری از کمپلکس مولتی آنزیمی پیرووات دهیدروژناز است که وظیفه دکربوکسیلاسیون اکسیداتیو پیرووات برای تشکیل استیل کوآنزیم A و

آنالیز مسیرهای متابولیسمی با استفاده از KEGG

آنالیز مسیرهای متابولیسمی، نقش بسیار مهمی در مطالعات گسترده ژنومیکس دارد. KEGG یکی از پایگاه‌های داده است که اطلاعات ارزشمندی از سیستم‌های بیولوژیک و روابط میان آن‌ها را در خود دارد (Kanehisa et al. 2015). در این آزمایش، تفسیر مسیرهای KEGG با استفاده از یونی‌ترانسکرپت‌های تراز شده با بیان متفاوت در دو بذر که در GOهای مختلف ترسیم شده بودند، صورت گرفت. در مجموع، تعداد EC 26 به 26 مسیر متابولیسمی اختصاص داده شدند و برای نمایش مسیرهای KEGG و ژن‌های فرضی یافت‌شده، مورد استفاده قرار گرفتند.

چرخه TCA

چرخه TCA دومین مرحله از تنفس هوازی است که در تمام موجودات هوازی به منظور آزادسازی انرژی

¹ Dihydrolipoamide dehydrogenase

² aconitate hydratase

افزایش مشاهده شده در رونوشت‌های مربوط به این آنزیم، می‌تواند بیانگر افزایش ایزومریزاسیون سیترات به ایزوسیترات در مسیرهای نام برده باشد که منجر به افزایش فراهمی انرژی در بذر بزرگ درون میوه گیاه توق خاردار خواهد شد.

فسفر یلاسیون اکسیداتیو

مجموعه داده‌های ترانسکریپتوم بذره‌های *X. stumarium* L. نشان داد که برخی از آنزیم‌های مسیر فسفر یلاسیون اکسیداتیو نیز از نظر سطح بیان، تفاوت معنی‌داری با یکدیگر داشتند (شکل 8). افزایش بیان ژن مربوط به آنزیم NADH دهیدروژناز (EC: 1.6.99.3) در بذر بزرگ تر گیاه توق خاردار مشاهده شد. این آنزیم، که یک فلاووپروتین با مرکز آهن-گوگرد است، کاتالیز کننده واکنش احیایی در ابتدای زنجیره انتقال الکترون میتوکندریایی است و با جداسازی یک الکترون از NADH و انتقال آن به پذیرنده الکترون (که در این بخش، یوبی کوینون است) سبب احیای پذیرنده و شکل‌گیری NAD^+ در بخش درونی میتوکندری می‌شود. افزایش بیان آنزیم NADH-یوبی کوینون ردوکتاز (EC: 1.6.5.3) نیز به عنوان بخشی از کمپلکس I انتقال الکترون میتوکندریایی، در این آزمایش مشاهده شد. نتایج به دست آمده از آزمایش حاضر، تغییر در سطح رونوشت مربوط به آنزیم دی‌فسفات فسفویدرولاز یا inorganic diphosphatase را نشان داد (شکل 8). این آنزیم در کمپلکس V زنجیره انتقال الکترون قرار دارد و هیدرولیز دی‌فسفات را به دو ملکول فسفات کاتالیز می‌کند. با توجه به افزایش بیان ژن‌های مربوط به آنزیم‌های درگیر در زنجیره انتقال الکترون میتوکندریایی و ماهیت عمل این زنجیره که آزادسازی انرژی را بر عهده دارد، به نظر می‌رسد که انرژی فراهم شده برای بذر بزرگ‌تر درون میوه توق خاردار بیش از بذر کوچک‌تر است و این موضوع می‌تواند در توجیه رشد بیشتر این بذر به ما کمک کند.

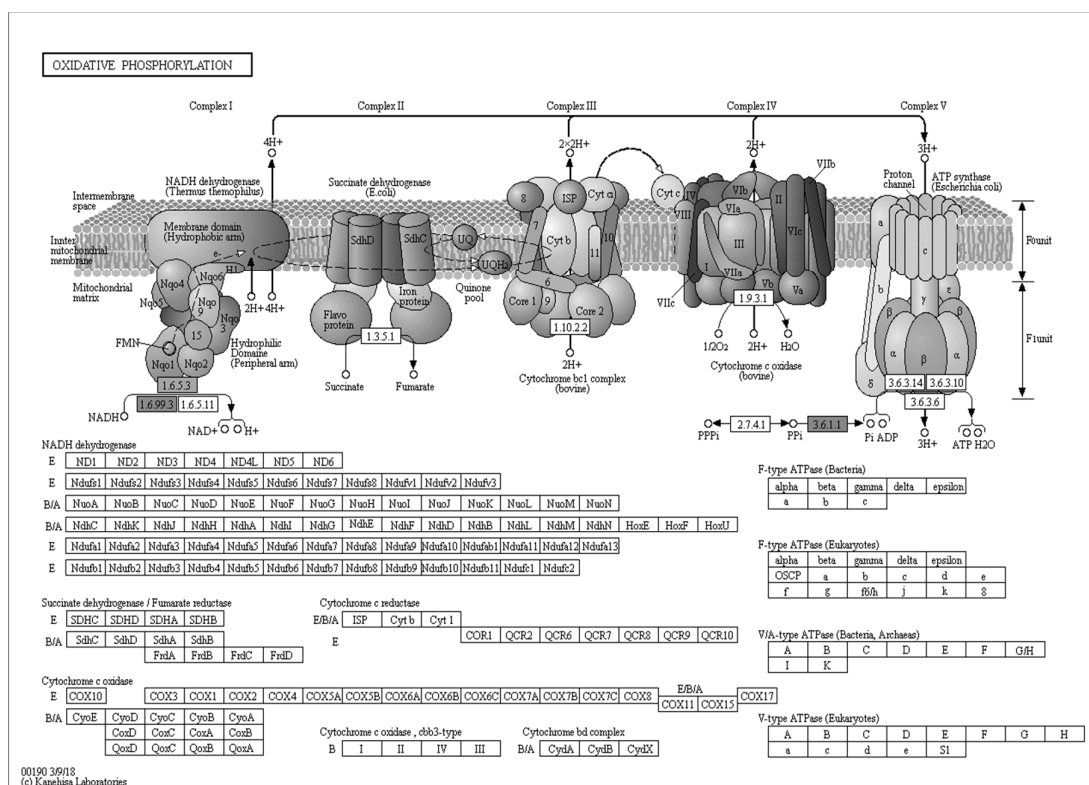
NADH در ابتدای چرخه TCA را بر عهده دارند. بخش عمده mtLPD در بافت‌های فتوسنتز کننده برای گلایسین دکربوکسیلاز (GDC) استفاده می‌شود که برای تبدیل گلایسین به سرین در شرایط تنفس نوری شدید، ضروری است. افزایش بیان mtLPD سبب افزایش فتوسنتز و به دنبال آن، تجمع زیست‌توده در *Arabidopsis* شده است (Timm et al. 2015). بر اساس تحقیقات انجام شده، فعالیت زیاد GDC سبب کاهش جریان کربن از طریق مسیر تنفس نوری می‌شود که سبب کاهش تجمع متابولیت‌های مربوط به تنفس نوری، به ویژه متابولیت‌های مختل کننده فعالیت رویسکو و افزایش احتمالی فعالیت سایر آنزیم‌های چرخه کلونین-بنسون می‌گردد. افزایش تثبیت CO_2 سبب افزایش کارایی به دام انداختن نور می‌گردد (Timm et al. 2015). به طور کلی این آنزیم با افزایش کارایی استفاده از نور سبب افزایش تثبیت CO_2 می‌شود که موجب افزایش تنفس نوری و در مجموع سبب افزایش فتوسنتز خالص می‌گردد. تنظیم متابولیسم کربن صورت گرفته نیز باعث افزایش رشد در لاین‌هایی می‌شود که افزایش بیان mtLPD در آن‌ها رخ داده باشد (Timm et al. 2015). آنالیز KEGG نتایج به دست آمده از این آزمایش تغییر در محتوای توالی مربوط به رونوشت این آنزیم را در مسیرهای متابولیسم گلایسین، سرین و ترئونین و گلوکونژونز نشان داد. دو ماده برای اکسیداسیون توسط کمپلکس پیرووات دهیدروژناز، اسیدهای چرب و گلوکز هستند. کاتالیز دکربوکسیلاسیون اکسیداتیو پیرووات توسط کمپلکس پیرووات دهیدروژناز منجر به تولید استیل کوآنزیم A می‌شود و از این طریق، متابولیسم گلوکز و اسیدهای چرب با یکدیگر مرتبط هستند (Timm et al. 2015).

آنزیم آکونیتاز (ACO) که به نام آکونیتاز هیدراتاز نیز شناخته می‌شود، یک پروتین آهن-سولفور است و آنزیمی کلیدی است که در چرخه‌های TCA و گلی اکسیلات، ایزومریزاسیون سیترات به ایزوسیترات را به وسیله cis-آکونیتاز به عنوان یک واسط، کاتالیز می‌کند.

است. افزایش مشاهده شده در میزان رونوشت‌های مربوط به آنزیم پکتین استراز می‌تواند سبب افزایش هیرولیز پکتین دیواره سلولی جنین شود. فعالیت این آنزیم سبب تضعیف دیواره سلولی جنین و رشد آن و در نتیجه موجب تجمع مواد ذخیره‌ای برای جنین خواهد شد (Levesque-Tremblay et al. 2015).

رونوشت مربوط به آنزیم DNA-directed RNA polymerase (EC: 2.7.7.6) در مسیره‌های متابولیسم پورین و پیریمیدین دارای بیان متفاوت بود. همچنین ژن‌های درگیر در متابولیسم پورین که در این آزمایش دارای بیان متفاوتی بودند، شامل آدنوزین تریبوزفسفاتاز (EC: 3.6.1.3) و نوکلئوزید تریفسفاتاز (EC: 3.6.1.15) بودند.

در مسیر پنتوز و گلوکورونات، تغییر در رونوشت مربوط به پکتین استراز (EC: 3.1.1.11) مشاهده شد. این آنزیم همواره در دیواره‌های سلولی و به شکل چندین ایزوفرم وجود دارد. آنزیم پکتین استراز، کاتالیزکننده هیدرولیز گروه‌های متیل استر پکتین‌های دیواره سلولی و تشکیل متانول و پکتات است. تغییر در فعالیت این آنزیم سبب تقویت یا تضعیف دیواره‌های سلولی می‌شود. بر اساس تحقیقات انجام شده، فعالیت این آنزیم بر HIGHLY METHYL ESTERIFIED SEEDS (HMS) در دیواره‌های سلولی حاوی پکتین، عاملی موثر بر تنظیم رشد سلول است (Levesque-Tremblay et al. 2015). پیشنهاد شده است که کاهش اندازه سلول‌های جنین طبیعی جنین می‌شود، در نتیجه عدم توسعه دیواره سلولی



شکل 8- آنزیم‌های کدشده با بیان متفاوت در ترانسکریپتوم بذره‌های توت در مسیر فسفوریلاسیون اکسیداتیو
Figure 8- Common cocklebur transcriptome differentially expressed encoded enzymes involved in oxidative phosphorylation

فعالیت می‌کند (Ponnu, Wahl, and Schmid 2011). سنتز ترهالوز در دو مرحله صورت می‌گیرد. در مرحله اول ترهالوز-6-فسفات (T6P) تشکیل می‌شود و محصول مرحله دوم، قند ترهالوز است. آنزیم کاتالیزکننده مرحله دوم سنتز ترهالوز، ترهالوز-6-فسفات فسفاتاز (EC: 3.1.3.12) (TPP) است. بر اساس آنالیز KEGG، رونوشت مربوط به TPP در مسیر متابولیسم نشاسته و ساکارز یافت شد و نتایج به دست آمده از آنالیز بیان افتراقی ژن نیز کاهش سطح رونوشت‌های مربوط به این آنزیم را در بذر بزرگ درون میوه تونق خاردار در مقایسه با بذر کوچک نشان داد. بر اساس نتایج تحقیقات انجام شده، تنظیم سطح T6P برای نمو طبیعی جنین گیاه ضروری است و جنین جهش یافته‌های ترهالوز-6-فسفات سینتاز (TPS) که مسئول کاتالیز واکنش تولید T6P است، در مرحله اژدری باقی مانده، چرخه سلولی این جنین‌ها مختل شده و دیواره سلولی نازک‌تری نسبت به جنین‌های عادی دارند. علاوه بر نمو جنین، TPS بر سایر مراحل نمو نیز تاثیر دارد. القای بیان TPS بر جهش یافته‌های حذف این آنزیم، سبب تاخیر در جوانه‌زنی، کاهش سرعت نمو، اختلال در رشد ریشه و کوتاه ماندن گیاه شد (Ponnu, Wahl, and Schmid 2011). گزارش شده است که T6P ممکن است از طریق مکانیسمی که با ABA و متابولیسم قندها درگیر است بلوغ جنین و رشد رویشی را تنظیم کند (Ponnu, Wahl, and Schmid 2011). سایر رونوشت‌های مشاهده شده در مسیر متابولیسم نشاسته و ساکارز شامل 1,3-beta-glucan synthase (EC: 2.4.1.34) و 1,4-alpha-glucan branching enzyme (EC: 2.4.1.18) بودند.

مسیر بیوسنتز فنیل آلانین، تیروزین و تریپتوفان

اسید آمینه‌های آروماتیک فنیل آلانین، تیروزین و تریپتوفان از جنبه‌های مختلف، دارای اهمیت هستند. این اسید آمینه‌ها علاوه بر اینکه در متابولیسم گیاهی دارای نقش مرکزی هستند، به عنوان اسید آمینه‌های ضروری در

لیزین اسید آمینه ضروری است که تجمع آن در بذر غلات اندک است زیرا این اسید آمینه به جای تجمع در بذر، از مسیر ساکاروپین تخریب می‌شود (Arruda et al. 2000). زنجیره‌های لیزینی در گیاهان به دنبال یک سری از واکنش‌های پی‌درپی تبدیل به کارنیتین می‌شوند. یکی از آنزیم‌های درگیر در این واکنش‌ها، هیستون لیزین-N-متیل ترانسفراز (EC: 2.1.1.43) است. توالی رونوشت مربوط به این آنزیم در این آزمایش با بیان متفاوت بین دو بذر درون یک میوه گیاه تونق خاردار مشاهده شد. بر اساس آنالیز بیان متفاوت ژن‌های دو بذر، بیان کمتر ژن مربوط به این آنزیم در بذر بزرگ‌تر مشاهده شد. آنزیم هیستون لیزین-N-متیل ترانسفراز در سه مرحله متوالی بر زنجیره لیزینی تاثیر گذاشته و در هر مرحله یک گروه متیل به زنجیره می‌افزاید. پس از واکنش‌های مراحل بعد، در پایان کارنیتین تشکیل می‌شود (Bourdin, Adenier, and Perrin 2007). با وجود این که محتوای کارنیتین در بافت‌های گیاهی بسیار کمتر از بافت‌های حیوانی است و سهم آسیل کارنیتین از کل محتوای کارنیتین نیز در گیاهان بسیار کمتر از حیوانات است، ولی این ماده در متابولیسم اسیدهای چرب گیاهی نقش مهمی دارد (Bourdin, Adenier, and Perrin 2007). در گیاهچه‌های *Arabidopsis* که ذخایر لیپیدی خود را از دست داده‌اند، افزایش سطح کارنیتین نیز مشاهده شده است که احتمال ارتباط میان کارنیتین و متابولیسم لیپیدها را افزایش می‌دهد (Bourdin, Adenier, and Perrin 2007). به نظر می‌رسد که کاهش مشاهده شده در رونوشت‌های مربوط به آنزیم هیستون لیزین-N-متیل ترانسفراز، عاملی در کاهش سنتز کارنیتین بوده و در مجموع ذخایر لیپیدی در بذر بزرگ‌تر درون میوه تونق، بیش از بذر کوچک‌تر می‌باشد. از سوی دیگر با توجه به افزایش فعالیت آنزیم‌های درگیر در متابولیسم هیدرات‌کربن، به نظر می‌رسد که منبع اصلی برای تامین انرژی جهت فعالیت‌های بذر بزرگ‌تر، در دوره نمو، کربوهیدرات‌ها باشد و نه چربی‌های درون بذر. ترهالوز یکی از قندهایی است که علاوه بر نقش متابولیک، به عنوان پیام‌رسان در ترانسسانی سیگنال نیز

چرخه TCA را تامین می‌کند. این آنزیم در فراهمی ملات مورد نیاز جهت سلول‌های محافظ روزنه و گره‌های ریشه لگوم‌ها نیز نقش مهمی دارد (Vidal and Chollet 1997). در برابر این نقش‌های بسیار گسترده، PEPC توسط یک خانواده ژنی کوچک کد می‌شود و بیان آن نیز در بافت‌های مختلف گیاهی صورت می‌گیرد (Gehrig, Heute, and Kluge 2001). مشاهده این آنزیم در مسیر متابولیسم متان در آزمایش حاضر نیز تاییدکننده نقش مهم این آنزیم در مسیرهای مختلف بیولوژیک است.

آنزیم‌های دیگر یافت شده در مسیرهای بیولوژیک بذره‌های تونق خاردار شامل 4- نیتروفنیل فسفاتاز (EC: 3.1.3.41) در مسیر تخریب آمینوبنزوات، آکونیتاز هیدراتاز (EC: 4.2.1.3) در مسیر متابولیسم گلی‌اکسیلات و دی‌کربوکسیلات، نوکلئوتید تریوزفسفات فسفاتاز (EC: 3.6.1.15) در مسیر متابولیسم تیمین، نیکوتینات فسفوریبوزیل ترانسفراز (EC: 6.3.4.21) و نیکوتینات- نوکلئوتید دی‌فسفوریلاز (EC: 2.4.2.19) در مسیر متابولیسم نیکوتینات و نیکوتینامید، استریکتوزیدین سینتاز (EC: 4.3.3.2) در مسیر بیوسنتز ایندول آلکالوید، لوسین tRNA لیگاز (EC: 6.1.1.4) در بیوسنتز مسیر آمینوآسیل tRNA و کربوکسیل استراز (EC: 3.1.1.1) در مسیر متابولیسم دارو و سایر آنزیم‌هاست.

تحقیق حاضر گام نخست در آنالیز ترانسکریپتوم بذر *Xanthium strumarium* L. است. آنالیز ترانسکریپتوم بذر این گیاه، نه تنها به دلیل ویژگی‌های منحصر به فرد آن در زمینه خواب بذر بلکه به دلیل اهمیت بسیار بالای این گیاه در تولید داروهای گیاهی بسیار ضروری است. بذر و بدنه رویشی تونق خاردار دارای متابولیسم‌های ثانویه متنوعی شامل ترکیبات فنلی، فلاونوئیدها، ساپونین‌ها، گلیکوزیدها و آلکالوئیدها است و آنالیز ترانسکریپتوم بذر آن گامی ضروری در جهت شناسایی سازوکارهای ملکولی درگیر در تولید این ترکیبات است. نتایج به دست آمده از تحقیق حاضر با استفاده از پلتفرم

آنزیم کوپینات/شیکیمات دهیدروژناز، آنزیمی است که هم از L-کوپینات و هم از شیکیمات به عنوان سوبسترا استفاده می‌کند. همچنین این آنزیم هم از NAD^+ و هم از $NADP^+$ به عنوان گیرنده استفاده می‌کند و از این نظر با شیکیمات دهیدروژناز (EC: 1.1.1.25) تفاوت دارد. در صورت استفاده از کوپینات یا شیکیمات به عنوان سوبسترا، محصول نهایی نیز به ترتیب، 3-دهیدروکوپینات یا 3-دهیدروشیکیمات خواهد بود. از سوی دیگر آنزیم شیکیمات دهیدروژناز فقط از شیکیمات به عنوان سوبسترا و $NADP^+$ به عنوان گیرنده استفاده می‌کند. 3-دهیدروکوپینات دهیدراتاز نیز تبدیل 3-دهیدروکوپینات به 3-دهیدروشیکیمات با جداسازی یک ملکول آب را کاتالیز می‌کند. واکنش تبدیل پری‌فئات به فنیل‌پیرووات با آزادسازی آب و دی‌اکسیدکربن نیز توسط پری‌فئات دهیدراتاز کاتالیز می‌شود. آرگونات دهیدراتاز نیز کاتالیز L-آرگونات به L-فنیل‌آلانین را انجام می‌دهد.

در مسیر متابولیسم پیرووات، علاوه بر آنزیم دی-هیدرولیپویل دهیدروژناز (EC: 1.8.1.4) که در مورد فعالیت آن توضیح داده شد، آنزیم فسفوانول پیرووات کربوکسیلاز (EC: 4.1.1.31) یا PEPC نیز مشاهده شد. این آنزیم کاتالیزکننده واکنش بازگشت‌ناپذیر بتاکربوکسیلاسیون فسفوانول پیرووات (PEP) در حضور HCO_3^- و Mg^{2+} است و محصول این واکنش نیز اگر الواستات (OAA) و فسفات غیرآلی (Pi) است (Izui et al. 2004). در تمام موجودات فتوسنتزکننده، شامل گیاهان آوندی، جلبک‌ها، سیانوباکترها و باکتری‌های فتوسنتزکننده و همچنین باکتری‌های غیرفتوسنتزکننده و پروتوزوآ گسترده هستند (Izui et al. 2004). در گیاهان آوندی، واکنش کاتالیز شده توسط PEPC مرحله تثبیت ابتدایی کربن در متابولیسم C_4 و CAM است. در اغلب بافت‌های غیر فتوسنتزی و در برگ‌های گیاهان C_3 ، عملکرد اولیه PEPC به صورت آناپلروتیک¹ است و مواد میانی در

¹ anaplerotic

بذر گیاهان وحشی به طور اعم و در مورد توق خاردار به طور اخص، این تحقیق رونوشت‌های زیادی را معرفی کرده است که می‌توانند در تحقیقات آتی جهت شناسایی مسیرهای بیولوژیک مربوط به نمو بذر، تولید داروها و شناسایی خواب بذر مورد استفاده قرار گیرند.

illumine HiSeq اطلاعات ارزشمندی از رونوشت‌ها و پروتئین‌های دخیل در فرایندها، عملکردها و مسیرهای متابولیک را در اختیار ما قرار داده که می‌توانند در زمینه تولید داروهای گیاه‌پایه و شناسایی سازوکارهای درگیر در خواب بذر در تحقیقات بعدی استفاده قرار شوند. با توجه به فقدان تحقیقات در زمینه آنالیز ترانسکریپتوم در مورد

Reference

منابع

- Arruda, P., E.L. Kemper, F. Papes, and A. Leite. 2000.** Regulation of lysine catabolism in higher plants. *Trends Plant Sci.* 5 (8): 324-330. [https://doi.org/10.1016/S1360-1385\(00\)01688-5](https://doi.org/10.1016/S1360-1385(00)01688-5)
- Autran, D., C. Baroux, M.T. Raissig, T. Lenormand, M. Wittig, S. Grob, A. Steimer, M. Barann, U.C. Klostermeier, O. Leblanc, J.P. Vielle-Calzada, P. Rosenstiel, D. Grimanelli, and U. Grossniklaus. 2011.** Maternal epigenetic pathways control parental contributionstons to Arabidopsis early embryogenesis. *Cell.* 145: 707-719. <https://doi.org/10.1016/j.cell.2011.04.014>
- Belmonte, M.F., R.C. Kirkbride, S.L. Stone, J.M. Pelletier, A.Q. Bui, E.C. Yeung, M. Hashimoto, J. Fei, C.M. Harada, M.D. Munoz, B.H. Le, G.N. Drews, S.M. Brady, R.B. Goldberg, and J.J. Harada. 2013.** Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. *PNAS.* 110 (5): 435-444. <https://doi.org/10.1073/pnas.1222061110>
- Bourdin, B., H. Adenier, and Y. Perrin. 2007.** Carnitine is associated with fatty acid metabolism in plants. *Plant Physiol. Biochem.* 45: 926-931. <https://doi.org/10.1016/j.plaphy.2007.09.009>
- Chen, H., F.W. Wang, Y.Y. Dong, N. Wang, Y.P. Sun, X.Y. Li, L. Liu, X.D. Fan, H.L. Yin, Y.Y. Jing, X.Y. Zhang, Y.L. Li, G. Chen, and H. Li. 2012.** Sequence mining and transcript profiling to explore differentially expressed genes associated with lipid biosynthesis during soybean seed development. *BMC Plant Biol.* 12: 122. <https://doi.org/10.1186/1471-2229-12-122>
- Curtin, S.J., F. Zhang, J. Sander, W.J. Haun, C. Starker, N.J. Baltus, D. Reyon, E.J. Dahiborg, M.J. Goodwin, A.P. Coffman, D. Dobbs, J.K. Joung, D.F. Voytas, and R.M. Stupar. 2011.** Targeted mutagenesis of duplicated genes in soybean with zinc-finger nucleases. *Plant Physiol.* 2(156): 466-473.
- Gehrig, H., V. Heute, and M. Kluge. 2001.** New partial sequences of phosphoenolpyruvate carboxylase as molecular phylogenetic markers. *Mol. Phylogenet. Evol.* 20 (2): 262-274. <https://doi.org/10.1006/mpev.2001.0973>
- Grabherr, M.G., B.J. Haas, M. Yassour, J.Z. Levin, D.A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B.W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman, and A. Regev. 2011.** Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotech.* 29 (7): 644-652. <https://doi.org/10.1038/nbt.1883>
- Izui, K., H. Matsumura, T. Furumoto, and Y. Kai. 2004.** PHOSPHOENOLPYRUVATE CARBOXYLASE: A new era of structural biology. *Annu. Rev. Plant Biol.* 55: 69-84. <https://doi.org/10.1146/annurev.arplant.55.031903.141619>
- Kanehisa, M., Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe. 2015.** KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44: 457-462. <https://doi.org/10.1093/nar/gkv1070>
- Le, B., A. Javier, A. Wagmaister, T. Kavashima, and A. Bui. 2007.** Using Genomics to Study Legume Seed Development. *Plant Physiol.* 144: 562-574. <https://doi.org/10.1104/pp.107.100362>
- Levesque-Tremblay, G., K. Müller, S.D. Mansfield, and G.W. Haughn. 2015.** HIGHLY METHYL ESTERIFIED SEEDS is a pectin methyl esterase involved in embryo development. *Plant Physiol.* 167: 725-737.

- Li, B., and C.N. Dewey. 2011.** RSEM : accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 12: 323. <https://doi.org/10.1186/1471-2105-12-323>
- Ponnu, J., V. Wahl, and M. Schmid. 2011.** Trehalose-6-phosphate: connecting plant metabolism and development. *Front. Plant Sci.* 2: 70. <https://doi.org/10.3389/fpls.2011.00070>
- Robinson, M.D., D.J. McCarthy, and G.K. Smyth. 2010.** edgeR : a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 26 (1): 139-140. <https://doi.org/10.1093/bioinformatics/btp616>
- Severin, A.J., J.L. Woody, Y.T. Bolon, B. Joseph, B.W. Diers, A.D. Farmer, G.J. Muehlbauer, R.T. Nelson, D. Grant, J.E. Specht, M.A. Graham, S.B. Cannon, G.D. May, C.P. Vance, R.C. Shomaker. 2010.** RNA-Seq Atlas of Glycine max: a guide to the soybean transcriptome. *BMC Plant Biol.* 10: 160. <https://doi.org/10.1186/1471-2229-10-160>
- Timm, S., M. Wittmi, S. Gamlien, R. Ewald, A. Florian, M. Frank, M. Wirtz, R. Hell, A.R. Ferinie, and H. Bauwea. 2015.** Mitochondrial dihydrolipoyl dehydrogenase activity shapes photosynthesis and photorespiration of *Arabidopsis thaliana*. *Plant Cell*. <https://doi.org/10.1105/tpc.15.00105>
- Troncoso-Ponce, M.A., A. Kilaru, X. Cao, T.P. Durrett, J. Fan, J.K. Jensen, ... and J.B. Ohlrogge. 2011.** Comparative deep transcriptional profiling of four developing oilseeds. *Plant J.* 68: 1014-1027. <https://doi.org/10.1111/j.1365-313X.2011.04751.x>
- Tzin, V., and G. Gad. 2010.** New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. *Mol. Plant.* 3(6): 956-972. <https://doi.org/10.1093/mp/ssq048>
- Vidal, J., and R. Chollet. 1997.** Regulatory phosphorylation of C4 PEP carboxylase. *Trends Plant Sci.* 2(6): 230-237. [https://doi.org/10.1016/S1360-1385\(97\)01046-7](https://doi.org/10.1016/S1360-1385(97)01046-7)
- Vogt, T. 2010.** Phenylpropanoid biosynthesis. *Mol Plant.* 3(1): 2-20. <https://doi.org/10.1093/mp/ssp106>

