

روابط تکاملی ساختار دمین پروتئین *ATAXIN-2* در گیاهان: بینش احتمالی دخالت این پروتئین در

بومی‌سازی ترجمه در فرایند جنین‌زایی گیاه برنج (*Oryza sativa*)

شادی حیدری^۱، پیوند حیدری^۲، بهارک حیدری^۳

چکیده

مقدمه: با توسعه فن آوری‌های تعیین توالی با توان بالا، داده‌های توالی با سرعت زیادی جمع‌آوری می‌شوند و تفسیر عملکرد آنها یک چالش اساسی است. ژن *ATXN2* کد کننده پروتئینی با عملکرد ناشناخته در برنج (*oryza sativa*) است. هدف از این مطالعه، تعیین ساختار دمین و پیش‌بینی عملکرد این پروتئین بود. **روش‌ها:** درخت فیلوژنی موتیف توالی با ۱۶۴ توالی اسید آمینه *ATAXIN-2* از ۴۵ گونه گیاهی توسط مجموعه *MEME* و نرم افزار *Phylyp* آنالیز شد. **نتایج و بحث:** نتایج، دو نوع کلاس کلی از پروتئین‌های *ATAXIN-2* را مشخص نمود. عدم وجود دمین *LsmAD* در یک کلاس ممکن است ویژگی‌های عملکردی متمایزی را نسبت به کلاس دیگر *ATAXIN-2* اعطا کند. در این پژوهش عملکرد *ATAXIN-2* به عنوان پروتئین‌های متصل شونده به *RNA (RNA binding)* پیش‌بینی شد که در تنظیم *mRNA* سیتوپلاسمی در جنین‌زایی برنج نقش دارد. یافته‌های این پژوهش می‌تواند اساس تحقیق در مورد عملکرد این پروتئین در هدف قرار دادن زیرمجموعه‌های جداگانه *mRNA* به مکان‌های زیر سلول در گیاهان باشد که توسط سلول‌های دو قطبی جنین برای ایجاد محفظه‌های عملکردی مشخص بکار گرفته می‌شود.

واژه‌های کلیدی: پروتئین‌های متصل شونده به *RNA*، پیش‌بینی عملکرد، ترجمه موضعی، فیلوژنتیک

۱ و ۲. دکتری، گروه اصلاح نباتات، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

۳. استادیار، گروه مهندسی کامپیوتر، واحد کرمانشاه، دانشگاه آزاد اسلامی، کرمانشاه، ایران (*نویسنده مسئول b_heidari@iauksh.ac.ir)

مقدمه

با توسعه فن‌آوری‌های تعیین توالی با توان بالا، داده‌های توالی‌های جدید با سرعت زیادی جمع‌آوری می‌شوند در حالیکه تفسیر عملکرد آنها یک چالش اساسی است (Wong *et al.*, 2019). بنابراین یکی از مراحل کلیدی در همه پروژه‌های تعیین توالی ژنوم، پیش‌بینی عملکرد ژن‌های موجود در آنها است. روش‌های آزمایشگاهی برای تعیین عملکرد پروتئین گامی دشوار و یک فرایند محدود کننده عمده برای تعیین عملکرد است. بسیاری از روش‌های محاسباتی و تجزیه و تحلیل‌های بیوانفورماتیک به منظور پیش‌بینی عملکرد ژن ایجاد شده است. به طور معمول، تفسیر عملکرد پروتئین به دنبال جستجوهای پایگاه داده و تکنیک‌های محاسباتی برای یافتن همولوگ‌های توالی پروتئین در آرگانیزم‌های مدل صورت می‌گیرد. این روش‌ها بر این اصل تکیه دارند که پروتئین‌های یک جد مشترک ممکن است عملکرد مشابهی داشته باشند. با این حال، بیشتر خانواده‌های پروتئینی دارای گروه‌هایی از پروتئین‌ها با عملکردهای مختلف مولکولی هستند و بنابراین رویکردهای بیوانفورماتیک سنتی قادر نیستند عملکرد مناسب را به پروتئین‌های بدون علامت اختصاص دهند (Engelhardt *et al.*, 2009). اکنون نرم افزارهای مختلفی برای بررسی دقیق توالی‌ها و ساختارهای پروتئین برای شباهت‌های تاخوردگی‌ها و موتیف‌ها در دسترس است. BLAST، Dali و PSI-BLAST نمونه‌هایی از رویکردهای کلاسیک برای استنباط عملکرد از روی توالی‌ها هستند و پایگاه‌های داده مانند PROSITE، Pfam و InterPro و از طرفی Profunc، SPRITE و PINTS به ترتیب، به منظور جستجوی عملکرد در سطح ساختاری، موتیف و الگوهای سه بعدی بکار می‌روند. از آنجا که ساختار پروتئین‌ها بیشتر از توالی‌ها حفاظت شده‌اند از این رو جستجوی شباهت در سطح ساختار معنای عمیق‌تری را نسبت به توالی فراهم می‌کند (Nadzirin *et al.*, 2012).

یافتن عملکرد باز هم می‌تواند با ادغام داده‌های عملکردی موجود با استفاده از روابط تکاملی یک خانواده پروتئینی معنای عمیق‌تری پیدا کند. این رویکرد فیلوژنتیک به منظور تفسیر عملکرد مولکولی (فیلوژنومیک) از یک درخت فیلوژنتیک صریح برای پیش‌بینی عملکرد استفاده می‌کند و یک روش موثر برای پیش‌بینی عملکرد مولکولی پروتئین است (Brown & Sjölander, 2006). این روش شامل شواهد عملکردی از همه اعضای یک خانواده است که با استفاده از یک درخت فیلوژنتیک مسیر تکاملی مجموعه‌ای از پروتئین‌های همولوگ را منعکس می‌کند و پیش‌بینی‌های قوی برای پروتئین‌های ناشناخته انجام می‌دهد. به دلیل تجدید ساختار مدولار دمین‌ها در طول تکامل پروتئین "دمین شافلینگ" بیشتر پروتئین‌ها از چندین دمین تشکیل شده‌اند که این امر روابط همسانی دوتایی بیشتری را بین توالی‌های دمین منفرد نسبت به توالی‌های پروتئین کامل به وجود می‌آورد. بهره‌برداری از این واقعیت می‌تواند باعث افزایش حساسیت رویکردهای پیش‌بینی عملکرد پروتئین شود (Chothia *et al.*, 2003).

به منظور تعیین عملکرد پروتئین *ATXN2* چندین دمین از این پروتئین و شرکای تعامل کننده متنوع آنها شناسایی و مورد پژوهش قرار گرفته‌اند. *ATAXIN-2* از دو دمین کروی به نام Lsm و LsmAD تشکیل شده است. دمین Lsm احتمالاً به RNA متصل می‌شود و دمین LsmAD شامل یک سیگنال ترانس گلژی با واسطه کلاترین و یک سیگنال خروجی شبکه آندوپلاسمی (ERD) است (Albrecht *et al.*, 2004). مشخص شده است که DEAD/H-Box RNA Helicase DDX6 به دمین Lsm/LsmAD در *ATAXIN-2* متصل می‌شود. DDX6 یک جز از اجسام P و گرانول‌های تنش (SG) است (Nonhoff *et al.*, 2007). SG ها و اجسام P محل تریاز RNA هستند. قرار گرفتن در معرض تنش سلولی می‌تواند باعث پاسخ تنش شود که شروع ترجمه را متوقف کرده و منجر به تشکیل SG ها می‌شود. SG ها کمپلکس‌های RNA-پروتئینی سیتوپلاسمی پویا هستند که حاوی پروتئین‌های اتصال‌دهنده RNA، mRNA ها و عوامل شروع ترجمه هستند. بعد از رفع تنش، SG ها جدا شده و ترجمه mRNA از سر گرفته می‌شود. همچنین mRNA های ترجمه نشده می‌توانند به اجسام P، گرانول‌های دیگری که محل تخریب mRNA هستند، هدایت شوند (Li *et al.*, 2013).

دمین *ATAXIN-2* Lsm تیپ معمول پروتئین‌های Sm هستند و در انواع رویدادهای اساسی پردازش RNA از جمله اصلاح RNA، اسپلایسینگ Pre-mRNA، تجزیه و تخریب mRNA نقش دارند. برخی از آنها همچنین اجزای مهم ریبونوکلوپروتئین‌های هسته‌ای کوچک اسپلایسوزومی (snRNP) هستند (Albrecht *et al.*, 2004). *ATAXIN-2* همچنین دارای یک موتیف C-terminal یا (PAM2) در انتهای کربوکسی است که با پروتئین متصل‌شونده به پلی (A) سیتوپلاسمی (PABP) که یکی دیگر از اجزاء SG است، تعامل می‌کند (Albrecht *et al.*, 2004). ساختار دمین پروتئین *ATAXIN-2* اطلاعات مربوط به عملکرد مولکولی آن را در متابولیسم RNA فراهم می‌کند. از جمله عملکرد آنها می‌توان به دخالت در چندین مسیر از جمله تنظیم ترجمه، اتصال، پاسخ تنش، آندوسیتوز گیرنده و متابولیسم لیپید اشاره نمود. *ATAXIN-2* از نظر تکاملی حفاظت شده است و عملکرد آن علاوه بر انسان به SG ها و تنظیم ترجمه در ارگانیسم‌های مدل نیز مرتبط است. علاوه بر این، *ATAXIN-2* غلظت درون سلولی شرکای تعامل کننده خود، پروتئین متصل‌شونده به پلی (A)، یکی از اجزاء دیگر SG ها و یک عامل اصلی برای کنترل ترجمه را تنظیم می‌کند. بنابراین، بیان ژن *ATAXIN-2* سلولی برای مونتاژ SG و اجسام P ضروری است که بخش‌های اصلی تنظیم و کنترل تخریب، پایداری و ترجمه mRNA هستند.

Jiménez-López & Guzmán *et al.*, (2014) در بررسی روابط تکاملی دمین‌ها، اطلاعات ارزشمندی را در مورد تکامل و ساختار دمین پروتئین‌های *ATAXIN-2* در سراسر یوکاریوت ها فراهم کردند. Heidari (2010) در مطالعه خود *ATAXIN-2* را به عنوان ژن کد کننده پروتئینی با عملکرد ناشناخته و افزایش بیان معنی‌دار در جنین‌زایی گیاه برنج گزارش نمود. مطالعات زیادی افزایش بیان این پروتئین را در بافت‌های جنینی ارگانیسم‌های مختلف گزارش کرده‌اند (Kiehl *et al.*, 2006; Kiehl *et al.*, 2000).

تاکنون عملکرد سلولی دقیق *ATXN2* در گیاهان، ناشناخته باقی مانده است. بنابراین در این تحقیق، از روابط تکاملی دمین‌های *ATXN2* در گونه‌های مختلف گیاهی برای پیش‌بینی عملکرد این پروتئین ناشناخته استفاده شد.

مواد و روش‌ها

به منظور بررسی دمین‌های حفاظت شده در خانواده *ATAXIN-2* در گیاهان تمام اعضای خانواده *ATAXIN-2* که در PlantGDB (<http://www.plantgdb.org/>) ثبت شده‌اند به عنوان توالی‌های جستجو (query) برای جستجو در Pfam (<https://pfam.xfam.org/>) (Finn *et al.*, 2016) و بانک اطلاعاتی دمین حفاظت شده (CDD) NCBI (Marchler-Bauer *et al.*, 2013) مورد استفاده قرار گرفتند. PlantGDB پایگاه داده‌ای از توالی‌های مولکولی برای همه گونه‌های گیاهی است. هدف این وب سایت ایجاد زمینه‌ای برای شناسایی مجموعه‌ای از ژن‌های مشترک در همه گیاهان یا گونه‌های خاص از طریق ادغام تعدادی از ابزارهای بیوانفورماتیک است که پیش‌بینی ژن و مقایسه بین گونه‌ها را تسهیل می‌کند. اسکن Pfam با استفاده از hmmscan، از مجموعه نرم‌افزار HMMer (Eddy, 2011) با استفاده از آستانه گرد آوری اجرا شد. $E\text{-value} < 10^{-5}$ ، با حداقل ۷۰٪ هم‌پوشانی توالی های دمین به منظور یافتن هم‌ردیفی معنی‌دار در نظر گرفته شد. اسکن CDD با استفاده از rpsblast از مجموعه بلاست NCBI انجام شد (Camacho *et al.*, 2009). گزینه‌های اجرای rpsblast آستانه $E\text{-value} < 10^{-5}$ ، آمار مبتنی بر ترکیب و پوشاندن نرم قطعات حاوی اطلاعات کم بود.

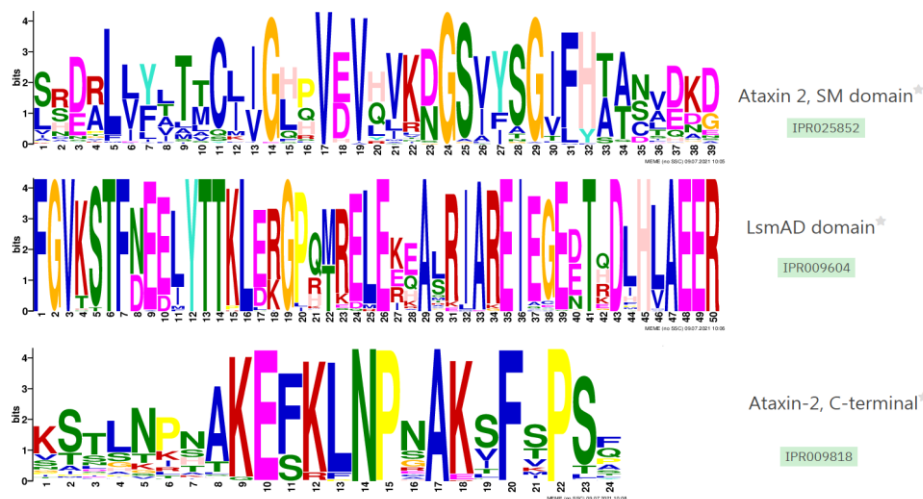
به منظور تجزیه خوشه و آنالیزهای فیلوژنتیک برای بررسی واگرایی نسبی در بزرگ خانواده *ATAXIN-2* از برنامه Phylyp (Retief, 2000) استفاده شد. هم‌ردیفی‌های چندگانه توسط MAFFT (<https://mafft.cbrc.jp/alignment/software/>) با استفاده از روش L-INS-i ایجاد شد (Kato & Standley, 2013). هم‌ردیفی‌های ضعیف و دارای گپ با استفاده از trimAL (<http://trimal.cgenomics.org/>) برداشته شدند (Capella-Gutierrez *et al.*, 2009). هم‌ردیفی‌های با شکاف کمتر در نظر گرفته شدند. درخت فیلوژنتیک با استفاده از مجموعه Phylyp با استفاده از برنامه‌های NEIGHBOR، FITCH و PROML با ۱۰۰ تکرار بوت استرپ ساخته شد. به منظور حفاظت از باقی‌مانده‌های عملکردی، توالی کامل تمام پروتئین‌های موجود در بزرگ خانواده *ATAXIN-2* برای اجرای MEME (<https://meme-suite.org/>) به منظور جستجوی ۳ موتیف برتر به طول بین ۲۰ تا ۶۰ باقیمانده (با افزایش ۵ باقیمانده و $E\text{-value} < 10^{-10}$) استفاده شد. برای دستیابی به همگرایی از حداکثر ۱۰۰۰ تکرار و حداقل فاصله ۱۰-۷ بین ماتریس‌های تناوب موتیف استفاده شد.

از ۵ موتیف شناسایی شده برتر در جستجوهای اولیه، فقط سه نمونه دارای $E\text{-value} < 10^{-100}$ MEME بودند و از این ۳ موتیف شناسایی شده توسط MEME برای اجرای MAST و مکان‌یابی موتیف (مقدار $E\text{-value} < 10^{-5}$) برای شناسایی سایر پروتئین

های حاوی این موتیف‌ها استفاده شد. وب سرور MEME یک پورتال واحد برای کشف و تجزیه و تحلیل آنلاین موتیف توالی ارائه دهنده ویژگی‌هایی مانند سایت‌های اتصال DNA و دمین‌های تعامل پروتئین فراهم می‌کند. سه الگوریتم اسکن توالی MAST، FIMO و GLAM2SCAN امکان اسکن چندین پایگاه داده DNA و توالی پروتئین برای موتیف کشف شده توسط MEME و GLAM2 را فراهم می‌کند. خروجی MEME حاوی توالی LOGOS برای هر موتیف کشف شده و همچنین گزینه‌هایی است که به کاربر امکان می‌دهد موتیف‌ها به راحتی به الگوریتم‌های اسکن پایگاه داده توالی و موتیف MAST، FIMO و TOMTOM یا GOMO برای تجزیه و تحلیل بیشتر ارسال شوند (Bailey *et al.*, 2015). در نهایت، عملکرد شرکای تعامل‌کننده با پروتئین ATAXIN-2 با استفاده از وب سایت STRING (<http://www.bork.embl-heidelberg.de/STRING/>) پیش بینی شد (von Mering *et al.*, 2003).

نتایج و بحث

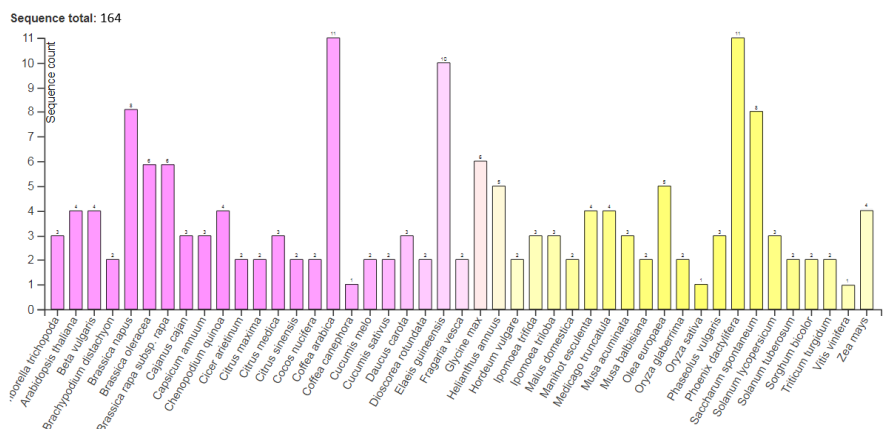
نتایج جست و جو Pfam با اعضای خانواده پروتئین ATAXIN-2 برای مطالعه ساختار دمین استفاده شد. خانواده پروتئین ATAXIN-2 در گیاهان ترکیبات مختلفی از دمین‌های Pfam قابل شناسایی دارند. ۳ دمین اصلی شناسایی شده (به ترتیب با طولی حدود ۳۹، ۵۰ و ۲۴ آمینواسید) در اغلب گونه‌ها در شکل ۱ به صورت لوگو نشان داده شده است. مجموعه‌ای از ۱۶۴ توالی حاوی حداقل یکی از ۳ دمین اصلی یعنی SM domain، LsmAD domain و C-terminal از ۴۵ گونه گیاهی بازبایی شدند. شکل ۲ توزیع اعضای خانواده پروتئین ATAXIN-2 در ۴۵ گونه گیاهی را نشان می‌دهد. دمین اصلی (اول) و دمین سوم در هر خانواده در اکثر گونه‌ها وجود داشت، در حالیکه دمین‌های دوم، چهارم، پنجم و ششم شناسایی شده در همه اعضا شناسایی نشدند. ساختار دمین‌ها از خانواده‌ها، شامل دمین‌های متمایز و همپوشان بودند. دمین اصلی (اول)، SM domain، ATAXIN-2 با اکسشن (*IPR025852*) در پایگاه داده InterPro بود. دمین دوم LsmAD domain با اکسشن (*IPR009604*)، دمین سوم C-terminal، ATAXIN-2 با اکسشن (*IPR009818*)، دمین چهارم LSM domain superfamily (*IPR010920*)، دمین پنجم UBA-like superfamily (*IPR009060*) و دمین ششم GBF-interacting protein 1, N-terminal (*IPR009719*) بودند.



شکل ۱- موتیف‌های اصلی شناسایی شده در اکثر خانواده پروتئین *ATAXIN-2* با استفاده از سایت MEME

Figure 1. Major motifs identified in most of the *ATAXIN-2* protein family using the MEME site

بیشترین تعداد پروتئین *ATAXIN-2* در گونه‌های قهوه *Coffea arabica*، نخل خرما *Phoenix dactylifera*، نخل روغنی *Elaeis guineensis*، کلزا *Brassica napus*، نیشکر *Saccharum spontaneum*، کلم *Brassica oleracea* و شلغم *Brassica rapa* *subsp. Rapa* مشاهده شد (شکل ۲).

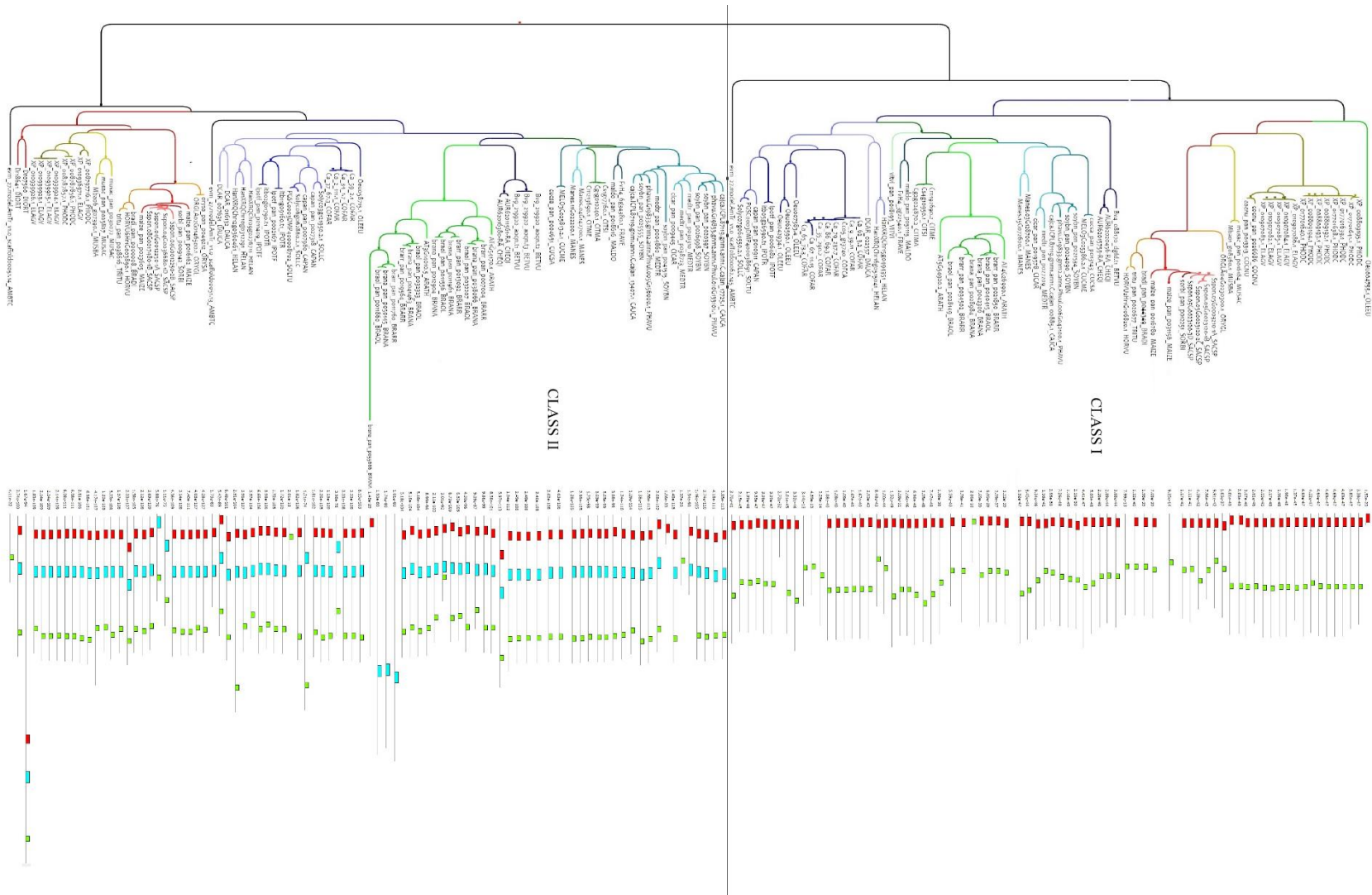


شکل ۲- توزیع اعضای خانواده پروتئین *ATAXIN-2* شامل ۱۶۴ پروتئین در ۴۵ گونه گیاهی

Figure 2. Distribution of *ATAXIN-2* protein family members including 164 proteins in 45 plant species

ارتولوگ‌های فرضی در اغلب گونه‌ها مشخص شدند، که نشان می‌دهد *ATAXIN-2* ژن‌های مهمی در ژنوم‌های گیاهی هستند. برای به دست آوردن فیلوژنی‌ها از ۱۶۴ پروتئین *ATAXIN-2* با پشتیبانی از طبقه‌بندی شاخه‌ها و رزولوشن مرسوم گونه‌ها، درختان

براساس ۳ دمین حفاظت شده مقایسه شدند (شکل ۳). فیلوژنی‌های تولید شده با روش‌های (NJ) neighbor-joining، maximum-likelihood و (MP) parsimony (ML) maximum-likelihood تولید و مقایسه شدند. درختان حاصل از هر ۳ روش توپولوژی مشابهی داشتند و گونه‌های گیاهان را به صورت کلادهای جداگانه گروه‌بندی کردند. در درخت تولید شده با روش NJ اتفاقات ناسازگاری (عدم تطابق) طبقه‌بندی در اکثر کلادها کمتر از فیلوژنی‌های بدست آمده از طریق ML و MP بود (درختان تولید شده با روش MP و ML نشان داده نشده است). برای به دست آوردن یک دید جامع از معماری دمین در میان پروتئین‌های *ATAXIN-2* در گونه‌های مختلف، نتایج جستجوی موتیف توالی در ۱۶۴ توالی اسید آمینه *ATAXIN-2* توسط مجموعه MEME در کنار درخت فیلوژنی آورده شده است.



شکل ۳- رسم درخت فیلوژنی از ۱۶۴ پروتئین *ATAXIN-2* و ۴۵ گونه ی گیاهی براساس ۳ دمین حفاظت شده
Figure 3. Phylogeny tree drawing from 164 *ATAXIN-2* proteins and 45 plant species based on 3 conserved domains

- ۱ برای هر دمین توالی از رنگ متفاوتی استفاده شده است. موتیف‌ها همراه با درخت فیلوژنتیک و کاتالوگ آنها در شکل ۳
- ۲ نمایش در آمده است. علاوه بر این، هم‌ردیفی پروتئین‌های گیاهان و تعیین مکان مناطق شامل موتیف توالی توسط مجموعه MEME
- ۳ ، حفاظت از موتیف‌ها را در میان گیاهان در سه منطقه با طولی حدود ۳۹، ۵۰ و ۲۴ باقیمانده را نشان داد که به ترتیب ۱۰، ۲۱ و ۱۰
- ۴ باقیمانده در آن غالب هستند (شکل ۱). تمام توالی‌های ATAXIN-2 در گیاهان به دو دسته اصلی تقسیم شدند. یک کلاس شامل کلاس
- ۵ I و کلاس دیگر شامل کلاس II بود (شکل ۳). در کلاس I اکثر توالی‌ها دارای دو موتیف حفاظت شده اصلی SM domain (موتیف با
- ۶ رنگ قرمز در شکل ۳) و موتیف سوم C-terminal (موتیف با رنگ سبز در شکل ۳) بودند. در کلاس II اکثر توالی‌ها دارای هر سه
- ۷ موتیف حفاظت شده SM domain، C-terminal و LsmAD domain (موتیف با رنگ آبی در شکل ۳) بودند. اکثر گونه‌های مورد
- ۸ مطالعه دارای هر دو نوع کلاس I و II پروتئین ATAXIN-2 بودند. به جز نخل نارگیل *Cocos nucifera*، *Coffea canephora*
- ۹ و *Solanum tuberosum* که دارای دو موتیف اصلی و *Oryza sativa* فقط دارای یک توالی با هر ۳ موتیف اصلی بود. وجود هر دو نوع
- ۱۰ کلاس I و II پروتئین ATAXIN-2 در اکثر گونه‌های مورد مطالعه نشان‌دهنده حفاظت شدید هر دو نوع توالی است و احتمالاً اهمیت
- ۱۱ عملکرد این خانواده پروتئینی را در گیاهان نشان می‌دهد. نتایج پیش‌بینی عملکرد پروتئین ناشناخته در گیاه برنج از طریق رسم
- ۱۲ درخت فیلوژنی بر مبنای موتیف‌های ساختاری به همراه پیش‌بینی عملکرد شرکای تعامل‌کننده آنها توسط سایت STRING، عملکرد
- ۱۳ مولکولی و فرایندهای بیولوژیکی آن در جدول ۱ نشان داده شده است. بالاترین همولوژی‌های معنی‌دار، عملکرد مولکولی پروتئین
- ۱۴ ATAXIN-2 را RNA binding و عملکرد بیولوژیکی آنها را تنظیم mRNA سیتوپلاسمی مشخص کرد. گزارش شده است که ATAXIN-
- ۱۵ 2 در موجودات مختلف عملکردهای بسیاری مانند سیکلینگ گیرنده‌های درون ریز، تنظیم ترجمه، رشد جنینی و متابولیسم انرژی
- ۱۶ دارد و با بسیاری از پروتئین‌ها برهم‌کنش دارد. (Lastres-Becker et al (2008) نقش اصلی پروتئین ATAXIN-2 را در تعدیل گسترده
- ۱۷ ترجمه موضعی mRNA بیان می‌کنند. این نقش ATAXIN-2 احتمالاً به مکانیسم جایگزین برای بومی‌سازی پروتئین اشاره دارد که
- ۱۸ در آن mRNA رمزگذاری کننده پروتئین هدف قرار گرفته می‌شود و متعاقب آن در محل ترجمه می‌شود. رایج‌ترین مکانیزم برای
- ۱۹ بومی‌سازی پروتئین شامل هدف‌گیری مستقیم خود پروتئین از طریق توالی‌های خاص مانند توالی‌های بومی‌سازی هسته‌ای یا
- ۲۰ میتوکندری است. بسیاری از پروتئین‌های سلولی در مکان‌های اختصاصی زیر سلولی قرار می‌گیرند. بومی‌سازی فضایی امکان تقسیم
- ۲۱ بندی عملکردی را فراهم می‌کند. با این حال شواهد فزاینده‌ای نشان می‌دهد که بومی‌سازی RNA، جایی که mRNA های خاص به
- ۲۲ زیر سلول منتقل می‌شوند و متعاقباً در پاسخ به سیگنال‌های موضعی ترجمه می‌شوند، مکانیزمی است که برای کنترل بومی‌سازی
- ۲۳ پروتئین حفاظت می‌شود. سنتز در محل، ویژگی‌های جدید علامت دهی را به یک پروتئین می‌بخشد و به حفظ هموستاز پروتئوم
- ۲۴ موضعی کمک می‌کند (Imai & Nakai, 2010).

جدول ۱- معنی دارترین همولوژی های عملکردی پیش بینی شده برای پروتئین ناشناخته به همراه پیش بینی عملکرد شرکای تعامل کننده و فرایند های بیولوژیکی آنها

Table 1. The most significant functional homologies predicted for the unknown protein along with the prediction of function of interacting partners and their biological processes.

Predicted function	Interacted protein	Molecular function	Biological function	E-value
Polyadenylate-binding protein-interacting protein	Polyadenylate-binding protein	RNA-binding	Translational initiation, Regulation of cytoplasmic mRNA, P-body assembly, Stress granule assembly	2.49e ⁻¹⁰⁸
SM-ATX -LsmAD domain-containing protein	Polyadenylate-binding protein	RNA-binding	Translational initiation, mRNA metabolic process	8.58e ⁻¹⁰¹
SM-ATX -LsmAD domain-containing protein	AD domain-containing protein	RNA binding	-	9.86e ⁻⁹⁹
LsmAD domain-containing protein	Belongs to the DEAD box helicase family	mRNA binding, helicase activity	Stress granule assembly, P-body assembly	1.74e ⁻⁸⁰
Polyadenylate-binding protein interacting protein	AD domain-containing protein	RNA binding	-	9.09e ⁻⁷⁸
SM-ATX domain-containing protein	Bromo domain-containing protein	-	Transcriptional activator	8.99e ⁻⁶⁹
SM-ATX -LsmAD domain-containing protein	TPR_MLP1_2 domain-containing protein	Structural constituent of nuclear pore	mRNA and Protein transport	3.68e ⁻⁴⁴
SM-ATX -LsmAD domain-containing protein	MI domain-containing protein	mRNA binding	Translation regulation, Translation initiation factor activity	3.70e ⁻⁴¹
SM-ATX -LsmAD domain-containing protein	CNOT1_X* -containing protein	Molecular adaptor activity, Repressor	Negative regulation of translation, Nuclear-transcribed mRNA catabolism, Deadenylation-dependent decay	1.98e ⁻³⁹
Polyadenylate-binding protein-interacting protein	Nuclear-pore anchor	Structural constituent of nuclear pore	Regulation of gene silencing, mRNA transport, Protein transport	1.45e ⁻²⁴
SM-ATX domain-containing protein	Ethylene-responsive transcription factor 3	DNA-binding, transcription factor activity	Ethylene-activated signaling pathway, Defense response	4.11e ⁻²²

* نشان دهنده دمین HEAT, TTP_bind و CAF1_bind می باشد.

پیشرفت‌های چشمگیر در روش‌های تشخیص RNA منجر به شناسایی تعداد فزاینده mRNA های موضعی شده است. تا اوایل دهه گذشته، مجموعه‌ای از mRNA های هدفمند توصیف شده تقریباً به ۱۰۰ عدد و فرایند بومی‌سازی mRNA به سلول‌های خاص محدود بود. با این حال، تجزیه و تحلیل‌های اخیر در سطح ژنوم این دیدگاه را به طرز چشمگیری تغییر داده است و نشان می‌دهد که هدف قرار دادن زیر سلول mRNA ها مکانیزمی رایج است که توسط سلول‌های قطبی برای ایجاد محفظه‌های عملکردی مشخص استفاده می‌شود (Palacios & Johnston, 2001). تجمع قطبی شده مولکول‌های RNA برای اولین بار نزدیک به ۳۰ سال پیش، هنگامی که مشخص شد mRNA β -اکتین به طور نامتقارنی در تخمک‌ها و جنین‌های برخی از موجودات قرار دارد، تجسم یافت (Jeffery *et al.*, 1983). به دنبال این، کشف اولین mRNA های بومی مادر در تخمک‌های زنبوپوس و دریزوفیلا شواهدی را برای پیشنهاد قبلی ارائه داد که تعیین‌کننده‌های موضعی RNA می‌توانند مسئول الگوی اولیه جنینی باشد (Kandler-Singer & Kalthoff, 1976). به زودی مشخص شد که mRNA ها به طور نامتقارن در سلول‌های سوماتیک متمایز مانند فیبروبلاست، الیگودندروسیت‌ها و سلول‌های عصبی توزیع می‌شوند و با پروتئین‌های رمزگذاری شده آنها هم بومی‌سازی می‌شوند و انتقال داخل سلولی mRNA ها به مکان‌های مجزا را فراهم می‌کند. یک مطالعه در مقیاس بزرگ هیبریداسیون در محل در جنین‌زایی دریزوفیلا نشان داد، ۷۰ درصد از mRNA ژن‌های بررسی شده (۲۰ از کل ژن‌ها) به بخش‌های متمایز زیر سلولی بومی‌سازی می‌شوند، جایی که در بسیاری از موارد، آنها با پروتئین‌هایی که رمزگذاری می‌کنند هم بومی‌سازی می‌شوند (Lécuyer *et al.*, 2007). همچنین صدها تا هزاران mRNA در محفظه‌های سلولی به اندازه دستگاه میتوزی، پای کاذب (pseudopodia)، سلول‌های کاملاً تخصصی مانند نورون‌های عصبی، دندریت یا آکسون متنوع تشخیص داده شده‌اند (Medioni *et al.*, 2012).

هدف‌گیری mRNA درون سلولی با شناسایی mRNA های موضعی در طیف وسیعی از ارگانسیم‌های به غیر از حیوانات، از جمله مخمر، باکتری‌ها، قارچ‌ها و گیاهان گسترش پیدا کرد. با این وجود اطلاعات زیادی در مورد نقش مکانیسم بومی‌سازی RNA در عملکردهای خاص سلولی گیاهی وجود ندارد. Heidari (2010) در بررسی بیان ژن و ژنومیکس کارکردی مسیرهای مختلف مورفوژنز با استفاده از تجزیه و تحلیل منابع EST افزایش بیان معنی‌دار توالی EST منتسب به *ATAXIN-2* در جنین‌زایی برنج را گزارش نمودند. معمولاً توالی‌هایی که در نشان دادن همولوژی معنی‌دار برای هر پروتئین در پایگاه داده‌های عمومی شکست می‌خورند، کاندیدای مناسبی برای ژنهای جدید هستند (Heidari *et al.*, 2012). ایجاد محور رأس-پایه یک رویداد مهم در جنین‌زایی گیاه است، که از اولین مراحل در کیسه جنین، سلول تخمک و زیگوت به بعد مشهود است. تعدادی از جهش‌یافته‌ها، به ویژه از ارگانسیم ژنتیکی مدل عملکرد ژنی جدیدی را که برای الگوسازی محور راس-پایه مورد نیاز است نشان داده‌اند. اکنون شواهد بیشتری وجود دارد که نشان می‌دهد نوع اختصاصی از علامت دهی نقش مهمی در هماهنگی برنامه‌های بیان ژن برای ایجاد قطبیت تعیین

می‌کنند. بنابراین، mRNA ها چیزی فراتر از "پیام رسان" های ساده هستند که اطلاعات ژنتیکی را از DNA به دستگاه سنتز پروتئین می‌رسانند جایی که، mRNA های مرتبط با عملکرد می‌توانند همزمان با توجه به نیازهای بیولوژیکی ترجمه شوند و علامت دهی کارآمد برای کنترل هماهنگ بیان ژن فراهم می‌کنند (Keene & Tenenbaum, 2002).

مشخص شده است که در جنین، از طریق بیان موضعی ژن‌های کلیدی، دو مریستم عملکردی متمایز در هر قطب تشکیل می‌شود. جلبک قهوه‌ای *Fucus* برخی از ویژگی‌های تجربی را در مطالعه وقایع اولیه قطب‌بندی زیگوت را فراهم می‌کند. مشاهده شده است که همراه با تشکیل محور، بومی‌سازی یا توزیع مجدد اجزای غشا پلازما از جمله کانال‌های یونی، توزیع مجدد کلسیم به انتهای پایه، بومی‌سازی F-اکتین در ریزودرم، توزیع نامتقارن مولکول‌های RNA در زیگوت (هر چند mRNA اکتین به طرز جالبی در قطب مخالف پروتئین F-اکتین جمع می‌شود) و یک ترشح قطبی از اجزای دیواره سلولی مشتق شده از گلژی به سمت منطقه "پایه" که سلول ریزوئید از آن توسعه می‌یابد، اتفاق می‌افتد (Shaw & Quatrano, 1996). این مکانیسم مبتنی بر RNA، شامل هماهنگی فرایندهای پیچیده متعدد، از جمله نقل و انتقال mRNA، هدف‌گیری و ترجمه است و کنترل دقیق ناشی از انواع محرک‌ها بر موقعیت پروتئین، فراوانی و تا حدودی عملکرد را قادر می‌سازد. از این رو می‌توان استدلال نمود که در گیاهان نیز مانند سایر موجودات مجموعه‌های متفاوتی از mRNA برای بخش‌های مختلف هدف قرار گرفته می‌شود و سلول‌ها را به سمت پاسخ دادن به محرک‌های مختلف و برنامه‌ریزی مجدد سلولی به سمت تکثیر، تمایز، رشد، آپوپتوز و غیره سوق می‌دهد. این لایه جدید از الگوی داخل سلولی که در ابتدا تصور می‌شد منحصر به سلول‌های بسیار تخصصی است ممکن است به طور گسترده‌ای در بسیاری از انواع سلول‌ها وجود داشته باشد و نقش *ATAXIN-2* را در این زمینه در گیاهان برجسته می‌کند. این بررسی نقش احتمالی *ATAXIN-2* را در ترجمه موضعی به عنوان یک حالت متمایز از کنترل بیان ژن نشان می‌دهد که عملکرد ژن را با دقت، کارایی و انعطاف پذیری مکانی-زمانی برنامه‌ریزی می‌کند. تنظیم بیان ژن یک فرآیند ضروری است که از طریق آن سلول‌های موجودات با تغییرات در محیط خود مقابله می‌کنند. تجزیه و تحلیل اطلاعات مربوط به ارتولوگ‌های کاوش نشده از طریق MAST زیر مجموعه MEME، وجود دمین LsmAD را در شبدر قرمز *Trifolium pratense*، دمین SM domain را در *Macleaya cordata*، هر دو دمین LsmAD و SM domain را در *Ipomoea nil* و دمین C-terminal را در *Artemisia annua* شناسایی کرد. اطلاعات مربوط به ارتولوگ‌های کاوش نشده می‌تواند برای درک بهتر نقش پروتئین‌های *ATAXIN-2* در متابولیسم mRNA مفید باشد.

نتیجه گیری

این مطالعه به بررسی ژنوم و تجزیه و تحلیل دمین پروتئین‌های ATAXIN-2 در گونه‌های گیاهی می‌پردازد. ۱۶۴ توالی پروتئینی در ۴۵ گونه انتخاب شده شناسایی شد. این توالی‌های ارتولوگ در گیاهان ممکن است در طی یک رویداد تکاملی ژنوم ایجاد شده باشند. در برخی گونه‌ها تا یازده ژن از این خانواده پروتئینی مشاهده شد و دو نوع کلی کلاس از پروتئین‌های ATAXIN-2 مشخص شد. بنابراین، عملکرد اختصاصی را می‌توان برای این ارتولوگ‌های ATAXIN-2 استنباط کرد. عدم وجود LsmAD در یک کلاس که به توزیع متفاوت زیر سلول و سایر ویژگی‌ها اشاره دارد ممکن است ویژگی‌های عملکردی متمایزی را نسبت به کلاس دیگر ATAXIN-2 اعطا کند. شناخت خوب از دامنه تنوعی که بر یک ویژگی مشخص تأثیر می‌گذارد، درک درستی از مجموعه ژن‌های موجود را فراهم می‌کند (Heidari *et al.*, 2019-2020). تعامل پروتئین-پروتئین در پیش‌بینی عملکرد پروتئین هدف نقش اساسی دارد. روش‌های تجربی مانند تصفیه افینیتی، Y2H (مخمر ۲ هیبریدی)، TAP (تصفیه افینیتی هیبریدی پشت سر هم) و غیره محدودیت‌های مانند هزینه، زمان و غیره خاص خود را دارند. بنابراین، روش‌های بیوانفورماتیک که شامل رویکردهای مبتنی بر ساختار و درخت فیلوژنتیک است به روشن شدن عملکرد پروتئینی و مسیرهای انتقال سیگنال کمک زیادی می‌کند.

در این پژوهش عملکرد ATAXIN-2 که در جنین‌زایی گیاه برنج نقش ایفا می‌کند به عنوان پروتئین‌های متصل‌شونده به RNA (RNA binding) مشخص شد که در تنظیم mRNA سیتوپلاسمی نقش دارد. نقش پروتئین‌های متصل‌شونده به RNA (RPBs) ممکن است در هدف قرار دادن زیرمجموعه‌های جداگانه mRNA به مکان‌های زیر سلول باشد که در آنجا منتظر سیگنالی برای ادامه سنتز پروتئین‌ها با توجه به شرایط ذاتی و خارجی هستند. ترجمه بومی انتقال mRNA ها به جای پروتئین‌ها چندین مزیت قابل توجه برای یک سلول دارد. کاهش هزینه‌های نقل و انتقال، جلوگیری از بیان ژن قبل از رسیدن mRNA به محل مناسب به ویژه در مورد عوامل تعیین‌کننده مادر بسیار مهم است، زیرا در تعدادی از موجودات مشاهده شده است که بیان نامناسب فضایی، الگوی جنینی را مختل می‌کند. یک مزیت عمده در هدف قرار دادن mRNA این است که امکان تنظیم دقیق بیان ژن را در فضا و زمان مناسب فراهم می‌کند. علاوه بر این به خاطر تغییرات پس از ترجمه یا از طریق مسیرهای تاخوردگی با کمک چاپرون‌های مولکولی، پروتئین‌های جدید شکل گرفته ممکن است دارای خواص متمایز از نسخه‌های قبلی باشند. علاوه بر این ترجمه بومی می‌تواند در کمپلکس‌های ماکرومولکولی غلظت پروتئین‌ها را تغییر دهد و امکان ترجمه همزمان زیرواحدهای مختلف را تسهیل کند که تمام موارد ذکر شده برای یک جنین دو قطبی از اهمیت زیادی برخوردار است. پیشرفت‌های اخیر در بررسی تنوع ژنتیکی، جداسازی و شناسایی ژن‌های مسئول مسیرهای مختلف مورفوژنز را تسهیل نموده است (Heidari *et al.*, 2017). از آنجا که پروتئین ATAXIN-2 با افزایش بیان و

تفاوت معنی‌دار در کتابخانه جنین‌زایی در گیاه برنج شناسایی شد و عملکرد آن در گیاهان پیش‌بینی شد می‌توان نسبت به جداسازی آن در راستای مطالعات آزمایشگاهی بیشتر اقدام نمود.

سپاسگزاری

نویسندگان مقاله از دانشگاه آزاد اسلامی واحد کرمانشاه برای حمایت از پژوهش حاضر سپاسگزاری می‌کنند.

تعارض منافع

نویسندگان اعلام می‌کنند که هیچ تضاد منافی ندارند.

منابع

- Albrecht, M., Golatta, M., Wüllner, U. and Lengauer, T. (2004). Structural and functional analysis of ataxin-2 and ataxin-3. *Eur. j. biochem.*, 271(15): 3155-3170.
- Bailey, T.L, Johnson, J., Grant, C.E. and Noble, W.S. (2015). The MEME Suite. *Nucleic Acids Res.*, 43: W39-49.
- Brown, D. and Sjölander, K. (2006). Functional classification using phylogenomic inference. *PLoS Comput. Biol.*, 30: 2(6): e77.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J. and Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinform.*, 10: 421.
- Capella-Gutierrez, S., Silla-Martinez, J.M, Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinform.*, 25(15): 1972-3.
- Chothia, C., Gough, J., Vogel, C. and Teichmann, S.A. (2003). Evolution of the protein repertoire. *Sci.*, 300(5626): 1701-1703.
- Eddy, S.R. (2011). Accelerated Profile HMM Searches. *PLoS Comput. Biol.*, 7(10): e1002195.
- Engelhardt, B.E., Jordan, M.I., Repo, S.T. and Brenner, S.E. (2009). Phylogenetic molecular function annotation. *Journal of physics. Conference series*, 180(1), 12024.
- Finn, R.D., Coghill, P., Eberhardt, R. Y., Eddy, S. R., Mistry, J., Mitchell, A. L. et al. (2016). The Pfam protein family database: towards a more sustainable future. *Nucleic Acids Res.*, 44(D1): D279-85.
- Heidari, Sh. (2010). A study of gene expression and functional genomics of different in vitro morphogenesis pathways by analyzing expressed sequence tags (ESTs). MSc. Thesis on Agricultural Science, Department of Plant breeding, Shiraz university, 380P.
- Heidari, Sh., Azizinezhad, R. and Haghparast, R. (2017). Investigation on genetic diversity in *Triticum turgidum* L. Plant var. durum using agro-morphological characters and molecular markers. *Indian J. Genet. Breed.*, 77(2): 242-250.
- Heidari, Sh., Azizinezhad, R., Haghparast, R. and P. Heidari. (2019). Evaluation of the association among yield and contributing characters through path coefficient analysis in advanced lines of durum wheat under diverse conditions. *J. Anim. Plant Sci.*, 29(5): 1325-1335.

- Heidari, Sh., Heidari, P., Azizinezhad, R., Etmnan, A. and Khosroshahli, M. (2020). Assessment of genetic variability, heritability and genetic advance for agro-morphological and some in-vitro related-traits in durum wheat. *Bulg. J. Agric. Sci.*, 26(1): 120–127.
- Heidary, P., Maleki Zanjani, B., Heidary, S. (2012). A study of gene expression and functional genomics of wheat, rice, cotton and festuca plants under drought stress by analyzing expressed sequence tags (EST). *modern genetics journal (mgj)*, 7(2 (29)): 129-140.
- Imai, K. and Nakai, K. (2010). Prediction of subcellular locations of proteins: where to proceed? *Proteomics*, 10: 3970-3983.
- Jeffery, W. R., Tomlinson, C. R. and Brodeur, R. D. (1983). Localization of actin messenger RNA during early ascidian development. *Dev. Biol.*, 99: 408-417.
- Jiménez-López, D. and Guzmán, P. (2014). Insights into the evolution and domain structure of ATAXIN-2 proteins across eukaryotes. *BMC research notes*, 7, 453.
- Kandler-Singer, I. and Kalthoff, K. (1976). RNase sensitivity of an anterior morphogenetic determinant in an insect egg (*Smittia* sp., Chironomidae, Diptera). *Proc. Natl. Acad. Sci. USA*, 73: 3739-3743.
- Katoh, K., Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, 30(4): 772-80
- Keene, J. D. and Tenenbaum, S. A. (2002). Eukaryotic mRNPs may represent posttranscriptional operons. *Mol. Cell*, 9: 1161-1167.
- Kiehl, T. R., Nechiporuk, A., Figueroa, K. P., Keating, M. T., Huynh, D. P. and Pulst, S. M. (2006). Generation and characterization of Sca2 (ATAXIN-2) knockout mice. *Biochem. Biophys. Res. Commun.*, 339(1): 17-24.
- Kiehl, T. R., Shibata, H. and Pulst, S. M. (2000). The ortholog of human ATAXIN-2 is essential for early embryonic patterning in *C. elegans*. *J. Mol. Neurosci.*, 15(3) :231-41.
- 7: 115-124. Lastres-Becker, I., Rüb, U. and Auburger, G. (2008). Spinocerebellar ataxia 2 (SCA2). *Cerebellum*, Lécuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T. R., Tomancak, P., Krause, H. M. (2007). Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell*, 131: 174-187.
- Li, Y. R., King, O. D., Shorter, J. and Gitler, A. D. (2013). Stress granules as crucibles of ALS pathogenesis. *J. Cell Biol.*, 201(3): 361-372.
- Marchler-Bauer, A., Zheng, C., Chitsaz, F., Derbyshire, M. K., Geer, L. Y, Geer, R. C. et al. (2013). CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Res.*, 41: D348-52.
- Medioni, C., Mowry, K. and Besse, F. (2012). Principles and roles of mRNA localization in animal development. *Development (Cambridge, England)*, 139(18): 3263-3276.
- Nadzirin, N. and Firdaus-Raih, M. (2012). Proteins of unknown function in the Protein Data Bank (PDB): an inventory of true uncharacterized proteins and computational tools for their analysis. *International journal of molecular sciences*, 13(10): 12761-12772.

- Nonhoff, U., Ralser, M., Welzel, F., Piccini, I., Balzereit, D., Yaspo, M. L., Lehrach, H. and Krobitsch, S. (2007). Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with P-bodies and stress granules. *Mol. Biol. Cell*, 18(4):1385-96.
- Palacios, I. M. and St Johnston, D. (2001). Getting the message across: the intracellular localization of mRNAs in higher eukaryotes. *Annu. Rev. Cell Dev. Biol.*, 17: 569-614.
- Retief, J. D. (2000). Phylogenetic Analysis Using PHYLIP. In: Misener S., Krawetz S.A. (eds) *Bioinformatics Methods and Protocols. Methods Mol. Biol.*, 132. Humana Press, Totowa, NJ.
- Shaw, S. L. and Quatrano, R. S. (1996). The role of targeted secretion in the establishment of cell polarity and the orientation of the division plane in *Fucus* zygotes. *Development*, 122: 2623-2630.
- von Mering, C., Huynen, M., Jaeggi, D., Schmidt, S., Bork, P. and Snel, B. (2003). STRING: a database of predicted functional associations between proteins. *Nucleic Acids Res.*, 31(1): 258-261.
- Wong, K. C. (2019). Big data challenges in genome informatics. *Biophys. Rev.*, 11(1): 51-54.

Evolutionary relationships of the *ATAXIN-2* protein domain structure in plants: A possible insight into the involvement of this protein in localized translation in rice (*Oryza sativa*) plant embryogenesis process

Sh. heidari¹, P. heidari², B. Heidari^{3*}

Accepted: 2022.05.29

Received: 2021.09.19

Abstract

Introduction: With the development of high-throughput sequencing technologies, sequence data is collected at a high rate and their functional annotation remains a major challenge. The *ATXN2* gene encodes a protein with an unknown function in rice (*Oryza sativa*). The aim of this study was to determine the structure of the domain and predict the function of this protein. **Methods:** The sequence motif phylogenetic tree with 164 *ATAXIN-2* amino acid sequences from 45 plant species was analyzed by MEME Suite and Phylip software. **Results and discussion:** The results revealed two general classes of *ATAXIN-2* proteins. The absence of the LsmAD domain in one class may confer distinct functional characteristics compared to another *ATAXIN-2* class. In this study, the function of *ATAXIN-2* was found to be RNA binding protein that plays a role in the regulation of cytoplasmic mRNA in rice embryogenesis. Findings from this study could be the basis for research on the function of this protein in targeting separate subsets of mRNA to subcellular locations in plants, which are recruited by embryonic bipolar cells to establish specific functional compartments.

Keywords: *Local translation, Phylogenetics, Prediction of function, RNA-binding proteins ing proteins*

1, 2. Department of Plant Breeding, Science and Research Branch, Islamic Azad University, Tehran, Iran

3. Assistant Professor, Department of Computer Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran

(*Corresponding author: b_heidari@iauksh.ac.ir)