Crop Biotech.
Spring (2019) 25:75-87. DOI: 10.30473/CB.2019.42249.1747

مجله علمی ـ پژوهشی زیست‌فناوری گیاهان زراعی
سال هشتم، شماره بیست‌وپنجم، بهار ۱۳۹۸ (۸۷-۷۵)

شناسایی و بررسی خوشه‌ژنی پارالوگ *DEVIL* در گیاه آلوروپوس لیتورالیس به روش ژنومیکس مقایسه‌ای

سیدحمیدرضا هاشمی‌پطرودی[۱*]، قربانعلی نعمت‌زاده[۲]، مارکوس کولمن[۳]

۱. استادیار، گروه مهندسی ژنتیک و بیولوژی، پژوهشکده ژنتیک و زیست‌فناوری کشاورزی طبرستان، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ایران.
۲. استاد، گروه مهندسی ژنتیک و بیولوژی، پژوهشکده ژنتیک و زیست‌فناوری کشاورزی طبرستان، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ایران.
۳. استادیار گروه ژنتیک مولکولی، موسسه ژنتیک گیاهی و تحقیقات گیاهان زراعی لیبنیز (IPK)، آلمان
(تاریخ دریافت: ۱۳۹۷/۷/۱۸ - تاریخ تصویب: ۱۳۹۸/۲/۹)

Identification and analysis of a DEVIL paralog gene cluster in *Aeluropus littoralis* by a comparative genomic approach

Seyyed Hamidreza Hashemi-petroudi[1*], Ghorbanali nematzadeh[2], Markus Kuhlmann[3]
1. Assistant Professor, Department of Genetic Engineering and Biology, Genetics and Agricultural Biotechnology Institute of Tabarestan (GABIT), Sari Agricultural Sciences and Natural Resources University (SANRU), Iran.
2. Professor, Department of Genetic Engineering and Biology, Genetics and Agricultural Biotechnology Institute of Tabarestan (GABIT), Sari Agricultural Sciences and Natural Resources University (SANRU), Iran.
3. RG Heterosis, Department Molecular Genetics, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Germany.

(Received: Oct. 10, 2018 - Accepted: Apr. 29, 2019)

Abstract

Genome-wide identification of orthologs and paralogs gene clusters across different species is considered as a common strategy for predicting gene function. Regarding to importance role of species-specific paralog genes in adaptation to specific environmental stresses, identification of paralog genes in the *Aeluropus littoralis*, halophyte plant, was considered in this study. For this purpose, the proteome data of four species including *A. littoralis*, *Oryza sativa*, *Brachypodium distachyon* and *Sorghum bicolor* was compared genome-widely. Based on OrthoMCL analysis, by comparing of 15916 protein sequences of *A. littoralis* to proteome of other species, 10312 orthologs gene cluster were identified that shared in all given species while 70 unique paralog gene clusters were devoted to *A. littoralis.* Gene ontology annotation of these paralog clusters showed that they are involved in key biological processes such as cellular processes, metabolic DNA processes, chromatin organization, response to environmental stimuli and cell growth and cycle. The study of the largest cluster of this set led to the identification of a family of small polypeptides (72-39 aa) that is called DEVIL (DVL). Analysis of *A. littoralis* transcriptome data in a Heatmap display a divergence in gene expression patterns of *DVL* gene family that could be an evident for their sub- functionalization in biological processes and molecular functions of the cell. Functional analysis of *AlDVL* peptide hormones (phytohormones) could be useful for identifying their potential role in the mechanisms involved in drought and salinity tolerance.

Keywords: *Aeluropus littoralis*, halophytes, small polypeptides, phytohormone, *DEVIL*.

چکیده

یکی از راهبردهای متداول جهت پیش‌بینی عملکرد ژن‌ها، تجزیه و تحلیل گستره ژنومی و شناسایی خوشه‌های ارتولوگ و پارالوگ در گونه‌های مختلف می‌باشد. با توجه به اهمیت ژن‌های پارالوگ در سازگاری یک گونه به شرایط خاص محیطی، در این تحقیق شناسایی ژن‌های پارالوگ در گیاه هالوفیت آلوروپوس لیتورالیس مدنظر قرار گرفت. بدین منظور پروتئوم این گیاه با برخی از گیاهان تیره گندمیان نظیر برنج، چمن‌جاوری و سورگوم در گستره ژنومی مقایسه شد. بر مبنای آنالیز OrthoMCL، بررسی مقایسه‌ای تعداد ۱۵۹۱۶ ژن آلوروپوس با توالی پروتئوم دیگر گیاهان، منجر به شناسایی بیش از ۱۰۳۱۲ گروه ژنی اورتولوگ گردید که در همه گونه‌های مورد بررسی مشترک بوده، در حالیکه۷۰ گروه ژنی پارالوگ به گیاه آلوروپوس اختصاص پیدا نمود. مستندسازی این خوشه‌ها بر مبنای ژنانتولوژی حاکی از کارکرد آنها در مسیرهای مهم زیستی نظیر فرایندهای سلولی، فرایندهای متابولیکی DNA، سازماندهی کروماتین، پاسخ به محرک‌های محیطی، رشد و چرخه سلولی بوده است. بررسی بزرگترین گروه این مجموعه، منجر به شناسایی خانواده‌ای از پلی‌پپتیدهای کوچک با اندازه ۷۲-۳۹ اسید آمینه به نام DEVIL (DVL) گردید. بررسی موتیف‌های این گروه ژنی نشان داد ۳ موتیف مختلف با مجموع ۱۰ جایگاه در این گروه پروتئینی وجود دارد. نمودار Heatmap بر مبنای آنالیز ترانسکریپتوم نشان داد، الگوی متنوعی از بیان ژن در خانواده ژنی *DVL* وجود داشته که گواهی بر بازآرایی وظایف این ژن‌ها در فرایندهای زیستی و کارکردهای مولکولی سلول می‌باشد. بررسی عملکردی پپتیدهای فیتوهورمونی *AlDVL* در مطالعات آتی می‌تواند به شناسایی نقش احتمالی آنها در مکانیسم‌های دخیل در تحمل به تنش خشکی و شوری سودمند باشد.

واژه‌های کلیدی: آلوروپوس لیتورالیس، هالوفیت، پلی‌پپتیدهای کوچک، فیتوهورمون، *DEVIL*.

* نویسنده مسئول: سیدحمیدرضا هاشمی‌پطرودی
E-mail: shr.hashemi@sanru.ac.ir

مقدمه

تغییرات آب و هوایی از یک سو، و تنش‌های غیرزنده نظیر شوری، خشکی و حرارت بالا از عوامل اصلی محدودکننده تولید محصولات کشاورزی در سراسر جهان می‌باشد (Mittler and Blumwald, 2010). هرچند که امروزه در کشاورزی مدرن تولید محصولات تراریخته مقاوم به این تنش‌ها یکی از راهبردهای اصلی افزایش تولیدات گیاهی لحاظ می‌گردد ولی ماهیت ژنتیکی و پیچیدگی‌های محیطی، اصلاح و معرفی ارقام مقاوم به این تنش‌ها را با کندی مواجه نموده است. در حال حاضر اصلاح و معرفی ارقام تجاری با عملکرد بالا با تمرکز بر بهبود تحمل به تنش شوری و خشکی یکی از راهبردهای مهم در کشاورزی نوین قلمداد می‌شود (Hashemipetroudi *et al.*, 2016). از سوی دیگر با گسترش تکنولوژی مهندسی ژنتیک امکان دست‌ورزی صفات کمی یعنی تولید گیاهان متحمل به تنش شوری و خشکی مهیا گردید. به‌طوری‌که جداسازی، انتقال و بررسی بیش‌بیان بسیاری از ژن‌ها در گیاهان زراعی مؤید این مطلب می‌باشد (Sreenvivasula, 2004). یکی از دغدغه‌های عمده که موفقیت این راهبرد را با چالش جدی مواجه می‌نماید بیان ژن‌هایی با منشا خارجی (از گیاهان دیگر) و زمینه ژنتیکی متفاوت از گیاه اصلی می‌باشد. از این منظر استفاده از گیاهان وحشی حاوی ژن‌های مقاومت که از قرابت بالایی با گیاهان زراعی برخوردارند بسیار مورد توجه می‌باشد (Faraji *et al.*, 2017).

گیاه هالوفیت تک‌لپه آلوروپوس لیتورالیس (*Aeluropus littoralis*) بومی مناطق کویری ایران و بسیاری از مناطق بیابانی کشورهای آسیایی و آفریقایی بوده، که به نظر می‌رسد با توجه به خاستگاه اکولوژیک خود علاوه بر تحمل به تنش شوری به طیف وسیعی از شرایط سخت محیطی متحمل می‌باشد (Sahar Faraji *et al.*, 2017; Hashemi-petroudi *et al.*, 2018). خصوصیات منحصربفرد فیزیولوژیکی و آناتومی این گیاه، قابلیت تحمل شرایط سخت محیطی را برای آن امکان‌پذیر ساخته است (Faraji *et al.*, 2018). از این‌رو بعلت دارا بودن طیف وسیعی از ژن‌های مقاومت به تنش‌های محیطی، می‌توان آن را بعنوان یک منبع ژنتیکی حاوی ژن‌های متحمل به تنش غیرزیستی لحاظ نموده (Hashemi *et al.*, 2016) و از سوی دیگر تعلق این گیاه به تیره غلات، به دلیل قرابت و خویشاندی، امکان بکارگیری ژن‌های مقاومت با حداقل تداخل ژنتیکی را فراهم می‌نماید (Mohsenzadeh *et al.*, 2006).

ظهور تکنولوژی‌های تعیین‌توالی ژنومی (توالی‌یابی شات‌گان) در اواخر قرن بیستم که با ارائه نسخه اولیه توالی ژنومی گیاه آرابیدوپسیس تالیانا (Meinke *et al.*, 1998; Cao *et al.*, 2011) و برنج (Goff *et al.*, 2002) همراه بود تسهیل فرایند شناسایی و تعیین‌مشخصه‌سازی ژن‌ها را موجب گردید. به‌طور کلی هدف از تعیین‌توالی ژنوم گیاهی، کشف تمامی ژن‌های موجود در گیاه و شناسایی تمامی عناصر تنظیمی مرتبط با ژن‌هاست که فرصت بسیار ارزشمندی در اختیار محققین علوم بیولوژی قرار خواهد داد و کمک شایان توجهی در درک چگونگی تکون خانواده‌های ژنی[1] تکثیر و انشعاب آنها و همچنین بروز عملکرد و خصوصیات جدید زیستی خواهد نمود. از این‌رو تعیین‌توالی کامل ژنوم[2] (WGS) گیاه هالوفیت *A. littoralis* با ارائه ۱۵۹۱۶ مدل ژن پیش‌بینی‌شده (داده‌های منتشرنشده) فرصت

1. Gene family
2. Whole-genome sequencing

بی‌نظیری را جهت تجزیه و تحلیل گسترده ژنومی (Genome-wide) این گیاه برای شناسایی ژن‌های درگیر در تحمل به تنش شوری و خشکی فراهم می‌نماید.

بررسی روابط خویشاوندی برای درک فرآیند تکامل در گیاهان و چگونگی پیدایش ژن‌ها و عملکرد زیستی آنها مفید می‌باشد. نظم بین ژن‌ها در گیاهان (غیر از فواصل فیزیکی) حفاظت شده بوده به‌طوریکه مقایسه و بررسی هم‌ردیفی ژن‌ها راه را برای کشف ژن‌ها یا شناسایی عملکرد آنها هموار می‌سازد (Proost *et al.*, 2009). به عبارت دیگر چنان‌چه اطلاعات اندکی از گیاه هدف در دسترس باشد با مقایسه توالی آن با دیگر گیاهان مدل (نظیر آرابیدوپسیس تالیانا و برنج که اطلاعات زیادی درباره ژنوم، ترانسکریپتوم، پروتئوم و تجزیه عملکردی آنها در دسترس می‌باشد)، شناسایی و تعیین مشخصه‌سازی ژن‌های همولوگوس[1] در گیاه هدف تسهیل می‌گردد (Lyons and Freeling, 2008; Emamjomeh *et al.*, 2015). ژن‌های همولوگ را بسته به رابطه بین ژن‌ها در گونه‌های مختلف می‌توان به دو گروه مختلف، ارتولوگ یا پارالوگ[2] دسته‌بندی نمود (Fitch, 1970). ارتولوگ یا ژن‌های ارتولوگوس[3] ژن‌های همولوگی بوده که از یک ژن مشترک اجدادی منشا گرفته و معمولاً در طول تکامل از عملکرد مشابه‌ای برخوردارند. در مقابل، پارالوگ‌ها، ژن‌های همولوگی بوده که از فرایند تکثیر ژن ایجاد شده و اغلب با بازآرایی عملکردی همراه بوده که برای سازگاری با شرایط خاص محیطی بسیار ضروری می‌باشند (Peters *et al.*, 2012). مضاعف‌شدگی برخی از ژن‌ها قبل از گونه‌زایی صورت گرفته و در جد مشترک به‌صورت یک جفت وجود داشته‌اند. این ژن‌ها به دلیل اینکه به‌صورت یک جفت در ژنوم اجدادی حضور داشته و در خلال فرایند گونه‌زایی نیز به صورت کو-اورتولوگ[4] به نسل‌های بعد منتقل گردیدند ژن‌های "خارج از پارالوگ[5]" نامیده می شوند، از سوی دیگر ژن‌هایی که فرایند مضاعف‌شدگی آنها بعد از فرایند گونه‌زایی صورت گرفته، ژن‌های درون-رپارالوگ[6]" نامیده می‌شوند (Koonin, 2005; Roth *et al.*, 2007). با توجه به اهمیت ژن‌های درون-رپارالوگ در سازگاری گونه به شرایط خاص محیطی، شناسایی ژن‌های پارالوگ در گیاه آلوروپوس لیتورالیس در این تحقیق مدنظر قرار گرفت. بدین منظور بررسی ژنوم این گیاه با دیگر گیاهان تیره گندمیان در گستره ژنوم مقایسه شده، و در ادامه با شناسایی خوشه ژنی پارالوگ، ساختار، همولوژی و الگوی بیانی ژن‌های آن مورد بررسی قرار گرفت.

مواد و روش‌ها

در این تحقیق از داده‌های ژنوم *Aeluropus littoralis* که اخیراً به روش تعیین‌توالی کامل ژنوم (WGS) شناسایی گردیده بود استفاده شد (داده منتشرنشده). ۱۲۵ میلیون جفت خوانش صورت گرفته در سیستم HiSeq 2500 شرکت Illumina، پوشش ژنومی ۶۲x را تامین نموده که امکان استفاده از اولین نسخه ژنوم رفرنس گونه *A. littoralis* را میسر نموده است. در این تحقیق از ۱۵،۹۱۶ مدل ژن پیش بینی شده در این گیاه جهت تجزیه و تحلیل گسترده ژنومی[8] و شناسایی خوشه‌های ارتولوگ در

1. Homologous genes
2. Orthologs or paralogs
3. Orthologs or orthologous genes
4. Co-orthologs
5. Out-paralogs
6. In-paralogs
7. Whole-genome sequencing
8. Genome-wide

گونه‌های مختلف غلات شامل برنج (*Oryza sativa*)، چمن‌جاوری جنگلی (*Brachypodium distachyon*) و سورگوم (*Sorghum bicolor*) استفاده شد. تعداد توالی پروتئین استفاده شده برای گیاهان برنج، چمن‌جاوری جنگلی و سورگوم به ترتیب ۴۰۷۴۵، ۲۶۵۵۲ و ۳۴۴۹۶ بوده که از بخش توالی رفرنس(RefSeq) موجود در پایگاه اطلاعاتی NCBI (https://www.ncbi.nlm.nih.gov/) اخذ گردید. برای بررسی همردیفی توالی پروتئین آلوروپوس با پروتئین‌های دیگر غلات مورد بررسی از نرم‌افزار UBLAST و BLASTP استفاده شد. آنالیز مقایسه ژنومی بر اساس همولوژی توالی با مقادیر پیش فرض، به روش الگوریتمOrthoMCL برای شناسایی خوشه‌های ژنی ارتولوگ استفاده شده است (Li *et al.*, 2003). به منظور مستندسازی[1] و بررسی عملکرد مولکولی هر ژن و نقش آن در فرایند زیستی و همچنین محل قرارگیری آنها در سلول، از پایگاه اطلاعاتی Gene Ontology (http.//www.geneontology.org) استفاده شد. بدین ترتیب هر ژن بر اساس عملکرد، نحوه فعالیت و مکان سلولی خود به گروه خاصی اختصاص یافت.

برای بررسی روابط فیلوژنی پروتئین‌های موجود در گروه‌های مختلف از نرم‌افزار PhyML 3.1/3.0 aLRT استفاده شد (Dereeper *et al.*, 2010). آنالیز همردیفی چندگانه[2] (MSA) با نرم‌افزار MUSCLE اجرا شد. برای شناسایی موتیف‌ها از نرم‌افزار MEME (http://meme.nbcr.net) استفاده شد (Liu *et al.*, 2004). مستندسازی موتیف‌ها و طبقه‌بندی آنها در خانواده‌های پروتئینی، پیش‌بینی دمین‌ها و جایگاه‌های مهم با نرم‌افزار InterPro (http://www.ebi.ac.uk/interpro/) صورت گرفت. جهت بررسی بیان ژن‌های پارالوگ شناسایی شده، از داده‌های RNA-seq (۳ تکرار) گیاه آلوروپوس لیتورالیس موجود در پژوهشکده ژنتیک و زیست‌فناوری کشاورزی طبرستان استفاده شد (داده منتشر نشده). اعمال تنش شوری ۶۰۰ میلی‌مولار (کلرید سدیم) بر روی گیاهان دو ماهه و به صورت تدریجی صورت گرفت (Hashemi, Nematzadeh *et al.*, 2016). نمونه‌برداری اول در زمان ۱ هفته پس از اعمال تنش شوری انجام شد. در ادامه به منظور ریکاوری، گیاهان باقی‌مانده به محیط هیدروپونیک هوگلند فاقد نمک کلرید سدیم منتقل و نمونه‌برداری دوم در مدت زمان یک هفته پس از انتقال به شرایط ریکاوری صورت پذیرفت. آنالیز تفاوت بیان ژن‌ها با استفاده از نرم‌افزار CLC (CLC Genomics workbench) Version 6.5 بررسی گردید. به این ترتیب که پس از تشکیل ماتریس بیان ژن که حاوی ژن‌ها و سطح بیان آنها می‌باشد، تغییرات بیان (FC) ژن‌های کاندید نسبت به کنترل بر مبنای Log2 FC محاسبه گردید. نمایش نمودار تغییرات بیان به‌صورت Heatmap بوده که در آن فاصله خوشه‌بندی بر مبنای همبستگی و خوشه‌بندی به روش متوسط (R package version 0.7.7) صورت گرفت.

نتایج و بحث

با توجه به تعیین‌توالی ژنوم گیاه هالوفیت آلوروپوس لیتورالیس و قرابت نزدیک آن با دیگر غلات مهم زراعی، شناسایی ژن‌های درون-پارالوگ در این گیاه که مختص این گونه باشند مدنظر قرار گرفت. بدین منظور تجزیه و تحلیل گروه‌های ژنی اورتولوگ در چهار ژنوم تیره گندمیان شامل برنج، سورگوم، چمن‌جاوری جنگلی و گیاه هالوفیت آلوروپوس

1. Annotation
2. Multiple sequence alignment

لیتورالیس انجام شد. بر مبنای آنالیز OrthoMCL، پروتئین‌های چهار گونه مورد بررسی در قالب ۲۴۰۳۷ گروه ژنی دسته‌بندی گردیدند که از این بین ۲۰۹۶۶ گروه ژنی اورتولوگ (گروه‌هایی که ژن‌های آن دستکم به دو گونه تعلق داشته‌اند)، و ۳۰۷۱ گروه ژنی پارالوگ بودند. در گیاه آلوروپوس از مجموع ۱۵۹۱۶ پروتئین مورد بررسی، ۱۳۰۷۵ پروتئین در گروه‌های مختلف دسته‌بندی شده، در حالی‌که ۲۲۳۱ پروتئین به‌صورت سینگلتون[1] مشاهده شدند که حاکی از عدم مضاعف‌شدگی این پروتئین‌ها در فرایند تکامل می‌باشد. بررسی مقایسه‌ای تعداد ۱۵۹۱۶ ژن آلوروپوس با توالی پروتئوم دیگر گیاهان مورد بررسی، منجر به شناسایی بیش از ۱۰۳۱۲ گروه ژنی اورتوگ گردید که در همه گونه‌های مورد بررسی (چهار گونه) مشترک بودند. نمودار توزیع گروه‌های ژنی مشترک اورتولوگ و یا گروه‌های گونه–اختصاصی (پارالوگ) در شکل ۱–الف نشان داده شده است. وجود ۱۰۳۱۲ ژن مشترک در چهار گونه مورد بررسی در درجه اول نشان‌دهنده تشابه عملکرد زیستی و مولکولی این ژن‌ها در هر یک از این گونه‌ها بوده که خود دلیلی بر بقای این ژن‌ها در روند تکاملی و گونه‌زایی است (Mahesh *et al.*, 2018). همانگونه که در شکل ۱–ب نشان داده شده از ۱۵۹۱۶ پروتئین مورد بررسی، ۱۳۰۷۵ گروه ژنی و ۲۲۳۱ سینگلتون شناسایی شد. تعداد پروتئین‌های پارالوگ در بین چهار گونه مورد بررسی ۳۰۷۱ بوده در حالیکه تعداد پروتئین‌های مشترک بین دو، سه و چهار گونه به ترتیب برابر ۳۱۱۰، ۷۵۴۴ و ۱۰۳۱۲ بود. ۷۰ گروه ژنی منحصربفرد پارالوگ به‌عنوان گروه‌های ژنی اختصاصی گیاه آلوروپوس لیتورالیس شناسایی گردید که احتمالاً در فرایند مضاعف‌شدگی، در این گیاه بوجود آمدند. اعتقاد بر این است تکون ژن‌های پارالوگ در یک گونه که اصولاً همراه با بازآرایی کارکرد و وظایف ژن‌ها می‌باشد به منظور سازگازی و تطابق گونه با شرایط اکولوژیکی خاص می‌باشد. این بازآرایی ناشی از انشقاق ژن، می‌تواند در جنبه‌های مختلف عملکرد ژن نظیر توالی یا ساختار پروتئین، الگوهای بیان ژن، یا میزان مشارکت در شبکه‌های مولکولی تظاهر یابد (Studer and Robinson-Rechavi, 2009).

اعتقاد بر این است میزان انشقاق ژن‌های پارالوگ در واحد زمان نسبت به ژن‌های اورتولوگ بیشتر بوده، در عین این‌که از تنوع عملکردی گسترده‌تری برخوردارند (Taylor and Raes, 2004). این دسته از گروه‌های ژنی درون– پارالوگ به‌دلیل اختصاصیت آنها به گیاه هالوفیت آلوروپوس و نقش آنها در فرایند سازگاری به تنش‌های غیر زنده شوری و خشکی بسیار حایز اهمیت می‌باشند.

مستندسازی این خوشه‌ها نشان داد بعضی از این خوشه‌های گونه–اختصاصی به‌طور بالقوه در مسیرهای زیستی مهم نظیر فرایندهای سلولی، فرایندهای متابولیکی DNA، سازماندهی کروماتین، پاسخ به محرک‌های محیطی، رشد و چرخه سلولی مشارکت دارند. در شکل ۲ طبقه‌بندی ژن‌های پارالوگ گیاه آلوروپوس لیتورالیس بر مبنای ژن انتولوژی نشان داده شده است که ژن‌ها بر مبنای فرایند زیستی (الف) و عملکرد مولکولی (ب) در آن طبقه‌بندی گردیدند. در طبقه‌بندی بر مبنای جانمایی پروتئین‌ها در سلول نیز ۲۰ درصد پروتئین‌ها در هسته، ۴۰ درصد در غشا و ۴۰ درصد باقیمانده در دیگر اجزای سلولی قرار گرفتند.

در بررسی اولین گروه از مجموع ۷۰ گروه پارالوگ آلوروپوس، ۶ پلی‌پپتید کوچک با اندازه ۷۲–۳۹ اسید آمینه شناسایی شد. بررسی همولوژی پروتئین‌ها در برنامه BLASTP 2.8.1 در پایگاه Swissprot (توالی Non-redundant UniProtKB/ SwissProt) هیچ رکورد مشابه‌ای برای این ژن‌ها مشاهده نشد در صورتی که در همردیفی در پایگاه Refseq (توالی

1. Single copy cluster

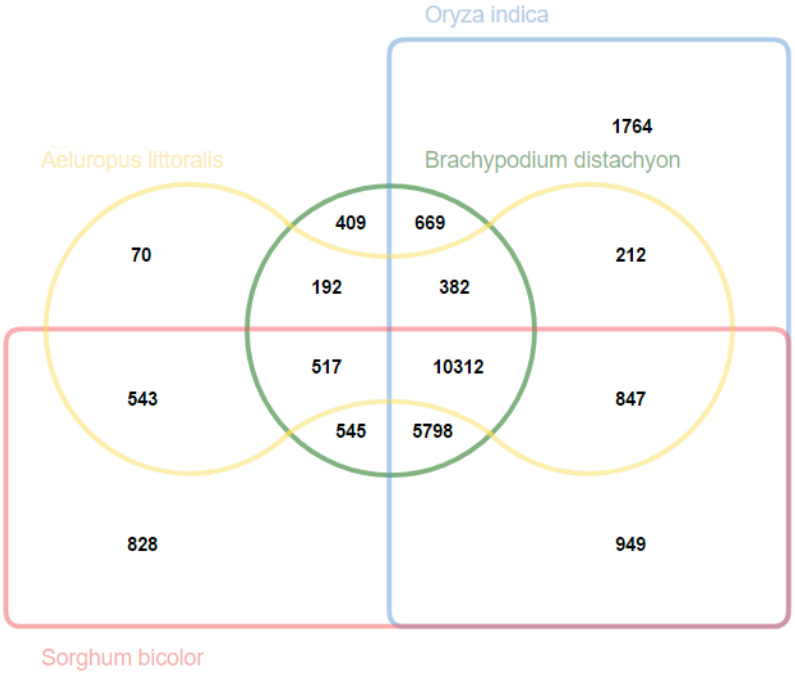
رفرنس پروتئین در NCBI)، پروتئین‌های مشابه در گیاهان *Oryza sativa* و *Panicum hallii* مشاهده گردید (جدول ۱) که از اندازه ژنی مشابه‌ای برخوردار بودند هرچند که هیچ کارکردی برای این پروتئین‌ها ارائه نشده بود[1]. روابط فیلوژنتیکی این پروتئین‌ها نیز در شکل ۳-الف نشان داده شده است. با توجه به همولوژی بالای این پروتئین‌ها نسبت به یکدیگر در مقایسه با پروتئین‌های دیگر گونه‌ها که منجر به قرارگیری آنها در یک گروه درون-پارالوگ گردیده، و با توجه به عدم حضور پروتئینی از گونه‌های دیگر می‌توان بیان نمود داشت که این پروتئین‌های گونه-اختصاصی در اثر پدیده مضاعف‌شدگی بعد از گونه‌زایی ایجاد گردیده‌اند. نمای گرافیکی آنالیز همردیفی چندگانه توالی پروتئینی نیز در شکل ۳- ب ارائه شده است. بررسی موتیف‌های این گروه ژنی به‌منظور پیش‌بینی دمین‌ها و جایگاه‌های مهم با نرم‌افزارMEME نشان داد ۳ موتیف مختلف با مجموع ۱۰ جایگاه در این گروه پروتئینی شناسایی شد (شکل ۴). موتیف شماره ۱ با E-value، 1.8e-049 در هر شش پروتئین مورد بررسی دارای یک جایگاه به طول ۲۹ اسیدآمینه بود. توزیع و پراکنش موتیف‌های شناسایی‌شده در شکل ۵- الف نشان داده شده است.

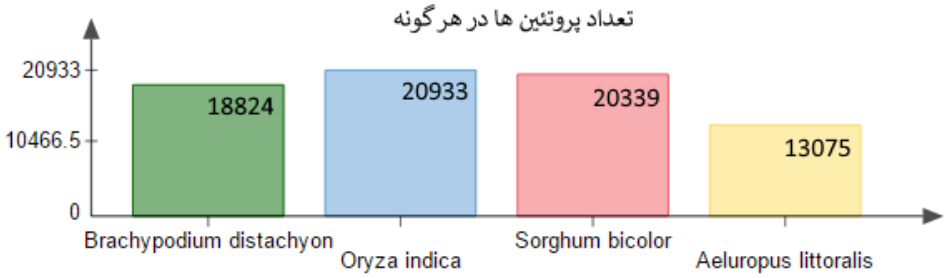
---

۱ Uncharacterized

Crop Biotech.
Spring (2019) 25:75-87. DOI: 10.30473/CB.2019.42249.1747

مجله علمی ـ پژوهشی زیست‌فناوری گیاهان زراعی
سال هشتم، شماره بیست‌وپنجم، بهار ۱۳۹۸ (۸۷-۷۵)

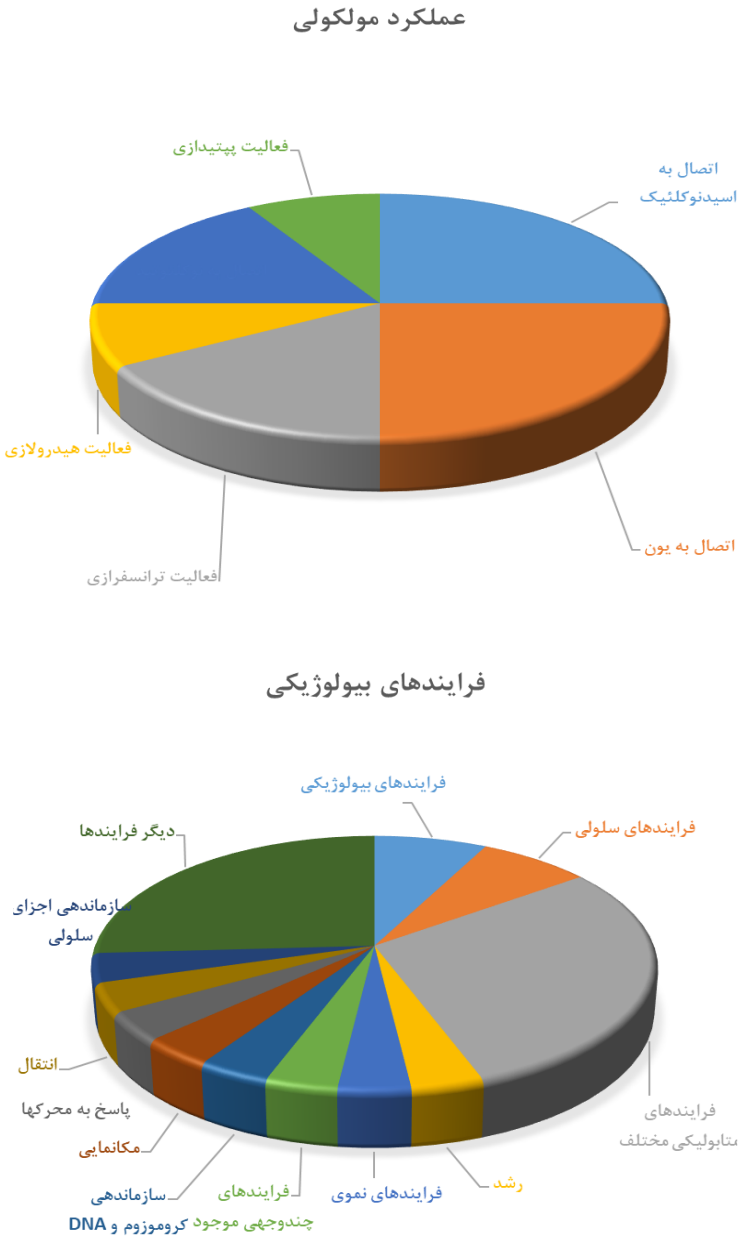




شماره‌ها بیانگر: تعداد پروتئین (۱) اختصاصی هر گونه، (۲) مشترک در دو گونه، (۳) مشترک در سه گونه و ..

شکل ۱. بررسی مقایسه‌ای ژنوم آلوروپوس لیتورالیس با دیگر گونه‌های غلات

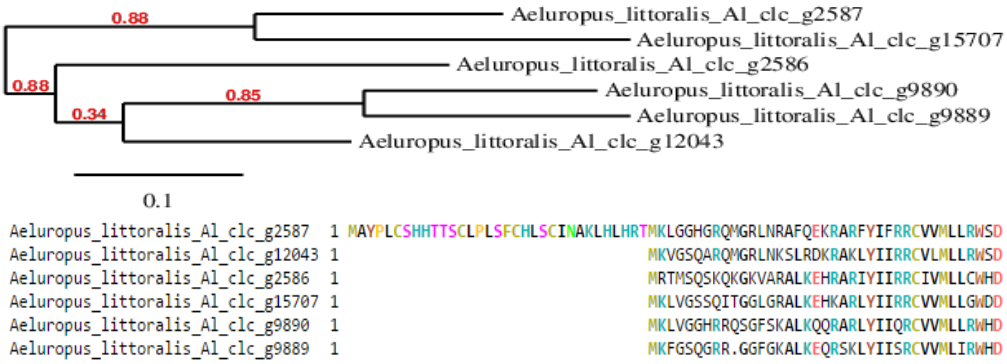
* نویسنده مسئول: سیدحمیدرضا هاشمی‌پطرودی
E-mail: shr.hashemi@sanru.ac.ir



شکل ۲. دسته‌بندی ژن انتولوژی ژن‌های درون- پارالوگ شناسایی‌شده در آلوروپوس لیتورالیس بر اساس فرایندهای زیستی (الف)، عملکرد مولکولی (ب).

جدول ۱. جدول خصوصیات همولوژی پلی‌پپتیدهای شناسایی شده در بزرگترین گروه درون- پارالوگ گیاه آلوروپوس. طول پروتئین بر مبنای طول اسید آمینه بوده و مقادیر شباهت و همپوشانی بر اساس درصد ارائه گردیده‌اند.

شماره دسترسی بانک ژن	شباهت با گیاه	شباهت	E value	همپوشانی	امتیاز کلی	طول پروتئین	کد دسترسی ژن
XP_025876235.1	*Oryza sativa*	58%	1e-18	98%	81.3	72	Al_clc_g2587
XP_025883274.1	*Oryza sativa*	82%	1e-14	100%	68.2	39	Al_clc_g12043
XP_025818286.1	*Panicum hallii*	87%	7e-16	100%	71.6	39	Al_clc_g9890
XP_025819138.1	*Panicum hallii*	90%	2e-15	100%	70.5	39	Al_clc_g15707
XP_025881936.1	*Oryza sativa*	84%	3e-15	100%	69.7	38	Al_clc_g9889
XP_025804388.1	*Panicum hallii*	95%	1e-18	100%	78.6	39	Al_clc_g2586



شکل ۳. روابط فیلوژنتیکی پروتئین‌های پارالوگ شناسایی‌شده در گیاه آلوروپوس لیتورالیس به همراه نمای گرافیکی از آنالیز همردیفی چندگانه (MSA) توالی پروتئینی

شکل ۴. شناسایی موتیف‌ها با نرم‌افزار MEME

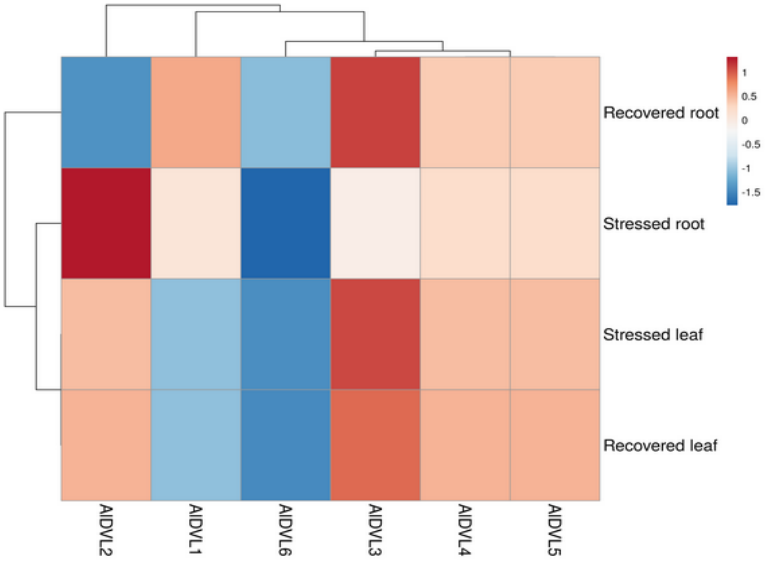
مستندسازی موتیف‌ها و تشخیص خانواده‌های پروتئینی متعلق به آن با نرم‌افزار InterPro حاکی از تعلق این موتیف به خانواده ژنی *DVL* با شماره Pfam، PF08137 و شماره IPR012552 بود درحالی‌که موتیف شماره ۲ با E-value 1.4e+001 و موتیف شماره ۳ با E-value 5.3e+002 که هر کدام با دو جایگاه در میان پروتئین‌های مورد بررسی، فاقد هیچ اطلاعاتی در پایگاه‌های بررسی دمین بودند. با توجه به تعلق این گروه ژنی به خانواده DVL، در ادامه نامگذاری ژن‌های مورد بررسی به‌صورت *AlDVL1*: Al_clc_g2587، *AlDVL2*: Al_clc_g12043، *AlDVL3*: Al_clc_g9890، *AlDVL4*: Al_clc_g15707، *AlDVL5*: Al_clc_g9889 و *AlDVL6*: Al_clc_g2586 بود. DEVILها (DVL) پلی‌پپتیدی به طول ۱۴۵–۴۰ آمینواسید بوده که اولین بار توسط Wen *et al.* (2004) شناسایی شده و انحصاراً در نهاندانگان وجود دارند (Wen et al., 2004). آنالیز همولوژی بر مبنای تشابه توالی در گیاه آرابیدوپسیس، منجر به شناسایی ۲۲ عضو خانواده ژنی *DVL* در این گیاه شده است (Jiangqi Wen and Walker, 2006). بیش‌بیان این ژن در آرابیدوپسیس، موجب پاکوتاهی، گل‌آذین

خوشه‌ای، برگ‌های روزت و غلاف‌های تیز گردید (Wen *et al.*, 2004). در مطالعه دیگری توسط Narita *et al.* (2004) در تعیین خصوصیات ژنی دیگر از این خانواده، عملکردی مشابه با ژن *DVL1* مشاهده شد که بنام ROTUNDIFOLIA4 (ROT4) نامگذاری گردیدند(Narita *et al.*, 2004; Andrews and Rothnagel, 2014). در مجموع ۲۲ همولوگ شناسایی شده از ژن‌های *DVL1* و *ROT4* در ژنوم آرابیدوپسیس، به‌عنوان خانواده ROT FOUR-LIKE/DEVIL (RTFL/DVL) نامگذاری گردیدند (Kastin, 2013). پپتیدهای RTFL/DVL، دارای یک دمین حفاظت شده متداول ۳۰ آمینواسیدی در نزدیکی منطقه C- ترمینال (RTF domain) می‌باشند (Narita, Moore *et al.*, 2004) که در این تحقیق در قالب موتیف ۱ شناسایی شده است.

بررسی منابع در خصوص بررسی بیان این خانواده در تنش‌های زیستی و یا غیرزیستی حاکی از عدم وجود تحقیقاتی در این زمینه بوده است. از این‌رو بمنظور بررسی الگوی بیان این گروه ژنی به روش RNA-seq، بیان این ژن‌ها در دو بافت برگ و ریشه و در دو شرایط تنش شوری و شرایط ریکاوری در گیاه آلوروپوس مورد بررسی قرار گرفت. ازآنجایی‌که در شرایط شوری، اثرات تنش‌های یونی در مقایسه با اثرات اسمزی بسیار دیرتر بروز نموده و از شدت تاثیر کمتری برخوردارند؛ علاوه بر این تنها در سطوح بالای تنش شوری، اثرات یونی بر اسمزی غلبه دارد (Munns and Tester, 2008). در این تحقیق اعمال تنش شوری در سطح ۶۰۰ میلی‌مولار، (برابر با ۳۵/۰۶۴ گرم بر لیتر) صورت پذیرفت. الگوهای بیانی حاصله از آنالیز RNA-seq در قالب نمودار Heatmap در شکل ۵ نشان داده شده است. بررسی آنالیز بیان ژن RNA-seq، برای دو ژن پیش‌بینی شده *AlDVL4* و *AlDVL5* نشان داد که هیچ بیانی چه در بافت‌های مورد بررسی (برگ و ریشه) و چه در تنش شوری و شرایط ریکاوری مشاهده نگردید که حاکی از خاموش بودن این ژن‌ها در بافت و تنش مذکور می‌باشد. در خصوص ژن *AlDVL2*، علی‌رغم مشاهده بیان متفاوت در بافت ریشه در مواجهه با تنش شوری و ریکاوری، هیچ بیانی در بافت برگی مشاهده نگردید که حاکی از بیان اختصاصی آن در ریشه می‌باشد. این ژن در بافت تحت تنش ریشه از افزایش بیان ۰/۸۲ بر مبنای loc2 FC نسبت به کنترل برخوردار بود در حالی‌که در شرایط یک هفته پس از ریکاوری بیان به میزان ۱/۸۲- نسبت به کنترل کاهش یافت. در خصوص ژن *AlDVL6*، الگوی تقریبا مشابه‌ای از کاهش بیان در بافت‌ها و تنش‌های مورد بررسی مشاهده شد. این کاهش بیان در سطوح مختلف بافتی و تنش، در واقع به خاموش بودن این ژن در این سطوح و فعالیت آن در شرایط کنترل دلالت داشت. افزایش بیان ژن *AlDVL3* در بافت برگی، و در هر دو تیمار تنش شوری و ریکاوری مشاهده شد در حالی‌که در بافت ریشه، این افزایش بیان تنها منحصر به شرایط ریکاوری بود. این در حالی است که در ژن *AlDVL1* تنها در بافت ریشه و شرایط ریکاوری افزایش بیان نسبت به کنترل مشاهده گردید. همان‌گونه که در نمودار Heatmap نشان داده شده، الگوی متفاوتی از بیان ژن در خانواده ژنی *DVL* مشاهده می‌گردد. این الگوی بیان متفاوت با توجه به همولوژی بالای این توالی نسبت به یکدیگر، احتمالاً از مکانیسم‌های تنظیمی متفاوت در کنترل فعالیت این ژن‌ها ناشی شده، ضمن این‌که می‌تواند گواهی بر وظایف متفاوت این ژن‌ها در فرایندهای زیستی و کارکردهای مولکولی سلول باشد.

تجزیه و تحلیل گسترده ژنومی و شناسایی خوشه‌های ارتولوگ در گونه‌های مختلف جزء مهمی از مطالعات ژنتیک مقایسه‌ای در گیاهان می‌باشد.

شناسایی خانواده‌های ژنی حفاظت‌شده در میان خوشه‌های ارتولوگ، امکان درک عملکرد و خصوصیات تکاملی پروتئین‌ها را در میان گونه‌های مختلف زراعی میسر می‌سازد.



شکل ۵. نمایش Heatmap بیان ژن‌های DVL در گیاه آلوروپوس لیتورالس

ژن‌های پارالوگ به‌دلیل این‌که در فرایند مضاعف‌شدگی عهده‌دار وظایف جدیدی در گونه انشقاق‌یافته گردیدند از کارکرد زیستی متفاوتی نسبت به ژن‌های ارتولوگ برخوردار هستند (Goffin and Ghuysen, 1998; Gabaldón and Koonin, 2013) که الگوی بیانی مشاهده شده از گروه ژنی *AlDVL* که از تنوع بالایی در بیان برخوردار بود، در همین راستا قابل توجیه می‌باشد. در بررسی خانواده ژنی *NAC* در گیاه *Gossypium raimondii* پنج ژن پارالوگ شناسایی شده، از الگوهای بیانی متفاوت برخوردار بودند که ناشی از بازآرایی عملکرد این ژن‌ها می‌باشد (Shang *et al.*, 2013). در اندک مطالعات صورت گرفته در خصوص بررسی عملکردی این پروتئین‌ها، کارکرد آنها در مراحل مختلف رشدی و نموی مورد بررسی قرار گرفت. بعنوان نمونه به بیش‌بیان شش عضو ژن *RTFL* در آرابیدوپسیس می‌توان اشاره نمود، که با ایجاد فنوتیپ گرد شدن برگ و ریشه‌های اولیه کوتاه‌تر همراه بود (Guo *et al.*, 2015). به‌طورکلی اعتقاد بر این است پپتیدهای کد شده توسط ژن‌های کدکننده کوچک، نقش مهمی در رشد گیاهان ایفا می‌کنند و به شیوه‌ای مشابه فیتوهورمون‌ها عمل می‌کنند. برخی از پپتیدهای کدشده توسط ژن‌های کوچک، در سلول‌های خاصی ترشح شده و به سلول‌های دیگر انتقال داده می‌شوند. این پپتیدهای ترشحی که در برخی منابع علمی به‌عنوان هورمون‌های پپتیدی نام‌گذاری می‌شوند (Andrews and Rothnagel, 2014; Matsubayashi, 2014)، نقش مهمی در رشد گیاهان، مشابه با نقش فیتوهورمون‌ها دارند (Matsubayashi and Sakagami, 2006).

تجزیه و تحلیل مقایسه‌ای خوشه‌های ژنی ارتولوگ و پارالوگ برای درک قوانین حاکم بر ساختار ژنوم و عملکرد ژن/ پروتئین بسیار حایز اهمیت است. اطلاعات به‌دست‌آمده از مقایسه خوشه‌های ارتولوگ و پارالوگ، پیش‌نیاز مطالعات طبقه‌بندی، تاکسونومی و مطالعات فیزیولوژیکی بوده که به نوبه خود درک جامعی از مکانیسم‌های تکامل مولکولی ژن‌ها و ژنوم ارائه می‌دهد. در این تحقیق مقایسه گستره ژنومی

پروتئین‌های گیاه آلوروپوس لیتورالیس با دیگر گیاهان تیره گندمیان برآوردی از ژن‌های اورتولوگ و پارالوگ را در این گونه هالوفیت ارائه نموده است ضمن اینکه به شناسایی گروهی از پروتئین‌های گونه-اختصاصی در گیاه هالوفیت منتج گردیده که احتمالا در فرایند استقرار این گونه در قالب فرایند مضاعف‌شدگی ژن‌ها، تکون یافتند. گروه ژنی *AlDVL* شناسایی شده در این تحقیق در زمره پپتیدهای فیتوهورمونی کوچک بوده، که بررسی عملکردی آن در مطالعات آتی به منظور شناسایی نقش احتمالی آن در تنش‌های غیرزنده خشکی و شوری می‌تواند در شناسایی مکانیسم‌های دخیل در تحمل به تنش سودمند باشد.

سپاسگزاری

این پژوهش با استفاده از اعتبارات پژوهشکده ژنتیک و زیست‌فناوری کشاورزی طبرستان، از محل طرح تحقیقاتی مصوب به شماره T215/96 انجام شده است که بدین‌وسیله قدردانی می‌گردد.

REFERENCES

Andrews SJ, Rothnagel JA (2014) Emerging evidence for functional peptides encoded by short open reading frames. Nat. Rev. Genet. 15(3): 193-204.

Cao J, Schneeberger K, Ossowski S, Gunther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C, Wang X, Ott F, Muller J, Alonso-Blanco C, Borgwardt K, Schmid KJ and Weigel D (2011) Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. Nat. Genet. 43(10): 956-963.

Dereeper A, Audic S, Claverie JM, Blanc G (2010) BLAST-EXPLORER helps you building datasets for phylogenetic analysis. BMC Evol. Biol. 10(1): 8.

Emamjomeh A, Goliaei B, Zahiri J, Ebrahimpour R (2015) Prediction of gene co-expression by quantifying heterogeneous features. Curr. Bioinform. 10(4): 414-424.

Faraji S, Hashemi-Petroudi SH, Najafi-Zarrini H, Ranjbar GA (2018) Characterization and expression profiling of AlPKL gene in response to salinity stress and recovery conditions in halophyte *Aeluropus littoralis*. Crop bitech. 7(20): 13-27.

Faraji S, Najafi-Zarrini H, Hashemi-Petroudi S, Ranjbar G (2017) AlGLY I gene implicated in salt stress response from halophyte *Aeluropus littoralis*. Russ. J. Plant Physiol. 64(6): 850-860.

Faraji S, Najafi-Zarrini H, Hashemi-Petroudi SH, Ranjbar GA (2017) Comparative expression profiling of four salt-inducible genes from *Aeluropus littoralis*. Iran. J. Genet. Plant Breed. 6(1): 1-7.

Fitch WM (1970) Distinguishing homologous from analogous proteins. Syst. Zool. 19(2): 99-113.

Gabaldón T, Koonin EV (2013) Functional and evolutionary implications of gene orthology. Nature Reviews Genetics. 14(5): 360.

Goff SA, Ricke D, Lan TH, Presting G, Wang R, Dunn M, Glazebrook J, Sessions A, Oeller P, Varma H, Hadley D, Hutchison D, Martin C, Katagiri F, Lange BM, Moughamer T, Xia Y, Budworth P, Zhong J, Miguel T, Paszkowski U, Zhang S, Colbert M, Sun WL, Chen L, Cooper B, Park S, Wood TC, Mao L, Quail P, Wing R, Dean R, Yu Y, Zharkikh A, Shen R, Sahasrabudhe S, Thomas A, Cannings R, Gutin A, Pruss D, Reid J, Tavtigian S, Mitchell J, Eldredge G, Scholl T, Miller RM, Bhatnagar S, Adey N, Rubano T, Tusneem N, Robinson R,

Feldhaus J, Macalma T, Oliphant A, Briggs S (2002) A draft sequence of the rice genome (*Oryza sativa* L. ssp. japonica). Science. 296(5565): 92-100.

Goffin C, Ghuysen JM (1998) Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs. Microbiol. Mol. Biol. Rev. 62(4): 1079-1093.

Guo P, Yoshimura A, Ishikawa N, Yamaguchi T, Guo Y, Tsukaya H (2015) Comparative analysis of the RTFL peptide family on the control of plant organogenesis. J. Plant Res. 128(3): 497-510.

Hashemi-Petroudi SH, Ghorbani H, Kuhlmann M (2018) Isolation Phosphoglycerate Dehydrogenase gene from *Aeluropus littoralis* halophyte plant and functional analysis of T-DNA mutant in *Arabidopsis thaliana*. Crop bitech. 8(23): 79-92.

Hashemi SH, Nematzadeh G, Ahmadian G, Yamchi A, Kuhlmann M (2016) Identification and validation of *Aeluropus littoralis* reference genes for Quantitative Real-Time PCR Normalization. J. Biol. Res. 3 (1): 2-18.

Hashemipetroudi S, Nematzadeh G, Ahmadian G, Yamchi A, Kuhlmann M (2016) Expression analysis of salt stress related expressed sequence tags (ESTs) from *Aeluropus littoralis* by quantitative real-time PCR. Biosci. Biotech. Res. Comm. 9(3): 445-456.

Kastin A (2013) Handbook of biologically active peptides, Academic press.

Koonin EV (2005) Orthologs, paralogs, and evolutionary genomics. Annu. Rev. Genet. 39: 309-338.

Li L, Stoeckert CJ Jr. Roos DS (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. Genome Res. 13(9): 2178-2189.

Liu FF, Tsai JJ, Chen R-M, Chen S, Shih S (2004) FMGA: finding motifs by genetic algorithm. Bioinformatics and Bioengineering, 2004. BIBE 2004. Proceedings. Fourth IEEE Symposium on, IEEE.

Lyons E, Freeling M (2008) How to usefully compare homologous plant genes and chromosomes as DNA sequences. The Plant Journal. 53(4): 661-673.

Mahesh HB, Subba P, Advani J, Shirke MD, Loganathan RM, Chandana SL, Shilpa S, Chatterjee O, Pinto SM, Prasad TSK, Gowda M (2018) Multi-Omics Driven Assembly and Annotation of the Sandalwood (*Santalum album*) Genome. Plant Physiol. 176(4): 2772-2788.

Matsubayashi Y (2014) Posttranslationally modified small-peptide signals in plants. Annu. Rev. Plant Biol. 65: 385-413.

Matsubayashi Y, Sakagami Y (2006) Peptide hormones in plants. Annu. Rev. Plant Biol. 57: 649-674.

Meinke DW, Cherry JM, Dean C, Rounsley SD, Koornneef M (1998) *Arabidopsis thaliana*: a model plant for genome analysis. Science. 282(5389): 662, 679-682.

Mittler R, Blumwald E (2010) Genetic engineering for modern agriculture: challenges and perspectives. Annu. Rev. Plant Biol. 61: 443-462.

Mohsenzadeh S, Malboobi M, Razavi K, Farrahi-Aschtiani S (2006) Physiological and molecular responses of Aeluropus lagopoides (Poaceae) to water deficit. Environ. Exp. Bot. 56(3): 314-322.

Munns R, Tester M (2008) Mechanisms of salinity tolerance. Annu. Rev. Plant Biol. 59: 651-681.

Narita NN, Moore S, Horiguchi G, Kubo M, Demura T, Fukuda H, Goodrich J, Tsukaya H (2004) Overexpression of a novel small peptide ROTUNDIFOLIA4 decreases cell proliferation and alters leaf shape in *Arabidopsis thaliana*. Plant J. 38(4):

699-713.

Peters AE, Bavishi A, Cho H, Choudhary M (2012) Evolutionary constraints and expression analysis of gene duplications in *Rhodobacter sphaeroides*. BMC Res. Notes. 5(1): 192.

Proost S, Van Bel M, Sterck L, Billiau K, Van Parys T, Van De Peer Y, Vandepoele K (2009) PLAZA: a comparative genomics resource to study gene and genome evolution in plants. Plant Cell. 21(12): 3718-3731.

Roth C, Rastogi S, Arvestad L, Dittmar K, Light S, Ekman D, Liberles DA (2007) Evolution after gene duplication: models, mechanisms, sequences, systems, and organisms. J. Exp. Zool. B Mol. Dev. Evol. 308 (1): 58-73.

Shang H, Li W, Zou C, Yuan Y (2013) Analyses of the NAC transcription factor gene family in *Gossypium raimondii* Ulbr.: chromosomal location, structure, phylogeny, and expression patterns. J. Integr. Plant Biol. 55(7): 663-676.

Studer RA, Robinson-Rechavi M (2009) How confident can we be that orthologs are similar, but paralogs differ? Trends Genet. 25(5): 210-216.

Taylor JS, Raes J (2004) Duplication and divergence: the evolution of new genes and old ideas. Annu. Rev. Genet. 38: 615-643.

Wen J, Lease KA, Walker JC (2004) DVL, a novel class of small polypeptides: overexpression alters Arabidopsis development. Plant J. 37(5): 668-677.

Wen J, Walker J (2006) DVL peptides are involved in plant development. Handbook of Biologically Active Peptides, Elsevier: 17-22.