



پیش‌بینی بیوانفورماتیکی توالی‌های حفاظت شده miRNA و مطالعه

روند ایجاد و از بین رفتن آنها در بقولات

بهزاد حاجی‌اقراری^{۱*}، مجاهد کمالی‌زاده^۲

۱-استاد، دانشگاه جهرم، گروه مهندسی تولید و ژنتیک گیاهی، جهرم، ایران

۲-استاد، دانشگاه جهرم، گروه مهندسی تولید و ژنتیک گیاهی، جهرم، ایران

*صندوق پستی ۷۴۱۳۵۱۱۱، جهرم، ایران
bheghrari@yahoo.com

پذیرش: ۱۴۰۱/۰۳/۲۲

دریافت: ۱۴۰۰/۰۶/۱۱

چکیده

خانواده بقولات یکی از مهمترین خانواده‌های گیاهان است. در این پژوهش توالی‌های پیش‌ساز miRNA ها در هفت گونه از بقولات شامل *Lotus japonicus*, *Glycine soja*, *Glycine max*, *Arachis hypogaea*, *Phaseolus vulgaris*, *Medicago truncatula* و *Vigna unguiculata* به شیوه جستجوی توالی‌های پیش‌ساز حامل توالی‌های همولوگ miRNA ای بالغ در میان توالی‌های نسخه‌برداری شونده و با در نظر گرفتن مشخصات اختصاصی توالی و ساختمان دوم توالی پیش‌ساز در miRNA های شناسایی شده، پیش‌بینی و مشخصات آنها مورد توجه قرار گرفت. همچنین جایگاه ژنومی توالی‌های پیش‌ساز تعیین شد. در این مطالعه، ۴۱۴ توالی پیش‌ساز miRNA متعلق به ۱۳۰ خانواده در هفت گونه مذکور که مشخصات توالی و ساختار دوم آنها با مشخصات توالی‌های پیش‌ساز miRNA گیاهی همخوانی داشت، پیش‌بینی و مشخصات آنها گزارش شد. روابط بین گونه‌ها براساس روند ایجاد و از بین رفتن خانواده‌های miRNA ای در هرگونه براساس بی‌شینه پارسیمونی به شیوه دولو به دست آمد و از این طریق درخت فیلوژنتیکی ارتباط گونه‌ها ترسیم شد. خانواده‌های miRNA به وجود آمده و از بین رفته در گونه‌های مورد مطالعه در طول پیدایش و تکامل گونه‌ها مشخص و گزارش شد. نتایج نشان داد روند ایجاد خانواده‌های miRNA در این خانواده اخیراً از شدت زیادی برخوردار بوده است. روابط بین گونه‌ها براساس روند ایجاد و از بین رفتن خانواده‌های miRNA به شیوه مدل بی‌شینه پارسیمونی همخوانی با روابط فیلوژنی ندارد. در این مطالعه تعدادی از خانواده‌های miRNA به عنوان خانواده‌های اختصاصی گونه و اختصاصی گروه شناسایی شد که امکان استفاده از آنها به عنوان شناسه گر گونه/گروه وجود دارد.

واژگان کلیدی: miRNA، خانواده بقولات، بی‌شینه پارسیمونی دولو، جستجوی توالی‌های همولوگ miRNA

۱- مقدمه

تنظیم بیان ژن شامل فرایندهایی است که توسط سلول برای تنظیم (افزایش یا کاهش) بیان یک ژن مشخص یا گروهی از ژن‌های مرتبط با هم انجام می‌شود. تنظیم بیان ژن در هر مرحله از فرایند بیان ژن شامل پیش از رونویسی، شروع رونویسی، رونویسی، پیرایش RNA^۱، ترجمه و تغییرات پس از ترجمه یک ژن انجام می‌شود [۱]. اغلب در یک شبکه تنظیم بیان ژن، برخی ژن‌ها قادر به کنترل عملکرد برخی دیگر می‌باشند. تنظیم و کنترل بیان ژن امری حیاتی در سلول‌های موجودات زنده مختلف به شمار می‌رود و به شیوه‌های مختلف انجام می‌شود. این فرایند سبب تولید فراورده‌های ژنی مورد نیاز در زمان مورد نیاز و به میزان مورد نیاز می‌شود. بنابراین قابلیت انعطاف پذیری و سازگاری با شرایط مختلف محیط برای موجود زنده امکان پذیر می‌شود. در سلول‌های یوکاریوتی مکانیسم‌های مختلفی برای تنظیم بیان ژن وجود دارد و تنظیم بیان ژن توسط miRNA ها یکی از مکانیسم‌های مهم و متداول است.

miRNA ها توالی‌های کوچک به طول ۱۸-۲۴ نوکلئوتید می‌باشند که در تنظیم بیان ژن در موجودات زنده یوکاریوتی در مرحله حین نسخه‌برداری و پس از نسخه‌برداری نقش دارند [۲]. miRNA ها از جایگاه‌های اختصاصی بر روی ژنوم رونویسی و طی فرایند بیوشیمیایی مشخصی در سلول و با مشارکت آنزیم‌های متعدد توالی آنها آزاد شده و به کمک سیستم آنزیمی پیچیده در تنظیم بیان ژن در حین نسخه‌برداری یا پس از نسخه‌برداری مشارکت می‌کنند [۳-۵]. این توالی‌ها با تاثیر بر روی تظاهر ژن‌ها در شبکه‌های بیان ژنی تاثیر خود را در رشد و نمو، تمایز و پاسخ گیاه به محرک‌های زنده و غیر زنده محیطی اعمال می‌کنند [۶]. همچنین، این توالی‌ها در دفاع از گیاه در برابر عوامل ژنتیکی مهاجم مانند

ویروس‌ها و توالی‌های ترانسپوزانی نقش دارند [۷ و ۸]. توالی بسیاری از پیش سازهای miRNA که حامل miRNA بالغ هستند، در برخی گونه‌های گیاهی شناسایی شده و در پایگاه ثبت توالی‌های miRNA میریس [۷ و ۹-۱۱] به ثبت رسیده‌اند. با این حال امروزه تلاش‌ها برای شناسایی و ثبت توالی‌های miRNA در یوکاریوت‌ها بخصوص در گیاهان همچنان ادامه دارد [۱۶-۱۲]. با شناسایی و ثبت کامل توالی‌های miRNA دانشمندان قادر خواهند بود شبکه بیان و تنظیم ژن را در هر گونه به طور کامل شناسایی و تکمیل کنند.

شیوه‌های آزمایشگاهی مختلفی برای شناسایی miRNA ها و اهداف ژنی آنها توسعه داده شده است [۴ و ۱۷-۲۰]. با استفاده از همین روش‌ها توالی تعداد زیادی از miRNA ها شناسایی و در پایگاه‌های ثبت توالی miRNA به ثبت رسیده‌اند. روش‌های آزمایشگاهی به دلیل نیاز به هزینه و زمان زیاد برای شناسایی miRNA های بیان شده در شرایط و استرس‌های مختلف محیطی در عمل کارایی مناسبی برای شناسایی miRNA ها ندارند [۲۱]. با توجه به مشکلات روش‌های آزمایشگاهی تشخیص جایگاه ژنی miRNA، امروزه استفاده از روش‌های جایگزین مانند روش‌های بیوانفورماتیکی تشخیص جایگاه ژنی miRNA متداول شده است [۲۲].

خانواده بقولات؛ یکی از مهمترین خانواده‌های گیاهان است. این خانواده شامل گونه‌های مهم غذایی و صنعتی می‌باشد. هدف از اجرای این پژوهش شناسایی و پیش‌بینی ژن‌های miRNA حفاظت شده در مجموعه توالی‌های نسخه‌برداری شونده گونه‌های خانواده بقولات و همچنین مطالعه روند ایجاد و از بین رفتن ژن‌های miRNA در گونه‌های این خانواده می‌باشد. شناسایی خانواده‌های ژنی miRNA در گونه‌های مورد مطالعه، گوناگونی ژن‌های miRNA در گونه‌ها و تعداد پارالوگ‌های آنها، منشاء

^۳ . miRBase

^۱ . Leguminosae

^۱ . RNA splicing

^۲ . miRNAs

ژن های miRNA موجود در گونه های مختلف، توزیع این ژن ها بر روی کروموزوم گونه ها، مدل روند تکاملی در گیاهان بر اساس روند ایجاد و از بین رفتن خانواده های miRNA، روند ایجاد و از بین رفتن ژن های miRNA در طول زمان تکامل آنها، موضوع مطالعه در این پژوهش است. این پژوهش بر اساس آنالیز و تحلیل مجموعه داده های چندین منبع پایگاه داده، همچون پایگاه داده میریس و پایگاه داده جامع بانک ژن و با استفاده از ابزارهای کامپیوتری آنالیز طرح ریزی و تحلیل شده است.

۲- مواد و روش ها

۲-۱. جستجوی miRNA های گیاهی بالقوه در گونه های

گیاهی خانواده بقولات

برای جستجو و استخراج توالی های پیش ساز miRNA بالغ همولوگ با توالی های شناسایی شده miRNA گیاهی، توالی های نسخه برداری شونده^۱ ثبت شده برای گونه های *Glycine soja*, *Glycine max*, *Arachis hypogaea*, *Phaseolus*, *Medicago truncatula*, *Lotus japonicus* و *vulgaris* از پایگاه داده توالی های نسخه برداری شونده بانک ژن^۲ (قابل دسترس در تار نامی: <http://www.ncbi.nlm.nih.gov/nucest>) در قالب فایل های مجزا جمع آوری شد. همچنین، توالی های پیش ساز و بالغ miRNA ثبت شده برای گونه های گیاهی از پایگاه داده میریس استخراج و در قالب فایل جداگانه ذخیره شد. توالی های نسخه برداری شونده تکراری ثبت شده برای هر گونه با استفاده از ابزار کامپیوتری CodonCode Aligner 6,0,2 شناسایی و حذف شد. همچنین، توالی های تکراری پیش ساز و بالغ miRNA ها نیز با این شیوه از فایل های ذخیره شده حذف شد. توالی های همولوگ با هر کدام از توالی های miRNA بالغ گونه های گیاهی شناسایی شده (که حداکثر ۴ اسید نوکلئیک

جایگزین در توالی آنها مشاهده می شد و دارای ابی ولیو کمتر یا برابر ۱۰ در طول همترازی بودند) بر روی توالی های نسخه برداری شونده هر گونه، توسط الگوریتم بلاست نوکلئوتید^۳ [۲۱] مشخص شد. این توالی ها به صورت جداگانه ذخیره شدند. مکمل توالی همولوگ توالی شناسایی شده بین توالی های استخراج شده که حداکثر ۴ نوکلئوتید جایگزین در توالی آنها مشاهده می شد و فاصله بین این دو توالی حداکثر ۳۰۰ توالی بود، جستجو و توالی حد فاصل بین توالی های miRNA و مکمل آن به اضافه دو نوکلئوتید از سمت ۳' از توالی مذکور استخراج شد. برنامه کامپیوتری این مرحله در بستر نرم افزاری متلب تهیه و استفاده شد (فایل ضمیمه ۱). برای حذف توالی های ترجمه شونده به پروتئین، توالی های پروتئین های گیاهی از پایگاه داده یونی پرات/ سوئیس پرات^۴ دریافت شد و توالی های مشابه توالی های پروتئینی در مجموع توالی های دارای همولوگ miRNA با انجام بلاست ایکس^۵ که سه قاب ترجمه شده هر کدام از توالی های مجموعه توالی های دارای همولوگ miRNA را با توالی های پروتئینی اخذ شده از پایگاه داده های مربوطه مقایسه می کند، شناسایی و متمایز شد (توالی هایی با ابی ولیو کمتر یا برابر ۲۰-۱۰ در طول همترازی، انتخاب شدند). در این مرحله، توالی های به دست آمده توالی هایی غیر ترجمه شونده به پروتئین و حامل توالی های همولوگ miRNA می باشند. سپس، توالی های به دست آمده با توالی های ذخیره شده در پایگاه داده RFEM که توالی های خانواده های مختلف RNA را در خود ذخیره می کند، با انجام بلاست نوکلئوتید مقایسه و توالی های مشابه با سایر خانواده های RNA از باقی توالی ها حذف شد (توالی هایی با ابی ولیو کمتر یا برابر با ۱۰^{-۸} در طول همترازی، انتخاب شدند). در مرحله بعد ساختار دوم توالی ها با استفاده از

^۴ Uniprot/Swissprot

^۵ BlastX

^۱ Expressed sequence tags (ESTs)

^۲ NCBI

^۳ BlastN

روش کمترین انرژی آزاد در ساختار دوم RNA توسط نرم‌افزار RNA فولد^۱ ترسیم و مشخصات ساختار دوم RNA ها مشخص شد. توالی‌هایی که مشخصات ساختار دوم آنها با مشخصات کلی ساختار دوم در miRNA ها مطابقت داشت [۱۱ و ۱۵]، به‌عنوان توالی پیش‌ساز توالی miRNA هم خانواده با miRNA همولوگ انتخاب و معرفی شدند. در مورد توالی‌های پیش‌ساز تکراری که از توالی‌های رونویسی شونده مختلف به دست آمد، یکی از توالی‌ها مورد توجه قرار گرفت و باقی توالی‌ها حذف شد. توالی‌های شناسایی شده با توجه به خانواده‌هایی که توالی‌های miRNA بالغ‌شان به آنها شباهت داشته و پارالوگ‌های شناسایی شده و ثبت شده برای هر خانواده miRNA و بازویی که توالی miRNA بالغ در ساختار دوم آن جای می‌گیرد، تعداد کپی‌های miRNA بالغ در توالی مختلف پیش‌سازها، پیش‌سازهای شناسایی و ثبت شده برای هر خانواده و بر اساس قانون نامگذاری miRNA ها ارائه شده توسط مایر و همکاران [۱۸] نامگذاری شدند. مشخصات آنها شامل شماره ثبت شده توالی نسخه‌برداری شونده‌ای که توالی miRNA از آن استخراج شده، جایگاه توالی پیش‌ساز بر روی توالی نسخه‌برداری شونده، بازویی که توالی بالغ miRNA در ساختار دوم توالی پیش‌ساز بر روی آن قرار می‌گیرد، انرژی آزاد و شاخص انرژی آزاد در ساختار دوم توالی پیش‌ساز و محتوی مجموع نوکلئوتید سیتوزین و گوانین توالی پیش‌ساز miRNA مشخص شد. در مرحله بعد توالی‌های پیش‌ساز و بالغ miRNA های شناسایی شده در این مطالعه و miRNA هایی که برای هر گونه گیاهی مورد مطالعه در پایگاه داده ثبت توالی‌های miRNA میریس ثبت شده بود، برای انجام مراحل بعدی مطالعه و آنالیز مورد توجه قرار گرفت.

۲-۲ شناسایی توالی‌های پیش‌ساز miRNA مشتق شده از توالی‌های ترانسپوزونی و رترو ترانسپوزونی و توالی‌های دارای نواحی تکرار شونده و ساده
از پایگاه داده سنسور^۲ (<http://www.girinst.org/censor/index.php>; [۲۳]) شناسایی توالی‌های مشابه ترانسپوزونی و رترو ترانسپوزانی در توالی‌های پیش‌بینی شده پیش‌ساز برای هر گونه گیاهی استفاده شد. در این پایگاه داده عناصر تکرار شونده ترانسپوزونی و رترو ترانسپوزونی به گروه‌های عناصر ترانسپوز شونده، عناصر تکرار شونده، ژن‌هایی در چندین کپی و ویروس‌های تلفیق شونده گروه بندی می‌شوند. با جستجو در این پایگاه داده، توالی‌های پیش‌ساز دارای چنین توالی‌هایی شناسایی شدند. همچنین، از نرم‌افزار تحت وب ریپید ماسکر^۳ ([۲۴] <http://www.repeatmasker.org/>) برای شناسایی توالی‌های تکرار شونده ساده و نواحی با ترکیب ساده نوکلئوتیدی در طول توالی‌های پیش‌ساز استفاده شد.

۲-۳ بررسی حضور یا عدم حضور خانواده‌های miRNA
حضور خانواده‌های ثبت شده در پایگاه داده میریس و شناسایی شده برای هر گونه گیاهی در این مطالعه و عدم حضور آنها در یک گونه گیاهی بررسی شد و به صورت فایل جداگانه برای انجام آنالیزهای بعدی ذخیره شد.

۲-۴ بررسی جایگاه ژنومی miRNA های پیش‌بینی بر روی ژنوم
جایگاه ژنومی توالی‌های شناسایی شده پیش‌ساز miRNA در مورد گونه‌هایی که توالی ژنومی آنها به طور کامل مشخص شده و در پایگاه داده بانک ژن به ثبت رسیده بودند و در دسترس قرار داشتند، تعیین شد. برای این منظور جایگاه توالی‌های پیش‌ساز miRNA استخراج شده به کمک برنامه نرم‌افزاری بلاست نوکلئوتیدی در ژنوم

^۳ Repeatmasker

^۱ RNAfold
^۲ Censor

تکمالی مشخص شد. برای انجام این بخش از مطالعه ابتدا وجود یا نبود خانواده‌های miRNA برای هر گونه به صورت دو خصوصیتی (عدد یک برای خانواده موجود و عدد صفر برای خانواده غایب) برای هر گونه مشخص شد.

۳-نتایج

در این مطالعه تعداد ۴۱۴ توالی miRNA متعلق به ۱۳۰ خانواده در گونه‌های مطالعه شده از خانواده بقولات شناسایی شد. از این تعداد، ۱۵۳ توالی پیش‌ساز حائز شرایط برای گونه *Glycine max* شناسایی شد. این تعداد توالی پیش‌ساز miRNA از ۱۲۱ توالی نسخه‌برداری شونده ژنومی به‌دست آمد. این توالی‌های پیش‌ساز miRNA حامل توالی‌های بالغ miRNA از ۶۶ خانواده miRNA گیاهی بودند (فایل ضمیمه ۲). در این میان خانواده miR5021 دارای بیشترین توالی‌های همولوگ miRNA شناسایی شده بود. توالی‌های پیش‌ساز پیش‌بینی شده برای این گونه طول‌های متفاوتی داشتند. بیشترین طول شناخته شده برای این گونه ۲۹۸ نوکلئوتید می‌باشد. دامنه طولی ۵۰ تا ۸۰ نوکلئوتید بیشترین توالی‌های پیش‌ساز پیش‌بینی شده را به خود اختصاص دادند. ۴۵ عدد از توالی‌های بالغ شناسایی شده برای این گونه در جایگاه اولین نوکلئوتید انتهایی^۱ ۵ توالی خود دارای نوکلئوتید آدنین، ۱۹ توالی نوکلئوتید سیتوزین، ۱۶ توالی نوکلئوتید گوانین و ۷۳ توالی نوکلئوتید اوراسیل بودند. همچنین، توالی‌های بالغ miRNA شناسایی شده برای این گونه نیز طول‌های متفاوتی دارند. توالی‌های بالغ که طول آنها ۲۱ نوکلئوتید بود، بیشترین تعداد را در بین توالی‌های بالغ شناسایی شده داشتند. کمینه انرژی آزاد و شاخص کمینه انرژی آزاد توالی‌های پیش‌ساز کاندید محاسبه شد. بیشتر توالی‌های شناسایی شده کمینه انرژی آزادشان در محدوده ۲۰ تا ۸۰ کیلو کالری بر مول و اکثر توالی‌های پیش‌ساز miRNA

گونه مورد مطالعه جستجو شد. پارامترهای نرم افزار بلاست به جزایی ولیو^۱ که به عدد 10^{-10} تعدیل شد، بقیه پارامترها بر اساس پیشنهاد نرم‌افزار تنظیم شدند. جایگاه‌های ژنومی که با کل طول توالی پیش‌ساز بیش از ۹۸ درصد یکسانی داشتند، مورد توجه قرار گرفته و شماره کروموزوم و جایگاه توالی بر روی کروموزوم به‌عنوان جایگاه ژنومی برای توالی پیش‌ساز miRNA معرفی شد. در مواردی که بیش از یک جایگاه برای توالی miRNA شناسایی شد، همه آنها به‌عنوان جایگاه‌های مضاعف ژن miRNA بر روی ژنوم شناسایی شدند.

۵-۲ مطالعه روند ایجاد و از بین رفتن خانواده‌های

miRNA در گونه‌های گیاهی خانواده بقولات

در این مطالعه روند ایجاد و از بین رفتن خانواده‌های miRNA در گونه‌های خانواده بقولات در طول زمان تکاملی به دو شیوه بررسی شد. در شیوه اول ابتدا بر اساس بیشینه پارسیمونی ایجاد و از بین رفتن خانواده‌های miRNA درخت فیلوژنی بر اساس روش دولو^۲ ترسیم و ایجاد و از بین رفتن خانواده‌های ژنی در هر مسیر تکاملی مورد توجه قرار گرفت. به عبارتی درخت فیلوژنتیکی برای گونه‌های مورد مطالعه، صرفاً بر اساس ایجاد و از بین رفتن خانواده‌های miRNA توسط برنامه دولو-پی^۳ از مجموعه برنامه‌های آنالیز فیلوژنتیکی در بسته فیلپ^۴ وارد و درخت فیلوژنتیکی با بیشترین پارسیمونی ترسیم شد. در ادامه روابط بین گونه‌ها بر این اساس با روابط فیلوژنتیکی موجود بین گونه‌ها مقایسه شد و از این طریق میزان نقش پدید آمدن و از بین رفتن خانواده‌های miRNA ها در ایجاد تنوع بین گونه‌ها ارزیابی شد. در شیوه دوم، درخت فیلوژنی حاصل از روابط فیلوژنتیکی گونه‌ها ترسیم و بیشینه پارسیمونی ایجاد و از بین رفتن خانواده‌های miRNA در مسیرهای این درخت فیلوژنتیکی تعیین و ایجاد شد و از بین رفتن خانواده‌های مختلف در مسیرهای

^۱ Dollo-p
^۴ Phylip

^۱ E-value
^۲ Dollo maximum parsimony

پیش‌بینی شده دارای شاخص کمینه انرژی در محدوده ۰/۸ تا ۱/۲ کیلو کالری بر مول می‌باشند. مطالعه کیفیت توالی‌های پیش‌ساز پیش‌بینی شده برای این گونه نشان داد برخی از توالی‌های شناخته شده دارای نواحی مشابه تکراری در ترکیب خود هستند. همچنین، برخی از این توالی‌های پیش‌ساز پیش‌بینی شده دارای نواحی تکرار شونده و نواحی با ترکیب ساده بودند که مشخصات آنها مشخص شد (فایل ضمیمه ۳).

تعداد ۲۸۶۱۷۶ توالی نسخه‌برداری شونده از ژنوم گونه *Medicago truncatula* از پایگاه داده EST به دست آمد که از بین این توالی‌ها، توالی‌های همولوگ با توالی‌های miRNA گیاهی مورد جستجو قرار گرفت (فایل ضمیمه ۲). از این مجموع توالی‌های نسخه‌برداری شونده تعداد ۹۹ توالی پیش‌ساز miRNA همولوگ با miRNA های گیاهی شناسایی شد. این توالی‌های پیش‌ساز شناسایی شده حامل miRNA های بالغ گیاهی متعلق به ۴۷ خانواده می‌باشند. از این میان ۲۵ خانواده جدید برای این گونه شناسایی شد. خانواده MIR1436 بیشینه عضو miRNA شناسایی شده (به تعداد ۱۰ توالی) را به خود اختصاص داد. طول توالی‌های پیش‌ساز متغیر بودند. بیشترین طول به دست آمده برای توالی پیش‌ساز در این گونه ۳۰۳ نوکلئوتید می‌باشد. پراکنش طول در توالی‌های پیش‌ساز نشان داد که miRNA های شناسایی شده برای این گونه نسبت به سایر گونه‌ها طول بیشتری دارند. اولین نوکلئوتید در توالی‌های بالغ پیش‌بینی شده در این گونه در ۲۹ توالی آدنین، در ۱۰ توالی سیتوزین، در ۹ توالی گوانین و ۵۱ توالی دارای نوکلئوتید یوراسیل بود. توالی‌های miRNA بالغ نیز طول‌های متغیری داشتند و بیشترین فراوانی به توالی‌های بالغ miRNA با طول ۲۱ نوکلئوتید مربوط بود. مطالعه ساختار دوم توالی‌های پیش‌ساز شناسایی شده نشان داد که این توالی‌ها کمینه انرژی آزادی بین ۲۱/۱ تا ۱۸۵/۹ کیلو کالری بر مول داشته و شاخص کمینه انرژی

آزاد توالی‌های پیش‌ساز شناسایی شده برای این گونه بین ۱/۹۱۸ و ۰/۸ کیلو کالری بر مول متغیر است. بررسی توالی‌های پیش‌ساز در این گونه نشان می‌دهد که تعداد زیادی از این توالی‌ها مشابه با توالی‌های ترانسپوزانی و رتروترانسپوزانی هستند. همچنین، تعدادی نواحی تکرار شونده و دارای ترکیب ساده توالی در برخی از توالی‌ها پیش‌ساز مشاهده شد (فایل ضمیمه ۳).

تعداد ۱۸۸۰۸ توالی نسخه‌برداری شونده از ژنوم *Glycine soja* در پایگاه داده EET جمع‌آوری و در این مطالعه استفاده شد. از این مجموع توالی‌های نسخه‌برداری شونده تعداد ۹ توالی پیش‌ساز miRNA همولوگ با miRNA های گیاهی شناسایی شد (فایل ضمیمه ۲). این تعداد توالی‌های شناسایی شده پیش‌ساز miRNA حامل miRNA های همولوگ گیاهی، متعلق به ۵ خانواده بودند. که همه این خانواده‌ها برای اولین بار شناسایی شدند. خانواده MIR169 بیشترین تعداد عضو miRNA شناسایی شده (به تعداد ۴ توالی) را دارد. توالی‌های پیش‌ساز miRNA دارای طول‌های متفاوتی بودند. MIR1520 دارای طولانی‌ترین توالی پیش‌ساز miRNA بود (۱۳۷ نوکلئوتید). پراکنش طول در توالی‌های پیش‌ساز این گونه نشان داد miRNA های شناسایی شده برای این گونه از میانگین طول کمتری نسبت به دو گونه قبلی برخوردار بودند. اولین نوکلئوتید توالی‌های بالغ miRNA پیش‌بینی شده در این گونه در ۳ توالی آدنین، در ۲ توالی سیتوزین، در ۳ توالی گوانین و در یک توالی یوراسیل بود. miRNA های بالغ شناسایی شده در این گونه دارای طول‌های متفاوتی می‌باشند و بیشترین تعداد توالی بالغ miRNA شناسایی شده در این گونه ۲۱ نوکلئوتید طول دارند. ساختار دوم توالی‌های پیش‌ساز شناسایی دارای کمینه انرژی آزاد این توالی‌ها بین ۱۵ تا ۴۹ کیلوکالری بر مول می‌باشد. نواحی تکرار شونده و دارای ترکیب ساده در ۴ پیش‌ساز miRNA مشخص شدند (فایل ضمیمه ۳).

گونه *Lotus japonicus* دارای 243067 توالی نسخه‌برداری شونده ثبت شده می‌باشد. در مطالعه حاضر توالی‌های بالقوه همولوگ با توالی‌های ثبت شده miRNA های گیاهی بر اساس دستورالعمل فوق‌الذکر جستجو شد. از این مجموعه توالی‌های نسخه‌برداری شونده، تعداد ۳۹ توالی پیش‌ساز miRNA همولوگ با miRNA های گیاهی شناسایی شد (فایل ضمیمه ۲). این تعداد از توالی‌های شناسایی شده پیش‌ساز حامل miRNA های همولوگ گیاهی، متعلق به ۲۳ خانواده بودند که ۱۳ خانواده از بین آنها برای اولین بار شناسایی شدند. خانواده MIR5202 با ۴ توالی بیشترین تعداد عضو شناسایی شده را داشت. توالی‌های پیش‌ساز دارای طول‌های متفاوتی بودند. MIR5205 با داشتن توالی پیش‌سازی به طول ۳۰۴- نوکلئوتید طولانی‌ترین توالی پیش‌ساز شناسایی شده می‌باشد. بررسی پراکنش طول توالی‌های این گونه مشخص نمود miRNA های شناسایی شده برای این گونه برعکس گونه *Glycine max* دارای طول کمتری هستند. به طوری که ۱۱ توالی از ۲۴ توالی پیش‌ساز شناخته شده برای این گونه، طولی کمتر از ۸۰ نوکلئوتید داشتند. اولین نوکلئوتید در توالی‌های بالغ miRNA این گونه در ۱۵ توالی یوراسیل، در ۵ توالی آدنین، در ۳ توالی سیتوزین و در یک توالی گوانین بود. توالی‌های بالغ miRNA نیز با طول‌های متفاوت شناسایی شدند. بیشترین تعداد توالی‌های بالغ در این گونه نیز دارای ۲۱ نوکلئوتید می‌باشد. ساختار دوم توالی‌های پیش‌ساز miRNA شناسایی شده برای این گونه دارای کمینه انرژی آزاد بین ۲۰ تا ۸۱/۲ کیلوکالری بر مول می‌باشند. شاخص کمینه انرژی آزاد این توالی‌ها نیز بین ۰/۸۰۳ تا ۱/۴۰۳ کیلوکالری بر مول محاسبه شد. مطالعه کیفیت توالی‌های شناسایی شده پیش‌ساز miRNA در این گونه گیاهی نشان داد که برخی از توالی‌های پیش‌ساز این گونه دارای

تعدادی نواحی با توالی‌های تکرار شونده و دارای ترکیب ساده می‌باشند (فایل ضمیمه ۳).
تعداد ۱۵۳۳۳۷ توالی نسخه‌برداری شونده از ژنوم گونه *Phaseolus vulgaris* در پایگاه داده EST ثبت شده است. از بین این توالی‌های نسخه‌برداری شونده از ژنوم، توالی‌های همولوگ با miRNA های گیاهی مورد جستجو قرار گرفت. از مجموع این توالی‌های نسخه‌برداری شونده تعداد ۳۷ توالی پیش‌ساز همولوگ با miRNA های گیاهی شناسایی شد. این تعداد توالی‌های شناسایی شده پیش‌ساز حامل miRNA هایی متعلق به ۲۰ خانواده می‌باشند که ۸ خانواده برای این گونه جدید شناسایی شدند. طول توالی‌های پیش‌ساز miRNA در این گونه نیز متغیر بود. بیشینه طول به دست آمده برای توالی پیش‌ساز miRNA در این گونه ۲۷۹ نوکلئوتید می‌باشد. ساختار دوم توالی‌های پیش‌ساز miRNA شناسایی شده این توالی‌ها کمینه انرژی آزادی بین ۶/۷ تا ۱۱۵/۲ کیلوکالری بر مول دارند. شاخص کمینه انرژی آزاد این توالی‌ها نیز بین ۰/۸۰۲ تا ۲/۹ کیلوکالری بر مول محاسبه شد. مطالعه کیفیت توالی‌های شناسایی شده پیش‌ساز miRNA در این گونه گیاهی نشان داد برخی از توالی‌های پیش‌ساز این گونه دارای تعدادی نواحی دارای توالی‌های تکرار شونده و دارای ترکیب ساده می‌باشند (فایل ضمیمه ۳).

با بررسی توالی نسخه‌برداری شونده ثبت شده در پایگاه داده EST در مورد گونه *Vigna unguiculata* تعداد ۱۸۹۵۹۳ توالی ثبت به دست آمد. پس از اعمال فیلترهای مورد نظر بر روی توالی‌ها در این گونه، تعداد ۸۴ توالی بالقوه و نامزد miRNA دارای شرایط مورد انتظار شناسایی شد (فایل ضمیمه ۲). این توالی‌ها در ۳۱ خانواده مشخص شدند که ۲۰ خانواده از آنها برای اولین بار برای این گونه شناسایی شدند. همانند سایر گیاهان خانواده فاباسه، خانواده MIR5021 در مورد این گونه هم شناسایی شد. همچنین، این خانواده دارای بیشترین عضو در بین

خانوده‌های شناسایی شده بود (۲۰ miRNA). طویل‌ترین توالی پیش‌ساز miRNA در این گونه ۲۵۶ نوکلئوتید طول دارد. بررسی طول توالی‌های بالغ در این گونه نشان داد که مانند سایر گونه‌ها طولی‌های با طول ۲۱ نوکلئوتید بیشترین تعداد را به خود اختصاص دادند. کمینه انرژی آزاد ساختار دوم توالی‌های پیش‌ساز miRNA شناسایی شده برای این گونه بین ۷/۲ تا ۱۸۷/۴ کیلوکالری بر مول محاسبه شد. همانند سایر گونه‌ها بیشتر توالی‌های پیش‌ساز انرژی آزاد تاخوردگی بین ۲۰ تا ۳۹ کیلوکالری بر مول دارند. شاخص کمترین انرژی تاخوردگی نیز با استفاده از فرمول موبوطه برای توالی‌های پیش‌ساز محاسبه شد. بررسی ساختار توالی‌های پیش‌ساز در این گونه نشان داد که این توالی‌ها دارای نواحی ساده تکرار شونده می‌باشند (فایل ضمیمه ۳). برای گونه *Arachis hypogaea* تعداد ۲۰۵۸۷۲ توالی نسخه‌برداری شونده از پایگاه داده مربوطه به دست آمد که برای شناسایی miRNA های بالقوه این گونه مورد جستجو قرار گرفتند. تعداد ۷۳ توالی بالقوه miRNA شناسایی شد که شامل ۳۳ خانواده بودند و ۲۳ خانواده برای اولین بار در این گونه شناسایی شد. خانواده MIR3509 دارای بیشترین تعداد عضو در این گونه می‌باشد (۱۳ عضو). طویل‌ترین توالی پیش‌ساز دارای ۳۳۲ نوکلئوتید می‌باشد. توالی‌های با طول ۵۰ تا ۷۹ نوکلئوتید بیشترین تعداد به خود اختصاص دادند. بررسی طول توالی‌های بالغ miRNA نشان داد که مانند سایر گونه‌ها توالی‌های بالغ miRNA با طول ۲۱ نوکلئوتید بیشترین تعداد را دارند. تقریباً همه توالی‌های پیش‌ساز انرژی تاخوردگی کمتر از ۷۹ کیلوکالری بر مول دارند و توالی‌های پیش‌ساز با انرژی تاخوردگی بین ۲۰ تا ۳۹ کیلوکالری بر مول بیشترین تعداد را به خود اختصاص دادند. وجود توالی‌های تکرار شونده در ساختار توالی‌های پیش‌ساز شناسایی شده توالی‌ها تکرار شونده در ساختار پیش‌سازها شناسایی شد (فایل ضمیمه ۳).

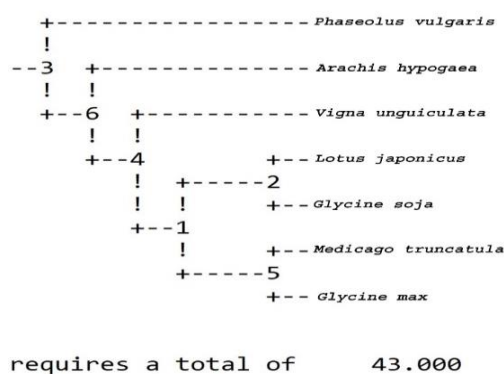
۳-۱ تعیین مکان ژنومی miRNA های شناسایی شده
مکان ژنومی توالی‌های پیش‌ساز miRNA های پیش‌بینی شده از طریق انجام هم‌ردیف‌سازی برای گونه‌هایی که ژنوم آنها به طور کامل توالی‌یابی شده بود، جستجو و مشخص شد. برای گونه‌هایی که ژنوم آنها به طور کامل توالی‌یابی شده بود (۳ گونه شامل: *Glycine max*، *Phaseolus vulgaris*، *Medicago truncatula*)، توالی‌های پیش‌ساز miRNA شناسایی شده مکان‌یابی شد. برای برخی از توالی‌های پیش‌ساز دارای یک جایگاه ژنومی و برخی بیش از یک جایگاه ژنومی شناسایی شد. فاصله جایگاه‌های ژنومی در روی کروموزوم مشابه، مختلف شناسایی شد. برخی از توالی‌ها دارای مکان ژنومی بر روی کروموزوم‌های مختلفی شناسایی شد. برای برخی از توالی‌های پیش‌ساز نیز مکان ژنی مشخصی بر روی ژنوم گونه مورد نظر پیدا نشد. همچنین ژن‌های miRNA بر روی هر دو رشته‌های مثبت و منفی ژنومی شناسایی شد (فایل ضمیمه شماره ۲).

۳-۲ روند ایجاد و از بین رفتن خانواده‌های miRNA در گونه‌های گیاهی خانواده بقولات

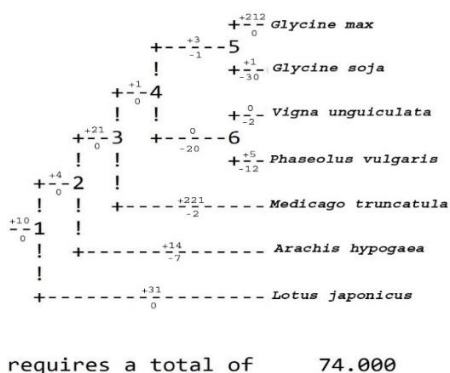
وجود و یا نبود خانواده‌های miRNA در گونه‌ها به صورت دو خصوصیتی (عدد یک برای خانواده موجود و عدد صفر برای خانواده غایب) در فایل جداگانه‌ای تهیه شد (فایل ضمیمه ۴). نکته قابل توجه در خصوص این لیست، نبود خانواده مشترک در لیست همه گونه‌های مورد مطالعه می‌باشد. بنابراین، می‌توان نتیجه گرفت که خانواده اختصاصی miRNA در خانواده بقولات تاکنون شناسایی نشده است. روابط فیلوژنتیکی گونه‌های مورد مطالعه بر اساس ایجاد و از بین رفتن خانواده‌های miRNA در بین گونه‌ها و جنس‌ها توسط مدل بیشینه پارسیمونی دولو مدل شد. نتایج به دست آمده روابط بین گونه‌ها بر اساس روند ایجاد و از بین رفتن خانواده‌های miRNA مشخص کرد (شکل ۱). نتیجه مقایسه درخت فیلوژنتیکی ترسیمی بر

در بین جنس‌ها و گونه‌های این خانواده بر اساس درخت فیلوژنتیکی خانواده مذکور بررسی شد و زمان تقریبی پدید آمدن و از بین رفتن خانواده‌های miRNA در طول زمان تکاملی آنها پیش‌بینی شد (شکل ۲).
بر این اساس این امکان فراهم آمد تا زمان تقریبی پدید آمدن و از بین رفتن خانواده‌های مختلف miRNA ها در طول زمان تکاملی قابل پیش‌بینی باشد. نیز خانواده‌های به وجود آمده و از بین رفته در اجداد این گونه‌ها در خانواده پیش‌بینی شود (جدول ۱).

اساس روند ایجاد و از بین رفتن خانواده‌های miRNA در گونه‌های مورد مطالعه با درخت فیلوژنتیکی آنها نشان داد توپولوژی درخت فیلوژنتیکی ترسیم شده بر اساس روند ایجاد و از بین رفتن خانواده‌های ژنی miRNA و درخت پیش‌بینی روابط فیلوژنتیکی آنها تفاوت نسبتاً زیادی باهم دارد. این مورد بیشتر در مورد گیاهان دولپه‌ای مشاهده و گزارش شده است [۱۴]. همچنین، در این مطالعه روند ایجاد و از بین رفتن خانواده‌های miRNA در گونه‌های گیاهی خانواده بقولات بر اساس بیشینه پارسیمونی دولو



شکل ۱ توپولوژی درخت فیلوژنتیکی ترسیم شده بر اساس روند ایجاد و از بین رفتن خانواده‌های miRNA



شکل ۲. روند ایجاد و از بین رفتن خانواده‌های miRNA در درخت روابط فیلوژنتیکی گونه‌های مورد مطالعه به شیوه مدل بیشینه پارسیمونی دولو. اعداد نمایش داده شده بر روی مسیرهای تکاملی نشان دهنده تعداد خانواده‌های ایجاد شده (+) و از بین رفته (-) در طول مسیر تکاملی را نمایش می‌دهد.

جدول ۱. لیست خانواده‌های miRNA به وجود آمده و از بین رفته در گونه‌ها و اجداد آنها در طول زمان تکاملی بر اساس مدل پیشینه پارسیمونی دولو بر اساس ارتباط فیلوژنتیکی گونه‌ها.

مسیر تکاملی	تعداد miRNA های اضافه شده	خانواده‌های میکروآر ان ای های اضافه شده	تعداد miRNA های حذف شده	خانواده‌های miRNA های حذف شده
ریشه به مسیر 1	10	167-171-172-390-396-397-408-1507-211-9722	0	-
ریشه درخت به گونه <i>Lotus japonicus</i>	10		0	-
مسیر 1 به 2	4	156-159-160-398	0	-
مسیر 2 به 3	21	162-164-166-168-169-319-1530-5037-393-395-399-482-530-1509-1510-1521-5281-5559-5761	0	-
مسیر 3 به 4	1	1533	0	-
مسیر 4 به 5	3	158-2109-3522	1	2595
مسیر 5 به <i>Glycine max</i>	218	1521-395-403-828-854-862-1534-1535-1536-2107-2108-1522-1523-1524-1525-1526-1527-1528-1529-1530-1531-1532-4340-4341-4342-4343-4344-4345-4346-4347-4348-4349-4350-4351-4352-4353-4354-4355-4356-4357-4358-4359-4360-4361-4362-4363-4364-4365-4367-4368-4369-4370-4371-4372-4373-4374-4375-4376-4377-4378-4379-4380-4381-4382-4383-4384-4385-4386-4387-4388-4389-4390-4391-4392-4393-4394-4395-4396-4397-4398-4399-4400-4401-4402-4403-4404-4405-4406-4407-4408-4409-4410-4411-4412-4413-4414-4415-4416-4992-4993-4994-4995-4996-4997-4998-5030-5031-5032-5033-5034-5035-5036-5038-5039-5040-5041-5042-5043-5044-5225-5368-5369-5370-5371-5372-5373-5374-5375-5376-5377-5378-5379-5380-5667-5668-5669-5670-5671-5672-5673-5674-5675-5676-5677-5678-5679-5762-5763-5764-5765-5766-5767-5768-5769-5770-5771-5772-5773-5774-5775-5776-5777-5778-5779-5780-5781-5782-5783-5784-5785-5786-6299-6300-9723-9724-9725-9726-9727-9728-9729-9730-9731-9732-9733-9734-9735-9736-9737-9738-9739-9740-9741-9742-9743-9744-9745-9746-9747-9748-9749-9750-9751-9752-9753-9754-9755-9756-9757-9758-9759-9760-9761-9762-9763-9764-9765-	0	-
مسیر 5 به <i>Glycine soja</i>	1	2218	30	156-159-160-162-163-166-168-169-171-172-319-390-393-395-396-397-398-399-408-530-1533-2111-2118-2119-5037-5281-5559-5761-9722-
مسیر 4 به 6	0	-	20	159-166-167-171-1509-1510-5559-390-393-

396-397-398-530-2114- 2119-2606-5037-5281- 5761-9722-				
1533-2595	2	-	0	مسیر 6 به <i>Vigna unguiculata</i>
160-162-164-168-169- 172-395-399-408-482- 1507-2118	12	779-1435-1438-2019-5140-5185	5	مسیر 6 به <i>Phaseolus vulgaris</i>
171-9722	2	2086-2087-2088-2089-2199-4114-5554-5555-5556-5557- 5558-5560-5561-5562-5563-2585-2586-2587-2588-2589- 2590-2591-2592-2593-2595-2596-2597-2598-2599-2600- 2601-2602-2603-2606-2607-2608-2609-2610-2611-2612- 2613-2614-2615-2616-2617-2618-2619-2620-2621-2622- 2623-2624-2625-2626-2627-2628-2629-2630-2631-2632- 2633-2634-2635-2636-2637-2638-2639-2640-2641-2642- 2643-2644-2645-2646-2647-2648-2649-2650-2651-2652- 2653-2654-2655-2656-2657-2658-2659-2660-2661-2662- 2663-2664-2665-2666-2667-2668-2669-2670-2671-2672- 2673-2674-2675-2676-2677-2678-2679-2680-5204-5205- 5205-5206-5207-5208-5209-5210-5211-5212-5213-5214- 5215-5216-5217-5218-5219-5220-5221-5222-5223-5224- 5225-5226-5227-5228-5229-5230-5231-5232-5233-5234- 5235-5236-5237-5238-5239-5240-5241-5242-5243-5244- 5245-5246-5247-5248-5249-5250-5251-5252-5253-5254- 5255-5256-5257-5258-5259-5260-5261-5262-5263-5264- 5265-5266-5267-5268-5269-5270-5271-5272-5273-5274- 5275-5276-5277-5278-5279-5280-5282-5283-5284-5285- 5286-5287-5288-5289-5290-5291-5292-5293-5294-5295- 5296-5297-5298-5299-5679-5740-5742-5743-5744-5745- 5746-5747-5748-5749-5750-5751-5752-5753-5754-5755- 5756-5757-5758-5759-5780-5781-5782-5783-5784-5785-	221	مسیر 3 به <i>Medicago truncatula</i>
171-172-390-396-397- 2111-9722	7	3508-3509-3510-3511-3512-3513-3514-3515-1516-3517- 3518-3519-3520-3521-	14	مسیر 2 به <i>Arachis hypogaea</i>
-	0	7516-7517-7518-7519-7520-7521-7522-7523-7524-7525- 7526-7527-7528-7529-7530-7531-7532-7533-7534-7535- 7536-7537-7538-7539-7540-7541-7542-7543-7544-7545- 7546-	31	مسیر 1 به <i>Lotus japonicus</i>

۴- بحث

مطالعه miRNA ها می‌تواند به‌عنوان زمینه‌ای با چشم انداز مطلوب و دارای کارایی مناسب برای مهندسی ژنتیک برای افزایش محصولات غذایی و تعدیل اثرات مربوط به تغییرات آب و هوایی ایجاد شده در سطح جهانی ناشی از فعالیت‌های جوامع انسانی بر روی گیاهان و نیز به منظور ایجاد امنیت غذایی از اهمیت ویژه‌ای برخوردار باشد [۲۵]. شناسایی miRNA ها بر اساس جستجوی توالی‌های حامل توالی‌های miRNA هومولوگ با توالی‌های بالغ miRNA شناسایی شده در درون توالی‌های نسخه‌برداری شونده یکی از شیوه‌های کارآمد در شناسایی توالی‌های miRNA به شمار می‌آید [۲۶] که در تلفیق با اطلاعات حاصل از توالی‌یابی، منجر به شناسایی تعداد زیادی ژن‌های miRNA شده است. توالی‌های نسخه‌برداری شونده در شرایط مختلف سلول از روی ژنوم رونویسی می‌شوند. تعداد توالی‌های miRNA قابل شناسایی برای یک گونه بستگی به تعداد توالی‌های رونویسی مطالعه شده و ثبت شده برای آن گونه در پایگاه داده مربوطه دارد. در گیاهان، هرچند توالی‌های پیش‌ساز miRNA حفاظت شده نمی‌باشد، ولی دارای خصوصیات مشخص و قابل تمایزی در ترکیب توالی، طول و ساختار دوم می‌باشند، که امکان تمایز چنین توالی‌هایی را از توالی‌های مشابه دیگر امکان‌پذیر می‌کند [۲۷]. در مطالعه حاضر برای بروز رسانی لیست miRNA های شناسایی شده برای گونه‌های مورد مطالعه بقولات، شناسایی توالی‌های miRNA و خانواده‌های آنها مورد توجه قرار گرفت. نتایج به‌دست آمده نشان داد دستورالعمل به کارگرفته شده از کارایی مناسبی در شناسایی توالی‌های دارای ساختارهای مشابه ساختار توالی‌های miRNA پیش‌ساز که حامل توالی‌های هومولوگ بالغ می‌باشند، برخوردار است. گرچه این شیوه دارای ایراداتی نیز می‌باشد. یکی از ایرادهایی که

دستورالعمل تهیه شده برای جستجوی اینگونه توالی‌ها دارد، این است که تعداد miRNA های شناسایی شده علاوه بر تعداد توالی‌های رونویسی مطالعه شده و ثبت شده بسیار وابسته به پارامترهای تعریف شده بر اساس ویژگی‌های توالی‌های miRNA می‌باشد. یکی دیگر از ایرادهای این دستورالعمل، عدم شناسایی توالی‌های پیش‌ساز miRNA ایی است که در اثر فرآوری آنزیمی RNA های غیر رمزگذار دیگر مانند RNA های ناقل^۱ [۱۷] ایجاد می‌شوند.

در این مطالعه تعداد زیادی توالی‌های ایزو miRNA (miRNA هایی که در یک یا چند توالی در دو انتهای توالی miRNA بالغ باهم اختلاف دارند) [۹] در توالی‌های به‌دست آمده مشاهده شد. به نظر می‌رسد ایجاد چنین توالی‌هایی از توالی‌های پیش‌ساز در نتیجه فعالیت انواع آنزیم‌های DCL^۲ باشد. که در گیاهان تعداد زیادی از انواع این آنزیم‌ها در طول زمان تکامل گونه‌های گیاهی پدید آمده است [۲۷] و یا ممکن است نتیجه اشتباه در برش توالی بالغ miRNA توسط این آنزیم باشد [۱۹ و ۲۸] و ارتباطی با فعالیت پارالوگ‌ها نداشته باشد [۲۸]. در این مطالعه ایزو miRNA ها به دلیل تعلق به یک جایگاه ژنی، چشم‌پوشی شدند.

توالی‌های پیش‌ساز هم‌خانواده مضاعف‌شده^۳ (خوشه شدن^۴) در گونه‌ها به فراوانی مشاهده شد (فایل ضمیمه ۲). فاصله بین توالی‌های مضاعف شده در گیاهان بسیار متنوع مشاهده شد. مضاعف شدن توالی‌های پیش‌ساز در گیاهان فراوان دیده می‌شود و گزارش‌های دیگری از آنها در منابع مختلف آمده است [۱۰ و ۲۹]. جایگاه‌های ژنومی توالی‌های پیش‌بینی شده متفاوت یک توالی پیش‌ساز این امکان را فراهم می‌سازد که در شرایط مختلف امکان رونویسی از جایگاه‌های متفاوت داشته باشد. توالی‌هایی که در جایگاه‌های مختلف به‌صورت یکسان مشاهده

^۳ Duplication

^۴ Clustering

^۱ Transfer RNA (tRNA)

^۲ Dicer like protein (DCL)

می‌شوند، نتیجه پدیده مضاعف شدن ژن در زمان‌های بسیار نزدیک می‌باشند. کپی‌های توالی با فاصله‌های مختلف بر روی یک کروموزوم و حتی بر روی کروموزوم‌های دیگر مشاهده می‌شود. مضاعف شدن یک ژن از فشار انتخاب طبیعی بر روی آن کاسته و امکان ایجاد تغییرات نوکلئوتیدی بیشتر را هم بر روی توالی‌های پیش‌ساز و هم بر روی توالی‌های بالغ فراهم می‌سازد و موجب ایجاد تعدد پارالوگ در miRNA بر روی ژنوم می‌شود. خوشه‌شدن توالی‌های پیش‌ساز غیرهم‌خانواده هر چند در جانوران معمول است [۳۰]، در گیاهان به‌ندرت دیده می‌شود [۳۱]. این گونه خوشه‌ها به دلیل اینکه باهم رونویسی می‌شوند، به‌نظر می‌رسد بر روی یک شبکه ژنی به‌طور هم‌زمان موثر باشند [۳۱].

تعدادی از توالی‌های پیش‌ساز پیش‌بینی شده در ترکیب توالی خود دارای توالی‌های مشابه به ترانسپوزونی و رتروترانسپوزونی می‌باشند (فایل ضمیمه شماره ۳). شواهد نشان می‌دهد ترانسپوزون‌ها و رتروترانسپوزون‌ها در ایجاد تعداد زیادی از خانواده‌های miRNA گیاهی موثر بوده‌اند [۳۳]. ممکن است یکی از شیوه‌های ترانسپوزون‌ها و رتروترانسپوزون‌ها در تکامل گیاهان به واسطه مشارکت آنها در ایجاد برخی توالی‌ها و خانواده‌های miRNA باشد. همچنین، ممکن است شناسایی توالی‌های ترانسپوزونی و رتروترانسپوزونی در توالی برخی miRNA ها که دارای تعداد زیادی توالی‌های مضاعف شده یا تعداد زیادی پارالوگ miRNA در ژنوم می‌باشند، ارتباط مستقیم با فعالیت ترانسپوزون‌ها و رتروترانسپوزون‌ها در این گونه داشته باشد. همچنین، اولین نوکلئوتید انتهایی ۵' در توالی‌های بالغ دارای اهمیت زیادی است. اهمیت نوع نوکلئوتید در ابتدای ۵' توالی بالغ در تمایل خانواده پروتئین خانواده ای‌جی‌او به اتصال با توالی miRNA بالغ عنوان شده است [۹ و ۳۲].

در مورد برخی توالی‌های پیش‌ساز، جایگاه ژنومی آنها بر روی ژنوم گونه مورد مطالعه یافت نشد. فقدان جایگاه ژنومی برخی توالی‌های پیش‌ساز که به شیوه‌های آزمایشگاهی یا بیوانفورماتیکی شناسایی شده است در منابع مختلف نیز گزارش شده است [۱۱ و ۱۵]. به‌نظر می‌رسد یکی از علل عدم شناسایی جایگاه ژنومی در مورد برخی از توالی‌های پیش‌ساز به دلیل تغییراتی است که پس از رونویسی جایگاه ژنومی و ایجاد توالی اولیه miRNA بر روی توالی رخ می‌دهد [۳۳ و ۳۴]. چنین تغییرات در توالی‌های رونویسی شونده رمزگذار پروتئین و تبدیل آنها به RNA پیک (پیامبر) بالغ مشاهده می‌شود [۳۵].

در این پژوهش روابط بین گونه‌ها بر اساس روند ایجاد و از بین رفتن خانواده‌های miRNA به شیوه مدل پیشینه پارسیمونی دلو مدل شد. بر اساس این مدل روند ایجاد و از بین رفتن خانواده‌های miRNA با ارتباط فیلوژنی گونه‌ها همخوانی نداشته است. این نتیجه شاید به این دلیل باشد که miRNA ها به تنهایی در تنوع ایجاد شده در طول زمان تکاملی موثر نبوده‌اند و وجود و نبود آنها به‌تنهایی نمی‌تواند عامل ایجاد تنوع در این گیاهان باشد. هر چند تداوم ایجاد آنها در گونه‌ها و نقش آنها به‌ویژه در بهبود واکنش گیاه به عوامل و استرس‌های محیطی باعث شده در طول زمان تکاملی موجب سازگاری بیشتر آنها شود [۳۶]. همچنین، روند ایجاد و از بین رفتن خانواده‌های miRNA در این گونه‌ها بر اساس مدل پیشینه پارسیمونی در درخت روابط فیلوژنتیکی شناخته شده نشان می‌دهد برخی از خانواده‌های شناخته شده اخیراً در این گونه‌ها ایجاد شده‌اند و اختصاصی گونه‌اند (جدول ۱). به‌نظر می‌رسد ایجاد گونه‌های جدید در طول تکامل آنها با ایجاد و از بین رفتن برخی خانواده‌های همراه بوده است (جدول ۱).

۵- منابع

- [16] Hajieghrari B, Farrokhi N, Goliaei B, Kavousi K (2016) Identification and characterization of novel miRNAs in *Chlamydomonas reinhardtii* by computational methods. *MicroRNA* 5:66-77.
- [17] Rother S, Meister G (2011) Small RNAs derived from longer non-coding RNAs. *Biochimie*, 93(11):1905-15.
- [18] Meyers BC, Axtell MJ, Bartel B, et al (2008) Criteria for annotation of plant Micro RNAs. *Plant Cell*, 20(12):3186-90.
- [19] Kurihara Y, Watanabe Y (2004) Arabidopsis micro- RNA biogenesis through Dicer-like 1 protein functions. *Proceedings of the National Academy of Sciences of the United States of America*. 101:2753–2758.
- [20] Kohany O, Gentles AJ, Hankus L, Jurka J (2006) Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics*, 25(7):474.
- [21] Altschul SF, Gish W, Miller W, et al (1990) Basic local alignment search tool. *J Mol Biol*, 215:403-410.
- [22] Budak H, Bulut R, Kantar M, Alptekin B (2015) Micro RNA nomenclature and the need for a revised naming prescription. *Brif Funct Gen*, 15(1): 1-7.
- [23] Kohany O, Gentles AJ, Hankus L, Jurka J (2006) Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinformatics*, 25(7):474.
- [24] Smit AFA, Hubley R, Green P (2013) RepeatMasker.
- [25] Nekrutenko A, Taylor J (2012) Next-generation sequencing data interpretation: enhancing reproducibility and accessibility. *Nat Rev Genet*, 13(9): 667-72.
- [26] Dezulian T, Remmert M, Palatnik JF, et al (2006) Identification of plant micro RNA homologs. *Bioinformatics*. 22(3):359-360.
- [27] Margis R, Fusaro AF, Smith NA, et al (2006) The evolution and diversification of Dicers in plants. *FEBS Letters*, 580:2442-2450.
- [28] Cuperus JT, Fahlgren N, Carrington JC (2011) Evolution and functional diversification of mi RNA genes. *The Plant Cell*, 23:431-442.
- [29] Axtell MJ (2013) Classification and comparison of small RNAs from plants. *Annu Rev Plant Biol*, 64:137-196.
- [30] Naqvi AF, Islam N, Choudhury NP, Haq QMH (2009) The fascinating world of RNA interference. *Int J Biol Sci*, 5(2): 97–117.
- [31] Li A, Mao L (2006). Evolution of plant micro RNA gene families. *Cell Res*, 1-7.
- [32] Rogers K, Chen X (2013) Biogenesis, Turnover, and mode of action of plant micro RNAs. *The Plant Cell*, 25:2383-2399.
- [33] Kruszka K, Barneche F, Guyot F et al (2013) Plant dicistronic tRNA–snoRNA genes: a new mode
- [1] Hajieghrari B, Farrokhi N (2022) Plant RNA-mediated gene regulatory network. *Genomics* 114 (1), 409-442.
- [2] Farrokhi N, Hajieghrari B (2020) Chronicles of dolos and apate in plant microRNAs. *Biologia*
- [3] Budak H, Akpinar BA (2015) Plant miRNAs: biogenesis organization and origins. *Funct Integr Genomics* 15:523–531. <https://doi.org/10.1007/s10142-015-0451-2>
- [4] Rogers K, Chen X (2013) Biogenesis, Turnover, and mode of action of plant micro RNAs. *The Plant Cell*, 25:2383-2399.
- [5] Voinnet O (2009) Origin biogenesis and activity of plant microRNAs. *Cell* 136:669–687. <https://doi.org/10.1016/j.cell.2009.01.046>
- [6] Khraiweh B, Zhu JK, Zhu J (2012) Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. *Biochemica et Biophysica Acta*, 1819:137-148.
- [7] Hajieghrari B, Farrokhi N, Goliaei B, Kavousi K (2019) The role of microRNAs in defense against viral phytopathogens. *Physiol Mol Plant Pathol* 107: 8–13.
- [8] Nozawa M, Miura S, Nei M (2012). Origins and evolution of micro RNA genes in plant species. *Genome Biol Evol*, 4(3): 230-239
- [9] Ebhardt H A, Fedynak A, Fahlman RP (2010) Naturally occurring variations in sequence length creates micro RNA isoforms that differ in Argonaut effector complex specificity. *Silence*, 1:12.
- [10] Dhandapani V, Ramchiary N, Paul P, et al (2011) Identification of potential micro RNAs and their targets in *Brassica rapa* L. *Mol Cells*, 32:21-37.
- [11] Hajieghrari, B, Farrokhi N, Goliaei B, Kavousi K (2019b) *In silico* identification of conserved miRNAs from *Physcomitrella Patens* ESTs and their target characterization. *Curr Bioinformatics*, 12(6)
- [12] Griffiths-Jones S, Grocock RJ, Van Dongen S, et al (2006) MiRBase: micro RNA sequences targets and gene nomenclature. *Nucleic Acids Res*, 34 D140–4.
- [13] Griffiths-Jones S, Saini HK, Van Dongen S, Enright AJ (2008). MiRBase: tools for micro RNA genomics. *Nucleic Acids Res*, 36 D154–8.
- [14] Hajieghrari B, Farrokhi N (2020) Investigation on the conserved microRNA genes in higher plants. *Plant Molecular Biology Reporter*. <https://doi.org/10.1007/s11105-020-01228-9>
- [15] Hajieghrari B, Farrokhi N, Goliaei B, Kavousi K (2015) Computational identification, characterization and analysis of conserved miRNAs and their targets in *Amborella Trichopoda*. *J Data Mining Genomics Proteomics* 6:2.

[35] Will CL, Lührmann R (2011). Spliceosome structure and function. Cold Spring Harb Perspect Biol, 3(7): a003707.

[36] Chen X (2005). Micro RNA biogenesis and Function in plants. FEBS Letter, 579:5923-5931.

of expression of the small nucleolar RNAs processed by RNase Z. EMBO J 22(3): 621–632.

[34] Mica E, Piccolo V, Delledonne M, et al (2009) High throughput approaches reveal splicing of primary microRNA transcripts and tissue specific expression of mature microRNAs in *Vitis vinifera*. BMC Genomics, 10:558.

***In silicon* prediction of conserved miRNAs in Leguminosae and their gains and losses during evolutionary process**

Behzad Hajieghrari^{1*}, Mojahed Kamalizadeh²

1. PhD. Department of Agricultural Biotechnology, College of Agriculture, Jahrom University, Jahrom, Iran

2. PhD. Department of Agricultural Biotechnology, College of Agriculture, Jahrom University, Jahrom, Iran.

bheghrari@yahoo.com

Receipt: 2021/09/02

Accepted: 2022/06/12

Abstract

Leguminosae is one of the most important plant families. In this study, we searched *Arachis hypogaea*, *Glycine max*, *G. soja*, *Lotus japonicas*, *Medicago truncatula*, *Phaseolus vulgaris*, and *Vigna unguiculata* conserved miRNAs through Expressed Sequence Tag (EST) based-homology method. All candidate sequences with appropriate fold-back structures were screened and characterized according to several filtering criteria. Chromosomal MIR locus and their distributions determined. The Dollo maximum parsimony was employed to construct the species relationship based on the MIR birth and death. Also, the number of MIRs gains and losses in the evolutionary process. In addition, we estimated the numbers of MIRs in their ancestral species using Dollo maximum parsimony in their phylogenetic tree. We found 414 novel miRNAs from 130 MIR families meeting the restricted filtering criteria. Either evolutionary time or the number of miRNAs gains and losses are estimated and characterized in the ancestral species based on the taxon-based phylogenetic tree. It speculates that gains of miRNA gene families within Leguminosae have accelerated during its evolutionary time. In addition, several taxon-specific MIR families find to assign diverse taxon and their species. Our thorough analyses resulted in the definition of some miRNA families as being lineage-specific. Therefore, they can use as markers in future systematic studies.

Keywords: microRNA, Leguminosae, Maximum Dollo parsimony, EST based-homology search.