



دانشگاه گوارش و سبزیجات

نشریه پژوهش‌های تولید گیاهی

جلد بیست و هفتم، شماره چهارم، ۱۳۹۹

۱۸۱-۱۹۱

<http://jopp.gau.ac.ir>

DOI: 10.22069/jopp.2020.17243.2587

مطالعه فیلوژنی مولکولی و ساختار پروتئین *matk* در سوسن چلچراغ

(*Lilium ledebourii* [Baker] Boiss)

مرتضی شیخ اسدی^۱، *روح‌انگیز نادری^۱، محسن کافی^۱، محمدرضا فتاحی‌مقدم^۱ و سید علیرضا سلامی^۱

^۱گروه مهندسی علوم باغبانی و فضای سبز، دانشگاه تهران، کرج، ایران

تاریخ دریافت: ۱۳۹۸/۰۷/۲۰؛ تاریخ پذیرش: ۱۳۹۸/۱۰/۰۹

چکیده

سابقه و هدف: ژن کلروپلاستی *matk* یکی از بهترین ژن‌های بارکد گیاهی است و از مناسب‌ترین ژن‌ها برای بررسی روابط فیلوژنتیکی و تکاملی در سطوح گونه تا خانواده معرفی شده است. سوسن چلچراغ از گونه‌های ارزشمند بومی ایران است که به‌شدت در معرض خطر انقراض قرار دارد. این گونه علی‌رغم ویژگی‌های ارزشمندی که دارد، تاکنون در هیچ مطالعه‌ای با سایر گونه‌های این خانواده از نظر جایگاه فیلوژنی مورد بررسی قرار نگرفته است. این پژوهش با به‌کارگیری توالی *matk* به بررسی جایگاه رده‌بندی و قرابت ژنتیکی سوسن چلچراغ با ۴۱ گونه دیگر از این جنس پرداخته است.

مواد و روش‌ها: استخراج DNA با استفاده از کیت Qiagen DNEasy انجام شد. بعد از تأیید کیفیت و کمیت DNA، توالی‌یابی به‌صورت Real-Time (SMRT) با استفاده از پلتفرم نسل جدید PacBio انجام و توالی کامل ژن *matK* با استفاده از ابزارهای بیوانفورماتیک استخراج و در پایگاه اطلاعات ژنی^۱ NCBI به شماره دسترسی MN557236 ثبت شد. به‌منظور بررسی روابط فیلوژنتیک، توالی *matk* ۴۱ گونه دیگر لیلیوم از پایگاه اطلاعات ژنی به‌دست آمد. همه توالی‌ها با روش ClustalW و با استفاده از نرم‌افزار مگا ۷ هم‌ردیف شدند. روش حداکثر درست‌نمایی جهت رسم درخت فیلوژنی به‌کار گرفته شد. فاصله ژنتیکی به روش K₂P محاسبه شد. میزان حداکثر درست‌نمایی مرکب الگوی جایگزینی نوکلئوتیدی با استفاده از ماتریس جایگزینی برآورد شد. با استفاده از ابزارهای تخصصی FFpred، InterProScan، TargetP و Phyre2 ساختارهای ثانویه‌ای مانند ماریچ‌های α ، β و ماریچ‌های تصادفی و ساختار سه‌بعدی پروتئین مورد بررسی قرار گرفت.

یافته‌ها: طبق نتایج درخت فیلوژنی، لیلیوم‌ها به چهار خوشه A، B، C و D تقسیم شدند که گونه ایران در خوشه B قرار گرفت. براساس توپولوژی فیلوژنی و هم‌چنین بررسی فاصله ژنتیکی، سوسن چلچراغ بیش‌ترین شباهت را با سه گونه *Lilium pyrenaicum*، *L. ciliatum* و *L. candidum* با ارزش پایداری ۹۷ درصد داشت. در بررسی بیوانفورماتیکی، در هم‌ردیفی این پروتئین درجه بالایی از حفاظت‌شدگی مشاهده شد. سوسن چلچراغ در جایگاه ۲۷۱ دارای اسید آمینه آرژنینین هست و از این نظر فقط با سه گونه *L. pyrenaicum*، *L. ciliatum* و *L. candidum* مشابه بود. سایر گونه‌ها در این جایگاه دارای اسید آمینه لیزین بودند. تفاوت‌های زیادی در جایگاه‌های اسیدآمینه‌ای مختلف از جمله در موقعیت‌های ۳۱۷، ۳۴۶، ۳۶۳،

* مسئول مکاتبه: maderi@ut.ac.ir

۴۱۷ در گونه‌ها مشاهده شد. بررسی تشابه ساختار دوم توالی *matk* در سوسن چلچراغ نشان داد که این پروتئین دارای ۲۳۳ اسید آمینه در مارپیچ α (۴۵/۵۱ درصد)، ۲۴ اسید آمینه در پیچ β (۴/۶۹ درصد) و ۱۵۶ اسید آمینه در مارپیچ تصادفی (۳۰/۷۴ درصد) است.

نتیجه‌گیری: پژوهش حاضر برای اولین بار این ژن را در سوسن چلچراغ مورد تجزیه و تحلیل قرار داد و ساختار دوبعدی و موقعیت α هلیکس‌ها و صفحات β مشخص شد. همچنین مدل ساختار سوم این پروتئین برای نخستین بار ارائه شد. بر اساس توپولوژی فیلوژنی و همچنین بررسی فاصله ژنتیکی، سوسن چلچراغ بیش‌ترین شباهت را با سه گونه *L. pyrenaicum*، *L. candidum* و *L. ciliatum* داشت. به‌طور کلی، در این پژوهش قرابت سوسن چلچراغ از نظر مولکولی (فیلوژنی، ساختار پروتئین) با دیگر گونه‌های لیلیوم بررسی شد.

واژه‌های کلیدی: فیلوژنی، گونه‌های لیلیوم، DNA بارکدینگ، *matk in silico*

مقدمه

از جمله ویژگی‌های این گونه می‌توان به ساقه گل‌دهنده قوی، تعداد گل زیاد و رایحه مطبوع اشاره کرد. این ویژگی‌های کم نظیر باعث می‌شود این گونه از لیلیوم از دیدگاه باغبانی بسیار ارزشمند باشد. علی‌رغم ویژگی‌های ارزشمند این گونه در معرض خطر، تاکنون در هیچ مطالعه‌ای، جایگاه و ارتباط فیلوژنی *L. ledebourii* با سایر گونه‌های این خانواده مورد بررسی قرار نگرفته است.

امروزه سعی می‌شود جهت بررسی روابط تکاملی و فیلوژنی بجای اطلاعات ریخت‌شناسی، از اطلاعات حاصل از توالی DNA و پروتئین‌ها استفاده شود. هر نوکلئوتید یا اسید آمینه می‌تواند به‌عنوان یک خصوصیت جهت بررسی روابط فیلوژنی در نظر گرفته شود. در مقایسه با اطلاعات ریخت‌شناسی، امتیاز دادن به این نوع اطلاعات ژنتیکی هم ساده‌تر است و هم اطلاعات بیش‌تری در اختیار قرار می‌دهند. برای مثال با مقایسه دو توالی ۱۰۰۰ نوکلئوتیدی، ۱۰۰۰ نقطه قابل مقایسه در دسترس است که باعث می‌شود نسبت به اطلاعات ریخت‌شناسی به حجم بیش‌تری از اطلاعات دسترسی پیدا کنیم (۸).

DNA بارکدینگ یک روش شناسایی گونه با استفاده از یک ناحیه کوتاهی از DNA است. در

راسته لیلیالس ۱۰ خانواده، ۶۴ جنس و حدود ۱۵۰۰ گونه را شامل می‌شود (۲ و ۱۸). بزرگ‌ترین خانواده این راسته، لیلیاسه می‌باشد که ۶۱۰ گونه را در خود جای داده است (۱۸). از دیدگاه باغبانی جنس لیلیوم با حدود ۱۰۰ گونه، مهم‌ترین جنس این خانواده است که چندین بار مورد طبقه‌بندی قرار گرفته است. از معتبرترین طبقه‌بندی‌ها می‌توان به طبقه‌بندی این جنس بر اساس ۱۵ ویژگی اشاره نمود که به هفت بخش تقسیم شده است: مارتاگون^۱، پسودولیریوم^۲، لیریوتیپوس^۳، آرچلیریون^۴، سینومارتاگون^۵، لئوکولیریون^۶ و دائورولیریون^۷ (۹). طبقه‌بندی خانواده لیلیاسه همواره با مطالعات ریخت‌شناسی، کاربولوژیکی و فیلوژنی مولکولی پیگیری می‌شود که رویکردهای جدید فیلوژنی مولکولی باعث تغییراتی اساسی در این طبقه‌بندی شده است (۱۱).

سوسن چلچراغ از گونه‌های ارزشمند بومی ایران است که به‌شدت در معرض خطر انقراض قرار دارد.

- 1- Martagon
- 2- Pseudolirium
- 3- Liriotypus
- 4- Archelirion
- 5- Sinomartagon
- 6- Leucolirion
- 7- Daurolirion

مواد و روش‌ها

به منظور استخراج DNA، نمونه برگ‌های سوسن چلچراغ از داماش گیلان جمع‌آوری شد. استخراج DNA از برگ با استفاده از کیت استخراج Qiagen DNEasy انجام شد. کیفیت و کمیت DNA با استفاده از نانودراپ ارزیابی و در نهایت با ژل آگاروز ۱ درصد بررسی و در نهایت نمونه برای توالی‌یابی ارسال شد. توالی‌یابی به صورت Real-Time (SMRT) با استفاده از پلتفرم نسل جدید PacBio انجام و توالی کامل ژن *matK* با استفاده از ابزارهای بیوانفورماتیک استخراج شد. جهت حصول اطمینان از نتایج دریافتی، بررسی، مقایسه و تعیین هویت توالی با استفاده از اطلاعات موجود در NCBI و عملیات بلاست انجام پذیرفت. توالی کامل ژن *matK* در پایگاه NCBI به شماره دسترسی MN557236 (انتشار نیافته) ثبت شد.

به منظور بررسی روابط فیلوژنتیک، توالی *matK* در ۴۱ گونه دیگر لیلیوم از پایگاه اطلاعات ژنی NCBI (<https://www.ncbi.nlm.nih.gov>) به دست آمد. بهترین مدل جهت رسم درخت فیلوژنی با استفاده از نرم‌افزار Mega 7 انتخاب شد. در مرحله بعدی توالی‌های به دست آمده با روش ClustalW و با استفاده از نرم‌افزار Mega7 با تنظیمات پیش فرض هم‌ردیف^۱ شدند. روش حداکثر درست‌نمایی^۲ جهت رسم درخت فیلوژنی به کار گرفته شد. هم‌چنین یک گونه از جنس *Tricyrtis* به عنوان اوت‌گروپ^۳ در نظر گرفته شد. فاصله ژنتیکی به روش K2P با استفاده از نرم‌افزار Mega7 محاسبه شد. به منظور بررسی احتمال انواع جایگزینی‌های انتقالی^۴ و متقاطع^۵ از ماتریس

حقیقت بارکدهای DNA حاوی تفاوت‌هایی هستند که می‌توانند به عنوان ویژگی تمایز گونه‌ها مطرح باشند. در طول دهه گذشته این روش به طور فزاینده‌ای برای تسهیل مطالعات تنوع زیستی به کار گرفته شده است (۱). تعدادی از مناطق ژنی هسته‌ای و پلاستییدی به عنوان بارکدی بالقوه برای گیاهان پیشنهاد شده‌اند و جستجو برای یک بارکد DNA گیاهی روی ژن‌های ژنوم کلروپلاستی مانند *rbcL* *psbK-psbI* *nhdJ* *matK* *atpF-atpH* *trnH-psbA* متمرکز شده است (۱ و ۱۹).

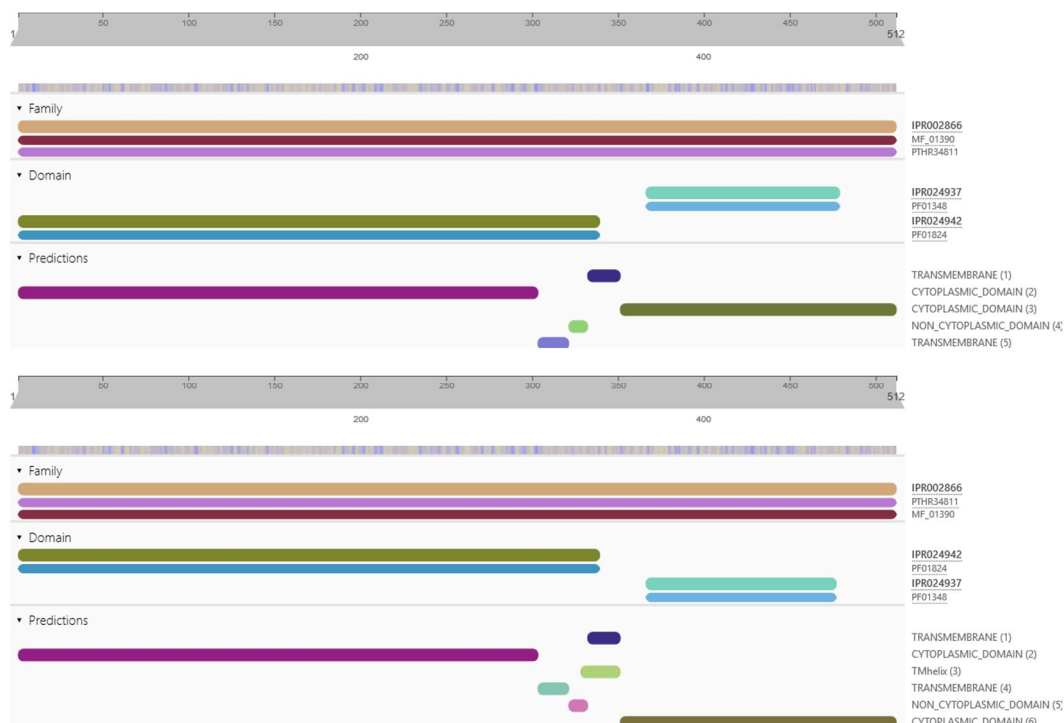
matK یکی از کاراترین ژن‌های بارکد گیاهی است. این ژن به طور تقریبی ۱۵۷۰ جفت باز طول دارد. ژن *matK* در درون اینترون *trnK* در ژنوم کلروپلاست بیش‌تر گیاهان به جز برخی از سرخس‌ها قرار دارد (۱۹). سرعت تکاملی بالای ژن *matK* باعث شده این ژن به عنوان یکی از مناسب‌ترین و کارآمدترین ژن‌های کلروپلاستی برای حل روابط فیلوژنی و تکاملی در سطوح مختلف به‌ویژه در سطح گونه و جنس همواره مطرح باشد (۱۷). این ویژگی منحصر به فرد ژن *matK* باعث شده این ژن به تنهایی یا همراه با سایر توالی‌ها به عنوان DNA بارکدینگ در نظر گرفته شود (۱۹). DNA بارکدینگ کار متخصصین رده‌بندی گیاهی را آسان‌تر می‌کند و هم‌چنین امکان تشخیص و طبقه‌بندی صحیح گونه‌ها را برای افراد غیرمتخصص نیز فراهم می‌آورد (۱۵).

این پژوهش با هدف بررسی *in silico* ژن *matK* به منظور تجزیه و تحلیل فیلوژنتیک، مشخص کردن عملکرد، ساختمان دوم و سوم پروتئین و بررسی تحلیل‌های بیوانفورماتیک در راستای تعیین مشخصات این ژن در گونه‌های مختلف لیلیوم به‌ویژه *L. ledebourii* انجام شد.

- 1- Multiple sequence alignments
- 2- Maximum likelihood
- 3- Outgroup
- 4- Transitional substitutions
- 5- Transversional substitutions

مشاهده شد و سایر گونه‌ها در این جایگاه دارای آرژنین بودند (شکل ۱).

دارای ایزولوسین بودند. به جز *L. lancifolium*، در موقعیت ۳۶۳ تمامی گونه‌ها دارای اسپارات بودند. در موقعیت ۴۱۷ نیز فقط در *L. concolor* ترئونین



شکل ۲- کارکرد موجود در پروتئین‌های matK و دامین‌ها در *L. ledebourii* (بالا) و *L. regale* (پایین).

Fig. 2. Domains and protein function of matK in *L. ledebourii* (Top) and *L. regale* (Below).

IPR024942, PF01824: The function of this regions is unknown. IPR024937: Group II introns use intron-encoded reverse transcriptase, maturase and DNA endonuclease activities for site-specific insertion into DNA. PF01348: This entry represents a sequence region known as domain X, which is conserved in group II introns and is required for maturase function.

بیشترین مقدار جابجایی متقاطع مربوط به تیمین با آدنین و تیمین با گوانین با مقدار ۷/۵۶ بود. نتایج نشان‌دهنده بیشتر بودن جابجایی انتقالی نسبت به جابجایی متقاطع است (جدول ۱). پژوهشگران دیگری نیز این شیوه تغییر را گزارش داده‌اند (۳). این تغییر به احتمال زیاد به علت متیلاسیون سیتوزین رخ می‌دهد (۱۶).

بررسی نتایج حاصل از درصد انواع جابه‌جایی‌های متقاطع و انتقالی در نوکلئوتیدهای ژن *matK* نشان داد که بیشترین میزان جایگزینی مربوط به بازهای پیریمیدینی است، به طوری که این مقدار برای تبدیل سیتوزین به تیمین ۲۲/۷۶ درصد بود. این مقدار جایگزینی برای بازهای پورینی، ۱۸/۷۵ درصد و مربوط به تبدیل گوانین به آدنین بود. هم‌چنین

جدول ۱- برآورد میزان حداکثر درست‌نمایی مرکب الگوی جایگزینی نوکلئوتیدی برای ژن *matk*

Table 1. Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution for *matk* gene.

	A	T	C	G
A	-	7.56	3.16	8.95
T	6.3	-	9.5	3.01
C	6.3	22.76	-	3.01
G	18.75	7.56	3.16	-

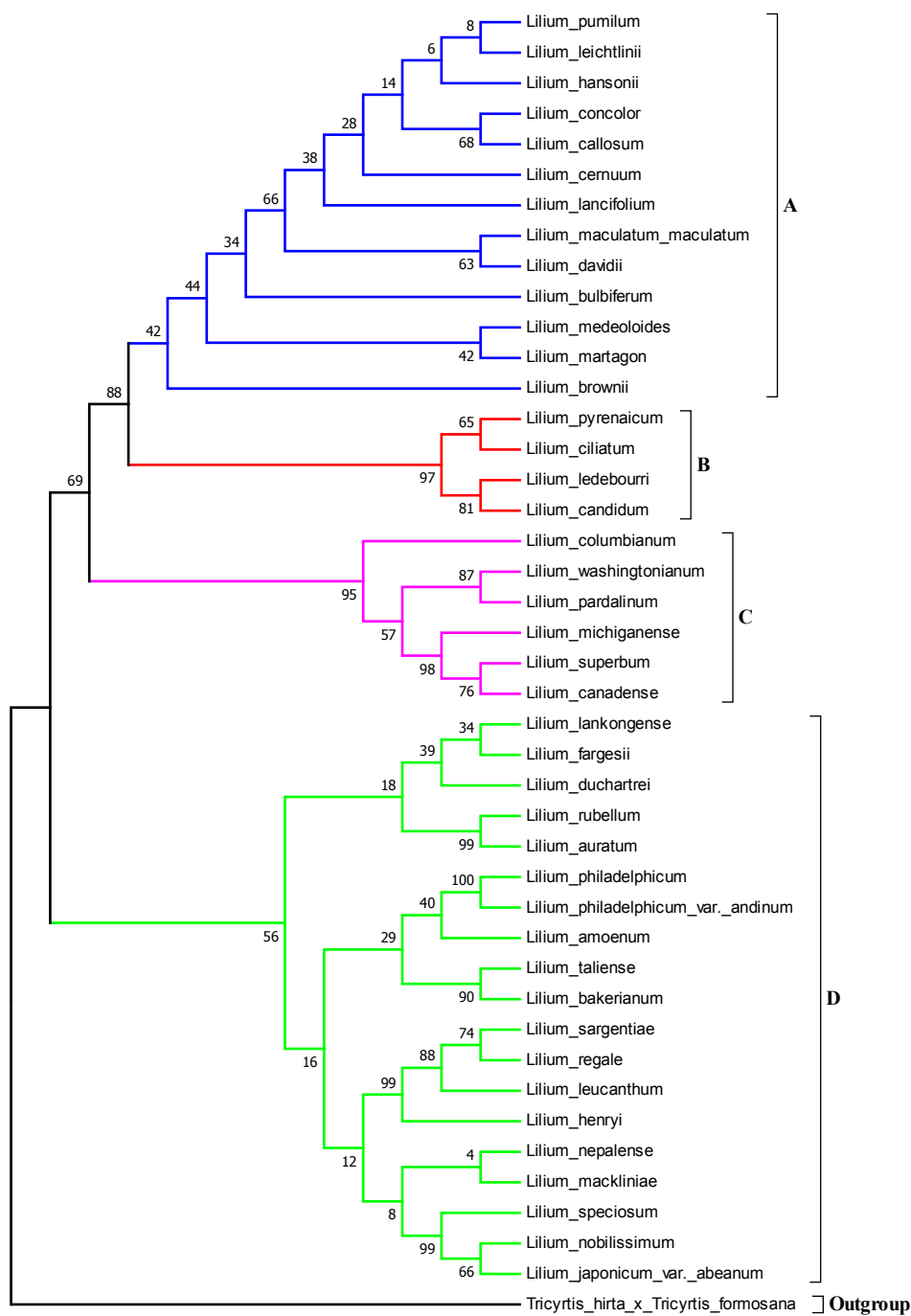
جزو لیلیوم‌های سینومارتاگون طبقه‌بندی می‌شوند (شکل ۳). نواحی پراکنش تمامی این گونه‌ها آسیای شرقی است. پژوهشگران دیگری هم نزدیکی گونه‌های با ناحیه پراکنش مشترک را نشان داده‌اند (۷).

سوسن چلچراغ به همراه سه گونه *L. pyrenaicum*، *L. candidum* و *L. ciliatum* با حمایت ۹۷ درصد، خوشه B را تشکیل دادند. نتایج ماتریس فاصله هم نشان داد که همین سه گونه نزدیک‌ترین فاصله ژنتیکی را با سوسن چلچراغ دارند. هم‌چنین با بررسی فاصله ژنتیکی گونه‌ها به روش K₂P بیش‌ترین فاصله ژنتیکی بین *L. regale* با گونه‌های *L. superbum*، *L. pumilum* و *L. concolor* مشاهده شد (جدول ۲).

به‌طورکلی در بین گونه‌های مطالعه حاضر، *L. japonicum* و *L. speciosum* با ارزش بوت استرپ ۹۹ درصد، *L. rubellum* و *L. auratum* با ارزش بوت استرپ ۹۹ درصد و *L. washingtonianum* و *L. pardalinum* با ارزش بوت استرپ ۸۸ درصد، طبق بالاترین امتیاز بوت استرپ، نزدیک‌ترین گونه‌ها به هم بودند (شکل ۳). این نتایج در مطالعات دیگر بر اساس ژن‌های دیگر هم مشاهده شده است؛ از جمله هایاشی و کاوانو (۲۰۰۰) در مطالعه خود ارتباط نزدیک *L. japonicum* و *L. speciosum*، *L. washingtonianum* و *L. pardalinum* را تأیید کردند (۷).

طول توالی ژن *matk* در گونه‌های مختلف لیلیوم بین ۱۵۲۱ تا ۱۵۴۹ جفت باز متغیر و در سوسن چلچراغ، این مقدار ۱۵۳۶ بود. طبق توپولوژی درخت فیلوژنی و امتیاز بوت استرپ، گونه‌های لیلیوم به چهار گروه تقسیم شدند (شکل ۳). نتایج این مطالعه با مطالعات دیگر پژوهشگران هم‌خوانی دارد (۱۳) و (۱۲). خوشه D بزرگ‌ترین خوشه و شامل *L. speciosum*، *L. japonicum*، *L. nobilissimum*، *L. henryi*، *L. mackliniae*، *L. nepalense*، *L. regale*، *L. sargentiae*، *L. leucanthum*، *L. amoenum*، *L. bakerianum*، *L. taliense*، *L. auratum*، *L. rubellum*، *L. philadelphicum*، *L. fargesii* و *L. lankongense*، *L. duchartrei* بود.

گونه‌های *L. pumilum*، *L. concolor*، *L. hansonii*، *L. leichtlinii*، *L. callosum*، *L. maculatum*، *L. lancifolium*، *L. cernuum*، *L. medeoloides*، *L. bulbiferum*، *L. davidii* و *L. martagon* با حمایت ۴۲ درصد خوشه مستقل A را تشکیل دادند. در این گروه، زیر گروهی با حمایت ۶۶ درصد، متشکل از *L. pumilum*، *L. leichtlinii*، *L. callosum*، *L. concolor*، *L. lancifolium*، *L. cernuum*، *L. hansonii*، *L. davidii*، *L. maculatum* وجود دارد که طبق مطالعات ریخت‌شناسی، همگی به‌جز *L. hansonii*

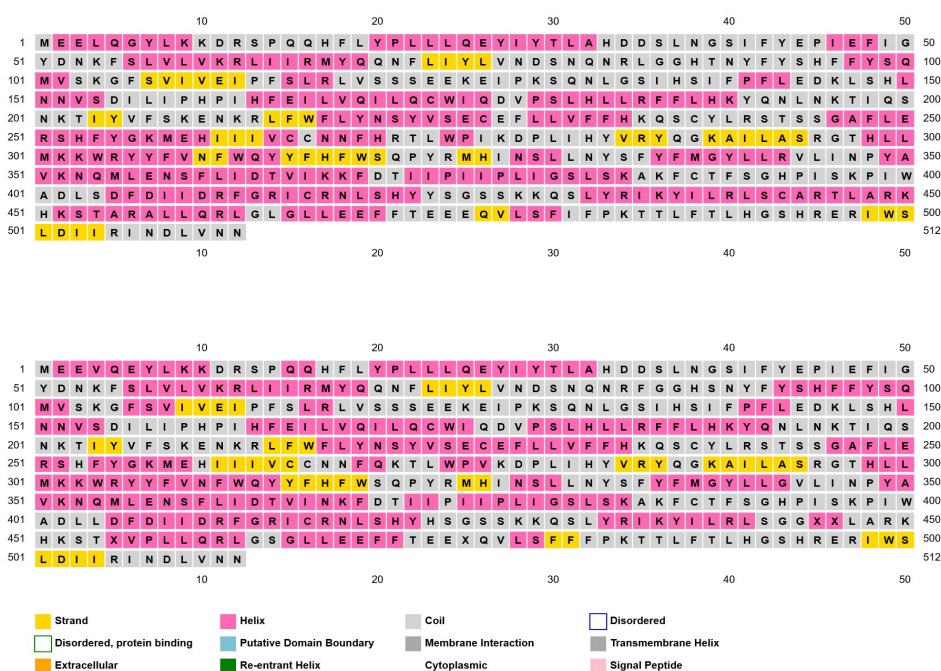


شکل ۳- درخت فیلوژنی بر اساس توالی *matK* به روش حداکثر درستنمایی با پشتوانه ۱۰۰۰ تکرار. اوت‌گروپ متعلق به جنس دیگری از خانواده لیلیاسه می‌باشد.

Fig. 3. The phylogenetic tree based on *matK* sequence according to the maximum likelihood with the support of 1000 replications.

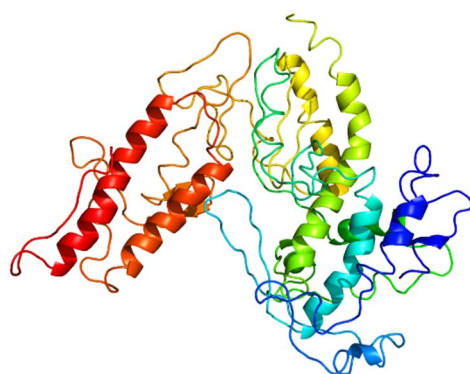
و *L. nobilissimum* *L. ciliatum* و *L. japonicum*
L. canadense و *L. rubellum* *L. ciliatum*
Lilium michiganense و *L. rubellum*
L. ciliatum *L. pyrenaicum* و *L. pyrenaicum*
L. superbum و *L. rubellum* *L. speciosum* و
 مشاهده شد (شکل ۶).

ساختار سه‌بعدی پروتئین اطلاعات مهمی در
 رابطه با درک بهتر عملکرد پروتئین، اثرات متقابل آن
 با اجزای دیگر (لیگاندها، پروتئین، DNA) و هم‌چنین
 جهت فهم اثرات فنوتیپیکی جهش‌ها در اختیار ما
 قرار می‌دهد. میانگین درصد تشابه بین گونه‌های مختلف
 مورد مطالعه در این پژوهش ۸۹/۰۴ بود. هم‌چنین
 کم‌ترین میزان تشابه به مقدار ۴۷/۰۳ درصد بین



شکل ۵- توالی پروتئینی matk در *L. ledebourii* (بالا) و *L. regale* (پایین).

Fig. 5. Protein sequence of matk in *L. ledebourii* (Top) and *L. regale* (Below).



شکل ۶- ساختار سه‌بعدی پروتئین matk در لیلیوم لدبوری.

Fig. 6. Tertiary structure of matk protein in *L. ledebourii*.

با *L. candidum* و *L. ciliatum* *L. pyrenaicum* ارزش پایداری ۹۷ درصد داشت. در این پژوهش مجموعه تحلیل‌های بیوانفورماتیکی را برای این ژن کلیدی در گونه‌های مختلف لیلیوم به‌ویژه سوسن چلچراغ برای توضیح ارتباط ساختاری، عملکردی و فیلوژنتیکی به‌کار گرفته شد.

نتیجه‌گیری کلی

پژوهش حاضر برای اولین بار این ژن را در سوسن چلچراغ مورد تجزیه و تحلیل قرار داد و ساختار دوبعدی و موقعیت α هلیکس‌ها و صفحات β مشخص شد. هم‌چنین مدل ساختار سوم این پروتئین برای نخستین بار ارائه شد. بر اساس توپولوژی فیلوژنی و هم‌چنین بررسی فاصله ژنتیکی به روش K_2P ، سوسن چلچراغ بیش‌ترین شباهت را با سه گونه

منابع

1. Amandita, F.Y., Rembold, K., Vornam, B., Rahayu, S., Siregar, I.Z., Kreft, H. and Finkeldey, R. 2019. DNA barcoding of flowering plants in Sumatra, Indonesia. *Ecol. Evol.* 9: 4. 1858-1868.
2. APG. 2009. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot. J. Linn. Soc.* 161: 105-121.
3. Bakker, F.T., Culham, A., Gomez-Martinez, R., Carvalho, J., Compton, J., Dawtrey, R. and Gibby, M. 2000. Patterns of nucleotide substitution in angiosperm cpDNA trnL (UAA)-trnF (GAA) regions. *Mol. Biol. Evol.* 17: 8. 1146-1155.
4. Combet, C., Blanchet, C., Geourjon, C. and Deleage, G. 2000. NPS@: network protein sequence analysis. *Trends Biochem. Sci.* 25: 3. 147-150.
5. Cozzetto, D., Minneci, F., Curren, H. and Jones, D.T. 2016. FFPred 3: feature-based function prediction for all Gene Ontology domains. *Sci. Rep.* 6: 31865.
6. Darabi, M., Masoudi-Nejad, A. and Nemat-Zadeh, G. 2012. Bioinformatics study of the 3-hydroxy-3-methylglutaryl-coenzyme A reductase (HMGR) gene in Gramineae. *Mol. Boil. Rep.* 39: 9. 8925-8935.
7. Hayashi, K. and Kawano, S. 2000. Molecular systematics of *Lilium* and allied genera (Liliaceae): phylogenetic relationships among *Lilium* and related genera based on the rbcL and *matK* gene sequence data. *Plant Spec. Biol.* 15: 1. 73-93.
8. Hollingsworth, P.M., Graham, S.W. and Little, D.P. 2011. Choosing and using a plant DNA barcode. *PloS One.* 6: 5. e19254.
9. İkinci, N. 2011. Molecular phylogeny and divergence times estimates of *Lilium* section *Liriotypus* (Liliaceae) based on plastid and nuclear ribosomal ITS DNA sequence data. *Turk. J. Bot.* 35: 4. 319-330.
10. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Prot.* 10: 6. 845.
11. Kim, J.S. and Kim, J.H. 2018. Updated molecular phylogenetic analysis, dating and biogeographical history of the lily family (Liliaceae: Liliales). *Bot. J. Linn. Soc.* 187: 4. 579-593.
12. Lee, C.S., Kim, S.C., Yeau, S.H. and Lee, N.S. 2011. Major lineages of the genus *Lilium* (Liliaceae) based on nrDNA ITS sequences, with special emphasis on the Korean species. *J. Plant Biol.* 54: 3. 159-171.
13. Muratovic, E., Hidalgo, O., Garnatje, T. and Siljak-Yakovlev, S. 2010. Molecular phylogeny and genome size in European lilies (Genus *Lilium*, Liliaceae). *Adv. Sci. Lett.* 3: 2. 180-189.
14. Nuka, G., Fraser, M., Mitchell, A., Potter, S., Yong, S.Y. and Finn, R.D. 2017. InterProScan: Protein sequence analysis and classification. *F1000 Research.* 6.

15. Pang, X., Song, J., Zhu, Y., Xu, H., Huang, L. and Chen, S. 2011. Applying plant DNA barcodes for Rosaceae species identification. *Cladistics*. 27: 2. 165-170.
16. Picoult-Newberg, L., Ideker, T.E., Pohl, M.G., Taylor, S.L., Donaldson, M.A., Nickerson, D.A. and Boyce-Jacino, M. 1999. Mining SNPs from EST databases. *Genom. Res.* 9: 2. 167-174.
17. Sanders, E.R., Karol, K.G. and McCourt, R.M. 2003. Occurrence of *matK* in a *trnK* group II intron in charophyte green algae and phylogeny of the Characeae. *Am. J. Bot.* 90: 4. 628-633.
18. Stevens, P.S. 2015. Angiosperm phylogeny website. <http://www.mobot.org/MOBOT/research/APweb/welcome.html>.
19. Yu, J., Xue, J.H. and Zhou, S.L. 2011. New universal *matk* primers for DNA barcoding angiosperms. *J. Syst. Evol.* 49: 3. 176-181.

Arc