مروری بر شناسایی ژن‌های هدف میکروRNAها با استفاده از ابزارهای بیوانفورماتیکی در تحقیقات زیست پزشکی

مریم کرائی[1] ، شمس الدین احمدی[2*]

1.کارشناسی ارشد سلولی و مولکولی گروه علوم زیستی، دانشکده علوم، دانشگاه کردستان، سنندج، ایران
2. دانشیار فیزیولوژی، گروه علوم زیستی، دانشکده علوم، دانشگاه کردستان، سنندج، ایران

*نویسنده مسئول: sh.ahmadi@uok.ac.ir

چکیده

میکروRNAها گروهی از RNAهای غیر کدکننده کوچک هستند که بیان ژن در یوکاریوت‌ها را در سطح پس از رونویسی تنظیم می‌کنند. میکروRNAها با تنظیم بیان تعداد زیادی از mRNAها، به‌عنوان تنظیم‌کننده‌های اصلی فرایندهای زیستی مختلفی مانند تکوین جنینی، تکثیر و تمایز سلولی و مرگ برنامه‌ریزی‌شده سلول عمل می‌کنند. بنابراین، شناسایی میکروRNAها و ژن‌های هدف‌شان در شناسایی مکانیسم‌های رشد و نمو و نیز فرایندهای دخیل در ایجاد و پیشرفت بیماری‌ها بسیار مؤثر است. به‌دلیل هزینه‌بربودن کارهای تجربی مولکولی، شناسایی میکروRNAهای مؤثر با استفاده از روش‌های بیوانفورماتیکی و زیست‌شناسی محاسباتی ارزان‌تر و سریع‌تر از روش‌های معمول آزمایشگاهی است. تعدادی بانک اطلاعاتی و نرم‌افزار برخط برای کمک به محققان در پیش‌بینی ژن‌های هدف میکروRNAها توسعه یافته و آزادانه در دسترس قرار گرفته‌اند. نرم‌افزارهای موجود برای پیش‌بینی ژن-های هدف میکروRNAها، از طیف وسیعی از اطلاعات توالی‌یابی، داده‌های مولکولی بیان ژن و نیز الگوریتم‌های محاسباتی مختلف استفاده می‌کنند. تعدادی از مهم‌ترین این ابزارهای برخط عبارت‌اند از: miRWalk، Target Scan، RNAhybrid، Diana-microT، miRanda و MirTarget. چهار ویژگی اصلی برهم‌کنش میان میکروRNA با mRNA هدف (جفت‌شدگی در ناحیه Seed، حفاظت‌شدگی توالی هدف، انرژی آزاد و دسترسی به مکان اتصال در رونوشت هدف) در الگوریتم تمامی این ابزارهای پیش‌بینی هدف میکروRNAها به کار گرفته شده است. هدف از این مطالعه بررسی آخرین یافته‌ها درباره ویژگی‌ها و توانایی‌های بیوانفورماتیکی پیش‌بینی ژن‌های هدف میکروRNA، مقایسه کارایی این ابزارها و درنهایت معرفی کارآمدترین ابزار در زمینه پیش‌بینی ژن هدف میکروRNAها برای تحقیقات بیوانفورماتیکی، زیست‌پزشکی و پزشکی مولکولی است.

کلید واژگان: میکرو RNA، بیوانفورماتیک، پایگاه‌داده‌های RNA، ژن هدف، پردازش‌های پس از رونویسی RNA.

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

مقدمه

میکروRNAها گروهی از مولکول‌های RNA غیرکدکننده کوچک هستند که 19 تا 23 نوکلئوتید طول دارند و بیان ژن در یوکاریوت‌ها را از طریق تخریب mRNA یا مهار ترجمه تنظیم می‌کنند. اتصال میکروRNA به mRNA هدف، از طریق چندین مکانیسم موجب می‌شود سطح mRNA هدف کاهش یابد. ازجمله این مکانیسم‌ها القای د-آدنیله‌شدن در mRNA که سبب می‌شود پایداری mRNA کاهش و میزان تجزیه آن افزایش یابد، مهار مرحله آغاز ترجمه و مهار مرحله طویل‌سازی در ترجمه است که در همه این موارد از ساخته‌شدن پروتئین جلوگیری می‌شود [1, 2]. به‌دلیل پتانسیل هر میکروRNA در هدف قراردادن تعداد زیادی از mRNAها، این دسته از اولیگونوکلئوتیدها به‌عنوان تنظیم‌کننده‌های اصلی فرایندهای زیستی زیادی که عبارت‌اند از تکوین جنینی، تمایز سلولی، تکثیر و آپوپتوز عمل می‌کنند [3, 4]. برهم‌خوردن تنظیم بیان میکروRNAها در ایجاد بیماری‌های مختلفی مانند انواع سرطان‌ها نقش دارد [5]. میکروRNAها به میزان زیادی در دستگاه عصبی مرکزی بیان می‌شوند و در تکوین دستگاه عصبی در دوران جنینی و رشد مغز بعد از تولد و نیز در عملکردهای فیزیولوژیک دستگاه عصبی نقش مهمی دارند [6]. امروزه مشخص شده است که تغییر در بیان میکروRNAها در مغز موجب القای بیماری‌های مختلفی مانند بیماری‌های تحلیل‌برنده دستگاه عصبی و اختلالات روانی مرتبط با دستگاه عصبی مرکزی می‌شود [7, 8]. نقش مهم این دسته از RNAهای غیرکدکننده در فرایند پیری و بیماری‌های مرتبط با آن مانند بیماری آلزایمر به‌خوبی معرفی شده است [9, 10]. در مطالعات جدیدی نیز قابلیت اتصال میکروRNAها به RNAهای غیرکدکننده دیگر مانند RNAهای غیرکدکننده طویل و نقش آنها در مکانیسم‌های مولکولی القای بیماری‌های تحلیل‌برنده عصبی مانند بیماری آلزایمر گزارش شده است [11]. همچنین، مطالعاتی در سال‌های اخیر میکروRNAها را به‌عنوان شناساگرهای زیستی در تشخیص زودهنگام بسیاری از بیماری‌ها ازجمله سرطان و بیماری‌های دستگاه عصبی مانند اختلالات روانی و بیماری‌های تحلیل‌برنده عصبی معرفی کرده‌اند [12-14]. بنابراین، این امید وجود دارد که با افزایش اطلاعات در زمینه میکروRNAها، یک روش مولکولی مطمئن برای تشخیص زودهنگام بیماری‌های مختلف مانند سرطان و بیماری مختلف دستگاه عصبی در دسترس جامعه پزشکی قرار گیرد.

خانواده ژنی میکروRNAها یکی از خانواده‌های بزرگ ژنی در یوکاریوت‌ها هستند. با توجه به دانش روبه‌رشد درباره میکروRNAها و معرفی توالی‌های آنها، پایگاه داده‌های مختلفی برای ارائه توالی‌های میکروRNAهای شناسایی‌شده، راه‌اندازی شده است که از مهم‌ترین آنها پایگاه داده miRBase است [15]. هم‌زمان با شناسایی و کشف میکروRNAها، روش‌های محاسباتی نیز توسعه یافتند و ابزارهایی برای فهم عملکرد میکروRNAها معرفی شدند [16]. با توجه به اینکه تغییرات بیان میکروRNAها و عملکرد آنها با ایجاد و نیز پیشرفت بسیاری از بیماری‌ها ارتباط دارد، در سال‌های اخیر محققان به روش‌هایی روی آورده‌اند که ژن‌های هدف میکروRNAها را پیش از انجام آزمایش‌های تجربی و صرف زمان و هزینه زیاد به دقت تعیین کنند. استفاده از ابزارهای بیوانفورماتیکی علاوه‌بر صرفه‌جویی در وقت و هزینه می‌تواند موجب درک هرچه بهتر عملکرد میکروRNAها در روندهای زیستی و درنتیجه پیش‌بینی اثرات احتمالی آنها شود. برهمین‌اساس ابزارهای متعددی برای پیش‌بینی ژن‌های هدف میکروRNAها ایجاد شده است تا محقق در ابتدا برهم‌کنش احتمالی ژن‌ها و میکروRNA را پیش‌بینی کند و سپس به کمک روش‌های تجربی و آزمایشگاهی، در جستجوی صحت این

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

برهم‌کنش‌ها و مؤثربودن آنها در تنظیم بیان ژن مورد نظر باشد [17]. در این مطالعه، ویژگی‌های مشترک و اختصاصی استفاده‌شده در توسعه هریک از ابزارهای پیش‌بینی هدف میکروRNA ها را بررسی می‌کنیم و چگونگی ترکیب آنها را با ابزارهای موجود برای پیش‌بینی ژن‌های هدف میکروRNAها مرور خواهیم کرد. سپس، نحوه عملکرد ابزارهای miRWalk، TargetScan، RNAhybrid، Diana-microT، miRanda و MirTarget در پیش‌بینی اهداف میکروRNAها مختصر توضیح داده خواهند شد. هدف ما ارائه اطلاعاتی به محققان برای تصمیم‌گیری آگاهانه مبنی بر استفاده از ابزارهای موجود براساس نیازهای یک پروژه خاص در زمینه تحقیقات زیست‌پزشکی است.

بیوسنتز میکروRNAها

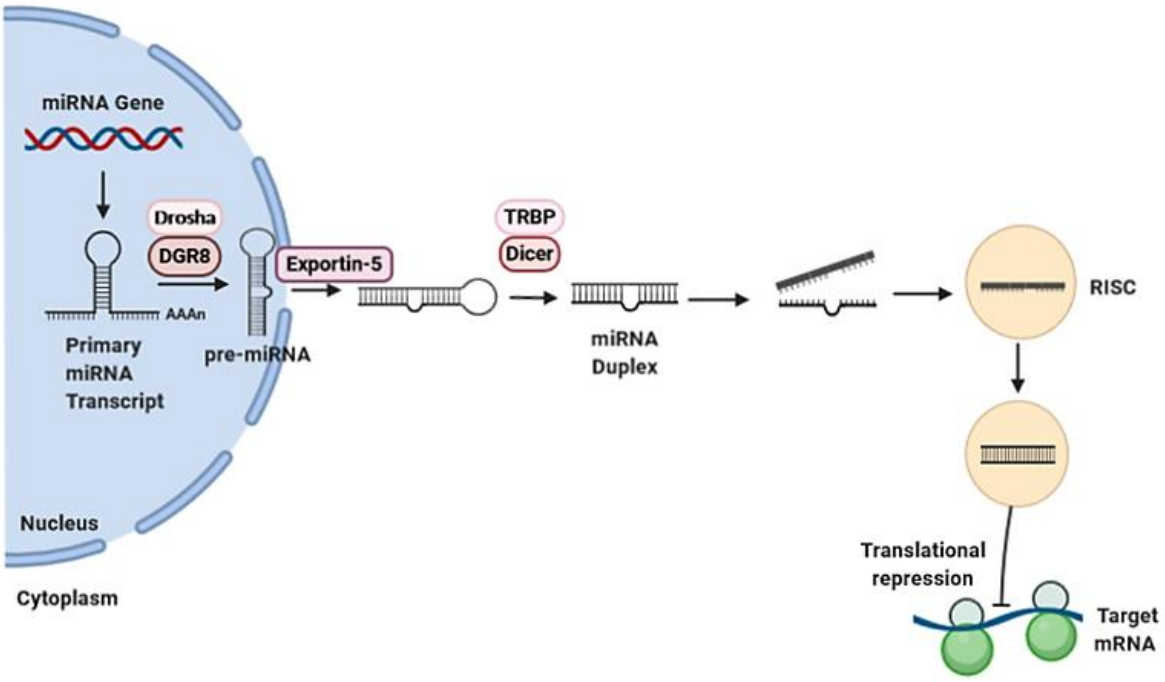
تولید میکروRNAها فرایندی چندمرحله‌ای است. ژن کدکننده میکروRNAها ابتدا توسط RNA پلیمراز II رونویسی می‌شود و رونوشت میکروRNA اولیه[1] (Pri-miRNA) را ایجاد می‌کند. سپس پردازش و بلوغ آنها توسط دو آنزیم اندونوکلئاز و ریبونوکلئاز نوع III (RNase III[2]) که Drosha و Dicer نام دارند، انجام می‌شود. ابتدا Drosha که یک ریبونوکلئاز هسته‌ای است و پروتئین همراه آن به نام DGCR8[3] مجموعه‌ای به نام Drosha/DGCR8 را تشکیل می‌دهند که در هسته سلول، Pri-miRNA را به پیش‌ساز ساقه-حلقه به نام Pre-miRNAپردازش می‌کنند. در مرحله بعدی این ساختار ساقه-حلقه که حدود 70 نوکلئوتید طول دارد، توسط Exportin-5 به سیتوپلاسم منتقل و در آنجا توسط آنزیم RNase IIIسیتوپلاسمی به نام Dicer و پروتئین همراه آن به نام TRBP[4]، به میکروRNA دورشته‌ای با تعداد تقریبی 21 نوکلئوتید برش داده می‌شود [18,19]. سپس یکی از دو رشته میکروRNA روی پروتئینی به نام آرگونات[5] سوار می‌شود و مجموعه‌ای به نام کمپلکس خاموش‌کننده RNA یا RISC [6] را به وجود می‌آورند. این مجموعه به قسمت 3′UTR در mRNA هدف متصل می‌شود و در تنظیم بیان ژن در سطح پس از رونویسی نقش دارد. اگر مکمل‌سازی به‌خوبی انجام شود، کمپلکس RISC ، mRNA هدف را برش می‌دهد؛ اگر جفت‌شدن کامل نباشد کمپلکس RISC با هدف قراردادن 3′UTR در ژن هدف باعث خاموش-کردن ترجمه می‌شود [19,20]. مراحل ساخت و سنتز میکروRNA در شکل- 1 نشان داده شده است.

انتهای 5′ در میکروRNAهای بالغ می‌تواند در خاموش‌کردن ژن هدف نقش داشته باشد. همچنین، عامل تعیین‌کننده در سرکوب ژن هدف است [21]. این ناحیه که Seed نامیده شده است، شش تا هشت نوکلئوتید را در انتهای 5′ به سمت 3′ در میکروRNA شامل می‌شود که رابطه مکملی واتسون – کریک با ناحیه 3′UTR در mRNA دارد. در مطالعات مختلفی نشان داده شده است، ناحیه Seed محدوده‌ای به‌شدت حفاظت‌شده است و براساس توالی این ناحیه میکروRNAها را به خانواده-های مختلفی طبقه‌بندی می‌کنند [22]. جهش در توالی seed درمیکروRNA به ازدست‌رفتن توالی برهم‌کنش میکروRNA با mRNA هدف منجر می‌شود و درنتیجه کارایی عملکردی میکروRNA برای کنترل بیان ژن از بین می‌رود [23].

1. Primary microRNA (Pri-miRNA)
2. Ribonuclease III (RNase III)
3. DiGeorge syndrome critical region 8 protein (DGCR8)
4. Trans activator RNA binding protein (TRBP)
5. Argonaute
6. RNA-induced silencing complex (RISC)

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]



شکل1. مراحل سنتز میکروRNAها

پس از رونویسی از ژن میکروRNA، رونوشت اولیه‌ای به نام (Pri-miRNA) در هسته تولید می‌شود و با اثر یک آنزیم ریبونوکلئاز هسته‌ای به نام Drosha به Pre-miRNA پردازش می‌شود که توسط Exportin-5 به سیتوپلاسم منتقل می‌شود. آنزیم Dicer یک ریبونوکلئاز سیتوپلاسمی است که Pre-miRNA را به میکروRNA دو رشته‌ای برش می‌دهد و سپس یک رشته از آن برای بارگذاری روی کمپلکس خاموش‌کننده القاشده توسط RNA (RISC) انتخاب می‌شود. این کمپلکس میکروRNA را به ناحیه 3′UTR در mRNA هدف هدایت می‌کند و سبب تجزیه mRNA هدف یا مهار ترجمه آن می‌شود.

نحوه نام‌گذاری میکروRNAها

در پایگاه داده miRBase برای نام‌گذاری میکروRNAها از یک روش سه قسمتی مانند dme-miR-100 استفاده می‌شود. پیشوند سه حرفی اول مربوط به اسم اختصاری گونه‌ای است که میکروRNA در آن شناسایی شده است (برای مثال dme مخفف Drosophila melanogaster و hsaمخفف Homo sapiens است). بخش دوم نام میکروRNA (یعنی miR) مخفف miRNA و بخش آخر، شمارهٔ مربوط به ترتیب شناسایی و معرفی میکروRNA را نشان می‌دهد. شایان ذکر است به میکروRNAهای شناسایی‌شده در جایگاه‌های ژنی هومولوگ در موجودات مختلف شمارهٔ مشابهی تعلق می‌گیرد [15].

امروزه در مطالعات زیادی بیان می‌شود که هریک از دو رشتهٔ مربوط به فرم ساقه-حلقه در رونوشت اولیهٔ میکروRNA می‌توانند به میکروRNA بالغ و نهایی تبدیل شوند. بنابراین، از پسوندهای 3p و 5p بعد از شمارهٔ میکروRNA (برای مثال dme-miR-100-5p و dme-miR-100-3p) برای مشخص‌کردن منشاءگرفتن میکروRNA از بخش 5′ یا 3′ رونوشت اولیه میکروRNA استفاده می‌شود [15]. همچنین، اگر توالی میکرو RNAجدید در یک یا دو نوکلئوتید با یک میکروRNA از قبل نام‌گذاری‌شده متفاوت باشد، برای نام‌گذاری آنها از حروف b و a پس از شمارهٔ میکروRNA (مانند miR-181a و miR-181b) استفاده می-

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

شود. اگر تفاوت توالی با یک میکروRNA از قبل نام‌گذاری‌شده بیشتر باشد، نام‌گذاری به‌صورت توافقی توسط پایگاه داده و شخص معرفی‌کنندهٔ میکروRNA جدید انجام می‌شود [24]. ثبت میکروRNA با شماره جدید در پایگاه داده miRBase به تفاوت توالی آن با موارد شناخته‌شده قبلی منوط است. همچنین، شخص مؤلف باید مقاله‌ای معتبر درباره توالی میکروRNA جدید منتشر کرده باشد [15].

پایگاه داده miRBase

پایگاه داده miRbase به آدرس www.miRBase.org از جامع‌ترین پایگاه‌هایی است که اطلاعات مربوط به توالی میکروRNA را از گونه‌های مختلف در خود جای داده است. براساس گزارش اکتبر 2018 تعداد 38589 میکروRNA در این پایگاه داده شناسایی شده‌اند که هرکدام از آنها پتانسیل اتصال به رونوشت صدها ژن هدف را دارند. در قسمت جستجوی این پایگاه می‌توان به دو صورت میکروRNAها را جستجو کرد. شیوه اول جستجو براساس اسم میکروRNAها و نوع دوم جستجو براساس توالی میکروRNAها است[15].

مبانی شناسایی ژن‌های هدف میکروRNA

پیش‌بینی درست ژن‌های هدف و اعتبار زیستی آنها یکی از مشکلات عمده محققان در مطالعات مرتبط با میکروRNAها است، زیرا میکروRNAها طول کمی دارند، به‌ویژه میکروRNAهای جانوران برای اتصال به ژن‌های هدف‌شان توالی مکملی محدودی دارند و این موضوع، پیش‌بینی ژن‌های هدف میکروRNAها را در جانوران بسیار چالش‌برانگیز می‌کند. درحالی‌که پیش‌بینی ژن‌های هدف در گیاهان آسان‌تر است، زیرا میکروRNAهای گیاهی کامل به mRNA هدف خود متصل می‌شوند. امروزه، از روش‌های محاسباتی متفاوتی برای پیش‌بینی جایگاه‌های هدف و احتمال اتصال میکروRNA به رونوشت‌های هدف استفاده می‌شود که الگوریتم‌های کامپیوتری را شامل می‌شود، اما چون جفت‌شدگی با توالی رونوشت ژن هدف ناقص و محدود است، پیش‌بینی دقیق هنوز دشوار است. ابزارهای موجود برای پیش‌بینی ژن‌های هدف میکروRNA معیارهای محاسباتی مختلفی را شامل می‌شوند؛ از مدل‌های برهم‌کنش‌های فیزیکی تا استفاده از یک ماشین یادگیری.[7] در ادامه به ویژگی‌های رایج و کمتر رایج که در این الگوریتم‌ها استفاده می‌شوند، اشاره می‌شود [25].

ویژگی‌های رایج استفاده‌شده در الگوریتم‌های پیش‌بینی ژن‌های هدف میکروRNA

چهار ویژگی معمول که در ابزارهای پیش‌بینی ژن‌های هدف میکروRNA استفاده می‌شود، مطابقت در ناحیه Seed، حفاظت‌شدگی، محاسبه انرژی آزاد و دسترسی به مکان اتصال هستند [26]. هرکدام از این ویژگی‌ها در بخش‌های بعدی جداگانه توضیح داده می‌شود.

جفت‌شدگی در ناحیه Seed

ناحیه seed به توالی شش تا هشت نوکلئوتیدی آغازی میکروRNA گفته می‌شود که از یکی از دو نوکلئوتید شماره یک یا دو از سر 5′ در میکروRNA آغاز می‌شود. در اکثر ابزارهای بیوانفورماتیکی، جفت‌شدن بازها در ناحیه Seed از نوع واتسون و کریک[8] بین یک میکروRNA و mRNA هدف آن صورت می‌گیرد. جفت‌شدگی واتسون-کریک بین نوکلئوتیدهای میکروRNA و mRNA هنگامی اتفاق می‌افتد که آدنوزین (A) با اوراسیل (U) و گوانین (G) با سیتوزین (C) جفت می‌شوند. انواع مختلفی از جفت‌شدگی‌ها در ناحیه Seed وجود دارد که باتوجه به الگوریتم بررسی می‌شوند [22].

ابزارهای پیش‌بینی‌کننده ژن‌های هدف میکروRNA، مکان‌های اتصال بالقوه میکروRNAها به توالی 3′UTR از mRNA را با توجه به الگوهای جفت‌شوندگی اختصاصی

7. Machine-learning
8. Watson-Crick

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

شناسایی می‌کنند. این مکان‌ها می‌توانند در سه دسته طبقه‌بندی شوند [25, 27, 28].

الف- ناحیه مرکزی غالب 5′ [9] که جفت‌شوندگی کاملی بین انتهای توالی 5′ و به‌صورت ادامه‌دار تا انتهای 3′ میکروRNA با mRNA هدف برقرار می‌کند، اما یک نوکلئوتید در وسط این محدوده جفت نمی‌شود و یک حباب را در وسط دو رشته ایجاد می‌کند (شکل2-الف) [25].

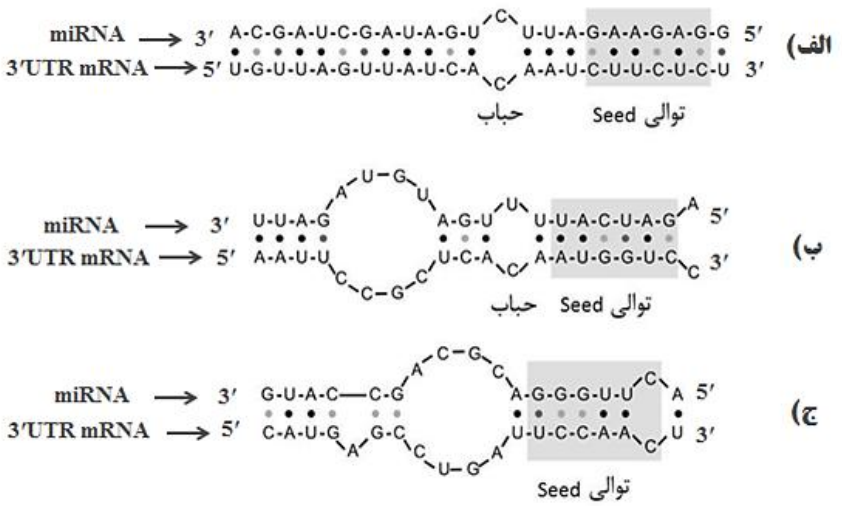
ب- ناحیه seed غالب 5′ [10] ، انتهای توالی seed با mRNA اتصال محکم و تقریباً کاملی برقرار می‌کند، اما انتهای 3′ از میکروRNA به‌صورت ناقص به mRNA متصل می‌شود. به‌طوری‌که سمت 3′ از میکروRNA هنگام اتصال به mRNA یک حباب چند نوکلئوتیدی را تشکیل می‌دهد (شکل2-ب) [25].

ج- ناحیه جبران‌کنندگی 3′ [11] که یک عدم تطابق در قسمت توالی seed با mRNA هدف دارد، اما اتصال تا انتهای 3′ از میکروRNA با mRNA برقرار می‌شود؛ اگر چه حباب بزرگی هم تشکیل می‌شود و درنهایت یک اتصال ضعیف بین میکروRNA و mRNA ایجاد می‌شود (شکل2-ج) [27].

9. 5′-dominant canonical site
10. 5′-dominant seed only site
11. 3′-compensatory site

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]



شکل 2. سه نوع ساختار ثانویه تقریبی حاصل اتصال بین miRNA و mRNA هدف

انواع جفت‌شدگی بین miRNA با قسمتی از mRNA در این شکل مشاهده می‌شود. الف- اتصال از نوع ناحیه مرکزی غالب ′5 است که انطباق کامل در ناحیه Seed وجود دارد، اما یک حباب در وسط تشکیل می‌شود. ب- اتصال از نوع ناحیه Seed غالب ′5 است. انطباق کامل در ناحیه Seed دارد، اما اتصال از سمت ′3 میکروRNA ناقص و یک حباب در آن قسمت تشکیل شده است. ج- ناحیه جبران‌کنندگی ′3 که در جفت‌های لغزنده G:U در ناحیه Seed عدم تطابق دارند، اما جفت‌شدگی تا انتهای ′3 میکروRNA به‌صورت ضعیف ایجاد شده است.

محاسبه انرژی آزاد

انرژی آزاد (یا انرژی آزاد گیبس) می‌تواند به‌عنوان معیار پایداری سیستم بیولوژیکی استفاده شود. اگر اتصال میکروRNA به یک mRNA هدف، پایدار پیش‌بینی شود به احتمال زیاد به‌عنوان یک هدف واقعی برای میکروRNA در نظر گرفته خواهد شد. با توجه به دشواری در اندازه‌گیری انرژی آزاد به‌طور مستقیم، معمولاً تغییر انرژی آزاد (ΔG) طی یک واکنش در نظر گرفته می‌شود. ازآنجایی‌که واکنش‌های دارای ΔG منفی انرژی کمتری برای فعال‌سازی مجدد دارند، به افزایش پایداری منجر می‌شوند. این انرژی آزاد تقریبی و ساختار دوم هیبرید میکروRNA-mRNA خود به وسیله برنامه‌های پیش‌بینی ساختار دوم RNA مانند Vienna package، RNAfold و Mfold محاسبه خواهند شد [29, 30]. نکته‌ای که وجود دارد که این است که صرفاً یک انرژی آزاد پایین هیبریداسیون نمی‌تواند دقت پیش‌بینی ژن‌های هدف میکروRNA را تضمین کند و همچنین، محاسبه آستانه‌های مناسب انرژی آزاد امری دشوار است [31]. بنابراین، درنظرگرفتن خصوصیات اضافی دیگری مانند تجزیه و تحلیل‌های حفاظت‌شدگی برای یک پیش‌بینی مطمئن از رونوشت‌های هدف بدیهی و اجتناب‌ناپذیر است [32].

حفاظت‌شدگی جایگاه‌های هدف

حفاظت‌شدگی[12] به حفظ توالی در بین گونه‌ها اشاره می‌کند. تجزیه و تحلیل حفاظت‌شدگی ممکن است در مناطق ′3UTR و ′5UTR در mRNA، میکروRNA یا هر ترکیبی از این سه مورد متمرکز شود. به‌طورکلی، درناحیه Seed میکروRNA نسبت به ناحیه غیر از Seed، حفاظت‌شدگی بالاتری وجود دارد. اغلب الگوریتم‌های پیش‌بینی‌کننده اهداف میکروRNA به حفاظت‌شدگی

12. Conservation

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

تکاملی توالی‌های میکروRNA بین گونه‌ها وابسته هستند. چنین الگوریتم‌هایی توالی‌هایی را به‌عنوان داده ورودی دریافت می‌کنند که بین دو گونه هومولوگ هستند و روش های مختلفی را برای پیش‌بینی میکروRNAهایی که در این دو گونه حافظت شده‌اند، به کار می‌گیرند. این روش کمک می‌کند تا نتایج مثبت کاذب کاهش پیدا کند، اما در عین حال نتایج به شناسایی میکروRNAهای حفاظت‌شده محدود می‌شوند [25]. همچنین، در زمینه پیش‌بینی اهداف میکروRNA در ناحیه 3′UTR، از تجزیه و تحلیل‌های توالی‌های حفاظت‌شده شواهدی به دست می‌آید که ناحیه هدف پیش‌بینی‌شده ویژگی عملکردی در اتصال به میکروRNA دارد و به این دلیل انتخاب شده است. بنابراین، نقش حفاظت‌شدگی در پیش‌بینی هدف میکروRNA گسترده است و تجزیه و تحلیل عناصر حفاظت‌شده در پیش‌بینی هدف میکروRNA می‌تواند از راه‌های مختلفی انجام شود.

دسترسی به مکان اتصال

قابلیت دسترسی به مکان اتصال درواقع اندازه‌گیری میزان سهولت در اتصال و هیبریدشدن یک میکروRNA با یک mRNA هدف است. پس از رونویسی، mRNA ساختار ثانویه به خود می‌گیرد که این امر می‌تواند با توان اتصال میکروRNA با مکان هدف تداخل ایجاد کند [33]. هیبریداسیون میکروRNA-mRNA فرایندی دو مرحله‌ای است که در آن ابتدا میکروRNA باید به یک ناحیه کوتاه قابل دسترس‌پذیر از mRNA متصل شود و سپس ساختار ثانویه mRNA باز[13] می‌شود تا میکروRNA بتواند کامل به آن متصل شود. بنابراین، برای قوی‌ترشدن احتمال اینکه یک mRNA هدف یک میکروRNA قرار می‌گیرد، پیش-بینی مقدار انرژی مورد نیاز برای ایجاد دسترسی به مکان اتصال می‌تواند ارزیابی شود [34].

ویژگی‌های کمتر رایج استفاده‌شده در الگوریتم‌های پیش‌بینی ژن‌های هدف میکروRNA

با پیشرفت‌های جدید در توصیف برهم‌کنش میکروRNA-mRNA، ویژگی‌های اضافی دیگری نیز ثبت شده است. ممکن است این ویژگی‌ها نیز در الگوریتم‌های پیش‌بینی ژن هدف برای پیش‌بینی اثربخشی هدف یا مستقیماً در پیش‌بینی ژن هدف در نظر گرفته شوند که موارد ذکرشده در زیر شامل می‌شوند [26].

فراوانی مکان‌های اتصال که اندازه‌گیری تعداد جایگاه‌های هدف در ناحیه 3′UTR مربوط به mRNA را شامل می‌شود [35].

محتوای AU مکانی که به تعداد نوکلئوتیدهای A و U موجود در ناحیه seed میکروRNA اشاره می‌کند [36, 37].

جفت‌های لغزنده G:U که در تطبیق با ناحیه Seed به میزان جفت‌شدن G با U به جای C اشاره می‌کند [23].

ثبات جفت‌شدگی در ناحیه Seed که محاسبه انرژی آزاد برهم‌کنش پیش‌بینی شده است [35].

آنالیز موقعیت که موقعیت جایگاه هدف را در mRNA تحلیل می‌کند [38].

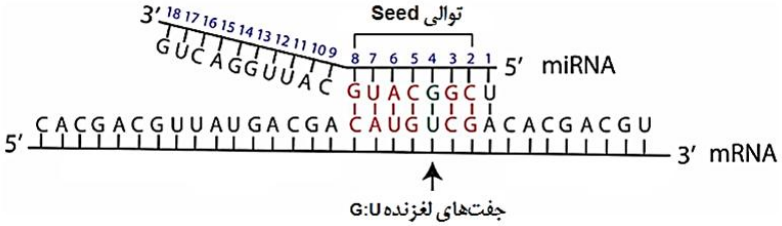
رویکردهای یادگیری ماشین: این روش کامپیوتری از داده‌های آزمایش‌های تجربی استفاده می‌کند تا الگویی از اهداف میکروRNA تهیه شود و سپس از مدل به‌عنوان بخشی از فرایند پیش‌بینی میکروRNA استفاده می‌کند. تکنیک‌های یادگیری ماشین به احتمال زیاد از ویژگی‌های بیشتری در پیش‌بینی‌های خود استفاده می‌کنند، زیرا می-توان آنها را آموزش داد تا قدرت پیش‌بینی‌کنندگی هر ویژگی را در مجموعه داده‌های مثبت و منفی تعیین کند. یک روش یادگیری ماشین که در آن از چندین ابزار بیوانفورماتیکی استفاده شده است، ماشین‌بردار پشتیبان [14] (SVM) نام دارد [26].

13. Unfold

14. Support Vector Machines

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]



شکل3. جفت‌های لغزنده G:U در مطابقت با توالی Seed

نوکلئوتیدهای 2-8 ناحیه Seed را نشان می‌دهند که اتصالات از نوع واتسون-کریک به رنگ قرمز نشان داده شده است و یک مثال از جفت‌های لغزنده G:U با رنگ سبز در توالی Seed مشاهده می‌شود.

نرم‌افزارهای پیش‌بینی اتصال میکروRNA به ژن هدف

در سال‌های اخیر روش‌های تشخیص و آنالیز میکروRNAها به سرعت توسعه یافته است. با این حال، آنالیز میکروRNAها با توجه به ویژگی‌های منحصربه‌فرد آنها مانند کوتاه‌بودن توالی، فراوانی کم و شباهت در توالی بین اعضای یک خانواده نیازمند شرایط ویژه‌ای است. اندازه کوچک میکروRNAها بسیاری از روش‌های سنجش آنها را که بر پایه PCR یا هیبریداسیون است، پیچیده کرده است، زیرا پرایمرهای استفاده‌شده در بسیاری از PCRهای معمولی از نظر طول به اندازه میکروRNAها هستند. درنتیجه، پرایمرهای بسیار کوتاه‌تری برای سنجش آنها نیاز هست که این عامل روی بازده PCR تأثیرگذار است. همچنین، در روش‌های تشخیصی بر پایه هیبریداسیون، نشان‌دارکردن پروب کوتاه برای شناسایی میکروRNA دشوار است. در ضمن دمای ذوب دورشته‌ای پروب و توالی هدف آن نیز پایین خواهد بود که به‌شدت باعث کاهش میزان هیبریداسیون و افزایش تلاقی هیبریداسیون می‌شود. بنابراین، افزایش راهکارهای جدید برای افزایش اختصاصی بودن اندازه‌گیری پروفایل میکروRNAها ضروری است [16]. در طول 15 سال اخیر، براساس معیارهای زیستی مشخصی، بسیاری از ابزارهای بیوانفورماتیک برای پیش‌بینی برهم‌کنش میکروRNA-mRNA ایجاد شده‌اند. این الگوریتم‌های پیش‌بینی ژن‌های هدف میکروRNA غالباً براساس میزان مطابقت ناحیه Seed با توالی هدف، حفاظت‌شدگی جایگاه‌های هدف، محاسبه انرژی آزاد هیبریدهای میکروRNA-mRNA و میزان دسترسی به مکان اتصال در توالی هدف پایه‌گذاری شده‌اند که همه این موارد غالباً برای صحت بیشتر پیش‌بینی هدف با هم ترکیب می‌شوند و این امر را ممکن می‌کنند که ژن‌های هدف میکروRNAها شناسایی و در مدل‌های زیستی تأیید شوند [35].

درمجموع، ابزارهای بیوانفورماتیک به سه دسته تقسیم می‌شوند: یک دسته از آنها خدمات را بر مبنای شبکه[15] ارائه می‌دهند، محیط کاربری آنها بیش از بقیه دوستانه و استفاده از آنها آسان‌تر است. دسته دوم نرم‌افزارهای دانلودشده[16] را شامل می‌شوند که ازنظر سهولت محیط کاربری پس از دسته اول قرار می‌گیرند. دسته سوم از نوع پکیج[17] هستند که استفاده از آن برای کاربر آسان نیست و به اطلاعات بیوانفورماتیکی نیاز دارند [16]. در جدول-1 اسامی خلاصه این نرم‌افزارها و آدرس دسترسی به آنها آورده شده است.

15. Web-based services
16. Downloaded software
17. Package

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

جدول1. انواع نرم‌افزارهای پیش‌بینی ژن‌های هدف میکروRNA

ردیف	نام نرم‌افزار یا پایگاه داده	آدرس سایت نرم‌افزار یا پایگاه داده
1	miRanda	http://www.microrna.org
2	TargetScan	http://www.targetscan.org
3	RNAhybrid	http://bibiserv.techfak.uni-bielefeld.de/rnahybrid
4	Diana-microT	http://diana.cslab.ece.ntua.gr/microT
5	miRWalk	http://mirwalk.uni-hd.de
6	RNA22	http://cbcsrv.watson.ibm.com/rna22.html
7	PITA	http://genie.weizmann.ac.il/pubs/mir07
8	picTar	http://pictar.mdc-berlin.de
9	MirTarget	http://mirdb.org

نرم‌افزار برخط و پایگاه داده miRanda

پایگاه داده miRanda از اولین برنامه‌های پیش‌بینی‌کننده ژن‌های هدف میکروRNA است که Enright و همکاران در سال 2003 آن را طراحی کردند و با گذشت زمان به‌روز شده است. در ابتدا به‌منظور شناسایی ژن‌های هدف میکروRNA در مگس سرکه[18] ابداع شد، اما پس از چندی این الگوریتم برای پیش‌بینی ژن‌های هدف در انسان نیز استفاده شد. این ابزار بیوانفورماتیکی هم به‌صورت آنلاین و هم دانلودشده در دسترس قرار دارد. الگوریتم miRanda از آنالیز سه‌مرحله‌ای استفاده می‌کند. سه قانون مکمل‌بودن توالی، انرژی آزاد هیبریدهای میکروRNA-mRNA و حفاظت‌شدگی جایگاه‌های هدف در انتخاب ژن‌های هدف برای هر میکروRNA در نظر گرفته می‌شود. ابتدا، توالی میکروRNA ارائه‌شده به‌عنوان ورودی، از سمت 3′UTR برای بررسی جفت‌شدن بازها طبق اصول واتسون ـکریک اسکن می‌شود. در مرحله بعد، انرژی آزاد هریک از هیبریدهای میکروRNA-mRNAهای جفت‌شده که از نمره تطبیق آستانه بالاتر باشد، محاسبه و هر هدفی که مقدار انرژی آزاد پیش‌بینی‌شده برای آن کمتر از حد آستانه باشد به مرحله بعد منتقل می‌شود. سرانجام، از حفاظت‌شدگی جایگاه‌های هدف به‌عنوان فیلتر نهایی استفاده می‌شود. ژن‌های هدف کاندیدشده بر-اساس میزان مطابقتشان با توالی میکروRNA امتیاز می-گیرند. ژن‌های هدف پیش‌بینی‌شده می‌توانند با کسب نمره بالا یا با داشتن چندین جایگاه اتصال پیش‌بینی‌شده در اولویت قرار گیرند. میکروRNAهای با جایگاه اتصال چندگانه درون ناحیه 3′UTR ارجح‌تر هستند که

هر بخش در افزایش ویژگی اختصاصی بودن اتصال میکروRNA به ژن هدف نقش دارد[32].

برخلاف بیشتر برنامه‌های پیش‌بینی‌کننده ژن‌های هدف میکروRNA، miRanda تطابق در کل توالی میکروRNA را در نظر می‌گیرد. این الگوریتم با اهمیت‌دادن ویژه به جفت‌شدگی در ناحیه Seed، جفت‌شدگی در منطقه Seed را مهم‌تر از ویژگی‌های دیگر در نظر می‌گیرد و حتی جفت‌شدگی در این ناحیه می‌تواند جفت‌های لغزنده G:U را نیز شامل شود؛ یعنی وجود یک حباب در توالی seed را مجاز در نظر می‌گیرد. همچنین، محاسبه انرژی آزاد هیبرید میکروRNA-mRNA در این الگوریتم با استفاده از Vienna package انجام می‌شود [30]. miRanda می-تواند از 10 هدف صحیح موجود برای هر میکروRNA حدود 9 عدد را شناسایی کند و نسبت نتایج مثبت کاذب

18. Drosophila melanogaster

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

در آن حدوداً 24% است. همچنین، برای بررسی دقیق‌تر و پیشرفته‌تر نیز می‌تواند تعدادی پارامتر را تنظیم ارائه دهد که ممکن است در بررسی اهداف میکروRNAهای خاصی مفید باشند [32].

نرم‌افزار برخط و پایگاه داده TargetScan

پایگاه داده TargetScan را Lewis و همکاران در سال 2003 توسعه دادند و اولین نرم‌افزاری بود که برای پیش-بینی اهداف میکروRNAهای انسانی به کار برده شد [22]. این الگوریتم با ادغام خصوصیات ترمودینامیکی برهم‌کنش‌های میکروRNA-mRNA و آنالیزهای مقایسه‌ای توالی، ژن‌های هدف میکروRNAها را در موجودات مختلف (مانند انسان، موش، رت و ماهی) پیش‌بینی می‌کند. برخلاف روش‌های دیگر پیش‌بینی ژن‌های هدف که در آنها ابتدا جفت‌شدگی احتمالی تمام قسمت‌های میکروRNA بررسی و سپس با استفاده از معیاری‌های دیگر نتایج مثبت کاذب فیلتر می‌شوند، در روش TargetScan ابتدا جفت‌شدگی کامل برای ناحیه seedجستجو، سپس به نواحی اطراف گسترش داده و با رسیدن به ناحیه عدم تطابق[19] یا ناحیه جفت‌های لغزنده G:U متوقف می‌شود. به‌این‌ترتیب، بسیاری از نتایج مثبت کاذب از همان ابتدا فیلتر می‌شوند. این الگوریتم انرژی آزاد ترمودینامیکی اتصالات بین میکروRNA و رونوشت ژن‌های هدف بالقوه را نیز محاسبه و سپس با برنامه RNAfold Package توالی‌های seed را بررسی می‌کند و به هر ناحیه UTR نمره می‌دهد. سپس تکرارها به‌منظور دسته‌بندی UTRها از موجودات دیگر مانند موش و رت و ماهی برای آنالیزهای فیلوژنتیکی پردازش می‌شوند. چندین ویژگی دیگر نیز شامل ناحیه جبران‌کنندگی 3′، محتوایAU موجود و موقعیت جایگاه اتصال در هنگام احتساب نمره پیش‌بینی اتصال میکروRNA به mRNA هدف در نظر گرفته می‌شوند. نتایج نشان داد این روش فقط جایگاه‌های اتصال میکروRNAهای شناخته‌شده را پیش‌بینی نمی‌کند، بلکه 451 جایگاه بالقوه جدید را نیز پیش‌بینی می‌کند. نسبت مثبت کاذب تخمین‌زده‌شده بین 22 تا 31 % است [32]. نرم‌افزار برخط TargetScan به کاربر اجازه می‌دهد با نام میکروRNA، نام ژن یا خانواده‌های میکروRNA کاملاً محافظت‌شده یا نسبتاً محافظت‌شده را در میان توالی‌های چندین گونه جستجو کند. استفاده از نرم‌افزار برخط TargetScan آسان است و به تنظیم ورودی توالی یا تنظیمات پیشرفته نیاز ندارد که این امر می‌تواند برای کاربران تازه‌کار مزیت باشد[26].

نرم‌افزار برخط و پایگاه داده RNAhybrid

Rehmsmeier و همکاران در سال 2004، یک نرم‌افزار پیش‌بینی هدف میکروRNA مبتنی بر تجزیه و تحلیل ساختار ثانویه هیبرید میکروRNA-mRNA را توسعه دادند که RNAhybrid نام گرفت. درحقیقت، RNAhybrid الگوریتم پیش‌بینی ساختار ثانویه RNA کلاسیک است. این الگوریتم نسخه توسعه‌یافته برنامه‌های پیش‌بینی‌کننده ساختارهای ثانویه RNA کلاسیک مانند MFold و RNAFold است و نواقص نرم‌افزارهای Mfold و RNAFold را برطرف کرده است. همچنین، می‌تواند انرژی آزاد جایگاه‌های اتصال را برای یک RNA کوچک درون توالی‌های هدف RNA بزرگ (درحالتی‌که RNA کوچک با RNA بزرگ مطابقت دارد) سریع و دقیق محاسبه کند. الگوریتم RNAhybrid جایگاه‌های اتصال با انرژی آزاد بسیار مطلوب را تخمین می‌زند، اما به تخمین انرژی آزاد جایگاه‌های اتصال درون‌مولکولی اجازه نمی-دهد؛ یعنی انرژی آزاد جفت‌شدن در بین نوکلئوتیدهای mRNA یا بین نوکلئوتیدهای میکروRNA را در نظر نمی‌گیرد. RNAhybrid اهداف میکروRNA را به‌وسیله جفت‌شدن شش نوکلئوتید آغازی seed (از شماره دو تا شماره هفت) در انتهای 5′ توالی میکروRNAی بالغ شناسایی می‌کند [25].

19. Missmatch

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

در سال 2006، نرم‌افزار RNAhybrid توسط Kruger و Rehmsmeier به‌روز شد [29]. این برنامه می‌تواند به‌عنوان سرویس‌دهنده تحت وب استفاده شود که این امر به کاربر امکان می‌دهد پارامترهای الگوریتمی و ماهیت یک Seed را آزادانه انتخاب کند. اول، موقعیت و طول Seed را در نظر می‌گیرد. دوم، مجازبودن یا نبودن وجود جفت‌های لغزنده W:B و G:U درون Seed و به این طریق انرژی آزاد اتصال بین میکروRNA و mRNA با ناحیه Seed را که کاربر تعریف کرده است، برای پیش‌بینی هدف در نظر می‌گیرد. RNAhybrid همچنین می‌تواند مقداری p-value براساس تعداد جایگاه‌های اتصال در انتهای 3′UTR برای اتصال میکروRNA-mRNA اختصاص دهد که فراوانی جایگاه‌های اتصال در محاسبه مقدار آن در نظر گرفته می‌شود [26].

این ابزار بیوانفورماتیکی برای کاربران پیشرفته در نظر گرفته شده است و گزینه‌هایی برای دستکاری تنظیمات پیشرفته دارد که مخصوص این ابزار است. از تنظیماتی که کاربر می‌تواند انجام دهد، تعداد جایگاه‌های اتصال در هر ژن هدف، محدودیت‌های مارپیچ، حداکثر اندازه حلقه داخلی، حداکثر اندازه حلقه برآمدگی و حداکثر انرژی آزاد cutoff است که این موارد در بخش راهنمای موجود در وب‌سایت RNAhybrid مفصل توضیح داده شده است [26].

نرم‌افزار برخط و پایگاه داده Diana-microT

الگوریتم Diana-microT در سال 2004 را Kiriakidou و همکاران با ترکیب‌کردن رویکردهای تجربی و بیوانفورماتیکی طراحی کردند [39]. الگوریتم Diana-microT برای شناسایی عناصر تشخیصی بالقوه میکروRNAها از یک محدوده 38 نوکلئوتیدی که در محدوده میانی 3′UTR رونوشت هدف بالقوه موجود است، استفاده می‌کند. Diana micro T را می‌توان برای پیش‌بینی ژن‌های هدف میکروRNAهای حاوی عناصر تشخیصی[20] (MRE) استفاده کرد. این الگوریتم، ژن‌های هدف میکروRNAها را به‌طورکلی بر اساس دو اصل تعیین می‌کند: اول، از ترمودینامیک به‌عنوان شاخص اولیه پیش‌بینی جایگاه‌های اتصال استفاده می‌کند و انرژی آزاد جفت‌شدن بازهای مکمل واتسون-کریک و همچنین جفت‌های لغزنده G:U را محاسبه می‌کند (برخلاف روش TargetScan که به اتصالات ضعیف در توالی Seed که جفت‌های لغزنده G:U را شامل می‌شود نیز اجازه ورود به پیش‌بینی را می‌دهد) و سپس با مقایسه انرژی آزاد نواحی کاندیدای جفت‌شدگی، ژن‌های هدف احتمالی را پیش‌بینی می‌کند [39,32]. دوم، پروتئین‌های مرتبط با میکروRNAها بر اندازه و موقعیت حباب مرکزی تأثیر می گذارند که به ایجاد اثراتی در اتصال میکروRNAها به ژن مورد نظر منجر می‌شوند. همچنین، برای جستجوی جایگاه‌های اتصال، این الگوریتم ناحیه Seed 5′ انتهایی، حباب مرکزی و اتصال انتهای 3′ میکروRNAها را نیز در نظر می‌گیرد [26].

این الگوریتم همه جایگاه‌های هدف میکروRNA در نماتود سی الگانس[21] را که به خوبی شناخته شده بودند، با موفقیت تشخیص داد و نیز هفت ژن هدف میکروRNAی پستانداران را که با روش‌های آزمایشگاهی پیش‌بینی شده بودند، تأیید کرد. علاوه‌براین، هنگامی‌که درستی این روش از طریق روش ریزآرایه[22] بررسی شد، سطح درستی یا صحت پیش‌بینی 66 % را نشان داد که بالاتر از میانگین بسیاری از برنامه‌های پیش‌بینی‌کننده اهداف میکرو RNA بود [32].

نرم‌افزار برخط و پایگاه داده miRWalk

برای بیش از ده سال، تلاش‌ها برای مطالعه برهم‌کنش میکروRNAها و اهداف آنها روی 3′UTR در mRNA متمرکز شده بود، اما بعدها مطرح شد که میکروRNAها

20. miRNA-recognition elements
21. *Caenorhabditis elegans (C. elegans)*
22. Microarray

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

ممکن است بیان ژن را با هدف قراردادن پروموتور، 5′UTR و همین‌طور ناحیه کدکننده ژن یا CDS نیز هدف قرار دهند. الگوریتم miRWalk در سال 2011 را Dweep و همکاران براساس روش‌های محاسباتی و بر پایه جفت‌شدن بازهای مکمل در قانون واتسون-کریک بین میکروRNA و توالی هدف طراحی و پایه‌گذاری کردند [40]. این الگوریتم به‌دنبال seedهایی می‌گردد که طبق قانون مکمل‌بودن واتسون–کریک با توالی ژن هدف در هفت نوکلئوتید seed مربوط در انتهای ′5 میکروRNA با mRNA هدف مکمل باشند. بنابراین، هپتامر seed و توالی مکمل هپتامر را شناسایی می‌کند و پیش می‌رود تا به اولین عدم تطابق می‌رسد. این‌کار را برای ژن‌های مختلف انجام می‌دهد. سپس، این جایگاه‌های اتصال براساس موقعیت قرارگیری‌شان که در موقعیت آغازی (Start)، موقعیت پایانی (N) یا بخشی از ژن باشند (Region) دسته‌بندی می‌شوند. درنهایت، نتایج پیش‌بینی‌شده در پنج گروه که عبارت‌اند از توالی پروموتور، 5′UTR، CDS، 3′UTRو ژن‌های میتوکندریایی دسته‌بندی خواهند شد [41]. الگوریتم miRWalk پیش‌بینی ژن‌های هدف را در انسان، موش و رت انجام می‌دهد. همچنین، این نرم‌افزار نتایج به‌دست‌آمده را با نرم‌افزارهای دیگر مانند Diana-microT، TargetScan، RNAhybrid، miRanda مقایسه می‌کند و نتایج به‌صورت جدول به نمایش گذاشته می‌شوند. این الگوریتم به کاربران اجازه می‌دهد کنترل بیشتری بر داده‌های پیش‌بینی داشته باشند و نوع ترکیب اطلاعاتی را که مایلند برای تحقیق‌شان در نظر بگیرند، انتخاب کنند. می‌توان گفت روش‌هایی که مبنای عمل آنها ترکیبی از الگوریتم‌های مختلف است، دید دقیق‌تری درباره احتمال هدف‌بودن ناحیه مورد مطالعه در اختیار محقق قرار می‌دهند [26].

نرم‌افزار برخط و پایگاه داده miRDB با ابزار بیوانفورماتیکی MirTarget

نرم‌افزار برخط miRDB با ابزار بیوانفورماتیکی MirTarget پیش‌بینی ژن‌های هدف میکروRNA را با استفاده از ماشین‌بردار پشتیبان یا SVM و ویژگی‌های استخراج‌شده از یک مجموعه داده بزرگ حاصل از روش ریزآرایه انجام می‌دهد. این رویکرد یادگیری ماشین، استفاده از چندین ویژگی رایج پیش‌بینی را تأیید کرد و ویژگی‌های جدیدی نشان داد که ارتباط مهمی با برهم‌کنش میکروRNA-mRNA دارند. برخی ویژگی‌های مورد استفاده که عبارت اند از حفاظت‌شدگی ناحیه seed، مطابقت ناحیه seed به‌ویژه در موقعیت‌های 2 تا 8، جفت‌شدگی بازها در مناطق حاوی جایگاه‌های اتصال به ناحیه seed و محل قرارگیری جایگاه‌های اتصال، در قسمت 3′UTR قرار دارند. MirTarget در تعامل با miRDB ایجاد شده است و دسترسی به آن از طریق سایت miRDB امکان‌پذیر است. پایگاه داده miRDB یک پایگاه اطلاعاتی برخط برای پیش‌بینی اهداف و تفسیر عملکردی اتصال میکروRNAها به رونوشت ژن‌های هدف است. همه ژن‌های هدف میکروRNAها در miRDB با ابزار MirTarget و آنالیز هزاران برهم‌کنش میکروRNA با ژن‌های هدف پیش‌بینی می‌شوند که از آزمایش‌های توالی‌یابی توان بالا به دست آمده‌اند. به‌طورکلی، miRDB کاربرپسند است و پیش‌بینی برهم‌کنش میکروRNA-رmRNA را با واردکردن نام هر میکروRNA یا نام ژن هدف می‌توان جستجو کرد. داده‌پایگاه miRDB اهداف میکروRNAهای گزارش‌شده در پنج گونه انسان، موش، رت، سگ و مرغ را هدف‌یابی می‌کند. همچنین، کاربران می‌توانند توالی دلخواه خود را برای پیش‌بینی سفارشی با استفاده از الگوریتم به‌روزشده در سایت انجام دهند. همچنین با ترکیب آنالیزهای عملکردی و استخراج داده‌های مقالات، میکروRNAهای

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

فعال و عملکردی در موش و انسان که موجب تغییراتی واقعی در بیان ژن با نتایجی عملکردی در سلول و تغییراتی در عملکرد فیزیولوژیک شده‌اند، در پایگاه داده miRDB شناسایی و معرفی شده‌اند [42].

در مطالعه‌ای که نتایج آن در سال 2019 منتشر شد، الگوریتم miRTarget به‌روز و 96 ویژگی مختلف برای ساخت مدل نهایی با هم ترکیب و مدل پیش‌بینی هدف نهایی آنها به نام MirTargetV4.0 معرفی شده است. نرم‌افزار برخط این الگوریتم امکانات رایجی دارد که برای هر دو ویژگی اتصال میکروRNA و تنظیم پایین‌دستی عملکردی ژن هدف مهم هستند. در آنالیزهای مقایسه‌ای نشان داده شده است که ابزار بیوانفورماتیکی MirTargetV4.0 از ابزارهای دیگری که برای پیش‌بینی ژن هدف میکروRNA استفاده می‌شوند کارایی بیشتری دارند [42].

نتیجه‌گیری

از زمان کشف اولین میکروRNAها و ژن‌های هدف آنها در سال 1993، رشد چشمگیری در تعداد میکروRNAها و شناسایی اهداف معتبر یا فرضی آنها حاصل شده است که با تعدادی از الگوریتم‌های محاسباتی پشتیبانی می‌شوند. پیش‌بینی ژن هدف میکروRNA به‌عنوان روشی محاسباتی برای پیش‌بینی عملکرد میکروRNAها، احتمالاً مهم‌ترین بخش در مطالعه میکروRNA است. از سال 2002 همراه توسعه روش‌های آزمایشگاهی، ابزارهای بیوانفورماتیکی برای تحقیق درباره میکروRNAها گسترش یافته‌اند. هرکدام از این نرم‌افزارها روی چندین ویژگی و اصول شناسایی ژن‌های هدف تأکید بیشتری می‌کنند. برای مثال، الگوریتم miRanda بر ویژگی اتصال ناحیه Seed متمرکز شده است و جفت‌شدن در این ناحیه را مهم‌تر از ویژگی‌های دیگر می‌داند. الگوریتم RNAhybrid انرژی آزاد برهم‌کنش‌های میکروRNA با ژن هدف را دقیق‌تر از الگوریتم‌های دیگر محاسبه می‌کند. ابزار بیوانفورماتیکی Dianamicro-T به جفت‌های لغزنده G:U نیز امکان قرارگیری در ناحیه seed را می‌دهد و انرژی آزاد پیوندهای آنها را محاسبه می‌کند؛ درحالی‌که TargetScan فقط روی جفت‌شدگی‌های کامل درون توالی seed تمرکز دارد. الگوریتم miRWlak که امکان مقایسه نتایج چندین نرم‌افزار را با هم برای کاربران فراهم می‌کند نیز جفت‌شدگی کامل درون توالی seed را مد نظر قرار می‌دهد. بااین‌حال، الگوریتم به‌روزشده MirTargetV4.0 که جدیدترین ابزار بیوانفورماتیکی ارائه‌شده تا به امروز به شمار می‌رود، به‌دلیل ادغام 96 ویژگی مختلف برای شناسایی ژن‌های هدف از الگوریتم های دیگردقیق‌تر و پیشرفته‌تر و نتایج حاصل از آن در مطالعات مقایسه‌ای انجام‌شده از الگوریتم‌های دیگر به واقعیت نزدیک‌تر و مستندتر است. در یک مطالعه نظری، ما با مقایسه پیش‌بینی ژن‌های هدف میکروRNAهای دخیل در بیماری آلزایمر نشان دادیم که قدرت پیش‌بینی ابزار MirTarget از صحت و کیفیت بیشتری برخوردار است [43]. بنابراین، می‌توان گفت این الگوریتم می‌تواند به‌عنوان جدیدترین و بهترین نرم‌افزار برخط برای پیش‌بینی عملکردی ژن‌های هدف میکروRNA ها معرفی شود.

شایان ذکر است که با وجود پیشرفت‌های فراوان در زمینه برنامه‌های پیش‌بینی ژن‌های هدف میکروRNAها، الگوریتم‌ها و ابزارهای بیوانفورماتیکی در دسترس هنوز کارایی کاملی ندارند و در پیش‌بینی ژن‌های هدف اشتباهات مکرر دارند. این امر مشکلاتی را در انجام‌دادن کارهای آزمایشگاهی برای محققان به وجود می‌آورد که گاهی از نظر زمانی و اقتصادی بسیار هزینه‌بر است [3, 41]. اگرچه این الگوریتم‌ها هنوز ویژگی‌های حساسیت و اختصاصیت کاملی ندارند، اما می‌توانند به محققانی که اهداف جدید میکروRNAها را بررسی می‌کنند، کمک‌های ارزشمندی کنند [18]. با ادغام رویکردهای محاسباتی و

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

آزمایشگاهی در کل ژنوم، تحقیقات درباره میکروRNAها از هر زمان دیگری موفق‌تر خواهد بود. ازآنجایی‌که الگوریتم‌های موجود برای پیش‌بینی ژن هدف بر فرضیات مختلفی متکی هستند، به نظر می‌رسد ترکیب نتایج حاصل از ابزارهای مختلف یک رویه مناسب باشد و اغلب در کاهش نتایج مثبت کاذب یا منفی تا حد امکان استفاده شود. با بررسی توانایی ابزارهای بیوانفورماتیکی جدید می‌توان انتظار داشت که الگوریتم‌های موجود به‌تدریج دقیق‌تر شوند و با افزودن دانش جدید بهبود یابند [25, 44]. بنابراین، پیشرفت‌های بیشتر در پیش‌بینی محاسباتی ژن-های هدف میکروRNAها به کمک ابزارهای بیوانفورماتیکی از موارد مهم تحقیقات حال و آینده خواهد بود [32, 44].

تشکر و قدردانی

مؤلفان مقاله از معاونت پژوهش و فناوری دانشگاه کردستان برای فراهم‌کردن دسترسی به متن کامل مقالات خارجی تقدیر و تشکر می‌کنند.

مجوزهای اخلاقی

مقاله از نوع مروری است و به اخذ مجوز اخلاقی نیازی نبوده است.

تعارض منافع

درخصوص این مقاله تعارض منافعی وجود ندارد.

سهم نویسندگان

مریم کرائی در جستجوی مقالات، آماده‌کردن شکل‌ها و نوشتن مطالب اولیه برای مقاله و شمس‌الدین احمدی در نوشتن مقاله و اصلاحات نهایی در این مقاله مشارکت داشته‌اند.

حمایت ها و منابع مالی

در تهیه این مقاله از منبع مالی خاصی استفاده نشده است.

منابع

1. Bartel DP. *MicroRNAs: genomics, biogenesis, mechanism, and function*. Cell 2004; 116 (2): 281-297.

2. Li LC, Okino ST, Zhao H, et al. *Small dsRNAs induce transcriptional activation in human cells*. Proc Natl Acad Sci U S A 2006; 103 (46): 17337-17342.

3. Ambros V. *The functions of animal microRNAs*. Nature 2004; 431 (7006): 350-355.

4. Parvini N, Ahmadi S. *Role of MicroRNAs in Development of Immune Cells and Nervous System and their Relation to Multiple Sclerosis*. Shefaye Khatam 2015; 3 (1): 131-144.

5. Ali Syeda Z, Langden SSS, Munkhzul C, Lee M, Song SJ. *Regulatory Mechanism of MicroRNA Expression in Cancer*. Int J Mol Sci 2020; 21 (5).

6. Yapijakis C. *Regulatory Role of MicroRNAs in Brain Development and Function*. Adv Exp Med Biol 2020; 1195: 237-247.

7. Karthikeyan A, Patnala R, Jadhav SP, Eng-Ang L, Dheen ST. *MicroRNAs: Key Players in Microglia and Astrocyte Mediated Inflammation in CNS Pathologies*. Curr Med Chem 2016; 23 (30): 3528-3546.

8. Van den Berg MMJ, Krauskopf J, Ramaekers JG, Kleinjans JCS, Prickaerts J, Briedé JJ. *Circulating microRNAs as potential biomarkers for psychiatric and neurodegenerative disorders*. Prog Neurobiol 2020; 185: 101732.

9. Danka Mohammed CP, Park JS, Nam HG, Kim K. *MicroRNAs in brain aging*. Mech Ageing Dev 2017; 168: 3-9.

10. Putteeraj M, Fairuz YM, Teoh SL. *MicroRNA Dysregulation in Alzheimer's Disease*. CNS Neurol Disord Drug Targets 2017; 16 (9): 1000-1009.

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

11. Ahmadi S, Zobeiri M, Bradburn S. *Molecular mechanisms underlying actions of certain long noncoding RNAs in Alzheimer's disease*. Metab Brain Dis 2020; 35 (5): 681-693.

12. Backes C, Meese E, Keller A. *Specific miRNA Disease Biomarkers in Blood, Serum and Plasma: Challenges and Prospects*. Mol Diagn Ther 2016; 20 (6): 509-518.

13. Swarbrick S, Wragg N, Ghosh S, Stolzing A. *Systematic Review of miRNA as Biomarkers in Alzheimer's Disease*. Mol Neurobiol 2019; 56 (9): 6156-6167.

14. Mazumder S, Datta S, Ray JG, Chaudhuri K, Chatterjee R. *Liquid biopsy: miRNA as a potential biomarker in oral cancer*. Cancer Epidemiol 2019; 58: 137-145.

15. Kozomara A, Griffiths-Jones S. *miRBase: integrating microRNA annotation and deep-sequencing data*. Nucleic Acids Res 2011; 39 (Database issue): D152-157.

16. Riffo-Campos Á L, Riquelme I, Brebi-Mieville P. *Tools for Sequence-Based miRNA Target Prediction: What to Choose?* Int J Mol Sci 2016; 17 (12).

17. Juźwik CA, S SD, Zhang Y, et al. *microRNA dysregulation in neurodegenerative diseases: A systematic review*. Prog Neurobiol 2019; 182: 101664.

18. Cullen BR. *Transcription and processing of human microRNA precursors*. Mol Cell 2004; 16 (6): 861-865.

19. Filipowicz W, Bhattacharyya SN, Sonenberg N. *Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?* Nat Rev Genet 2008; 9 (2): 102-114.

20. Zamore PD, Haley B. *Ribo-gnome: the big world of small RNAs*. Science 2005; 309 (5740): 1519-1524.

21. Lai EC. *Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation*. Nat Genet 2002; 30 (4): 363-364.

22. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. *Prediction of mammalian microRNA targets*. Cell 2003; 115 (7): 787-798.

23. Doench JG, Sharp PA. *Specificity of microRNA target selection in translational repression*. Genes Dev 2004; 18 (5): 504-511.

24. Griffiths-Jones S. *The microRNA Registry*. Nucleic Acids Res 2004; 32 (Database issue): D109-111.

25. Dweep H, Sticht C, Gretz N. *In-Silico Algorithms for the Screening of Possible microRNA Binding Sites and Their Interactions*. Curr Genomics 2013; 14 (2): 127-136.

26. Peterson SM, Thompson JA, Ufkin ML, Sathyanarayana P, Liaw L, Congdon CB. *Common features of microRNA target prediction tools*. Front Genet 2014; 5: 23.

27. Brennecke J, Stark A, Russell RB, Cohen SM. *Principles of microRNA-target recognition*. PLoS Biol 2005; 3 (3): e85.

28. Mazière P, Enright AJ. *Prediction of microRNA targets*. Drug Discov Today 2007; 12 (11-12): 452-458.

29. Wuchty S, Fontana W, Hofacker IL, Schuster P. *Complete suboptimal folding of RNA and the stability of secondary structures*. Biopolymers 1999; 49 (2): 145-165.

30. Mathews DH, Sabina J, Zuker M, Turner DH. *Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure*. J Mol Biol 1999; 288 (5): 911-940.

31. Watanabe Y, Tomita M, Kanai A. *Computational methods for microRNA target prediction*. Methods Enzymol 2007; 427: 65-86.

32. Min H, Yoon S. *Got target? Computational methods for microRNA target prediction and their extension*. Exp Mol Med 2010; 42 (4): 233-244.

33. Mahen EM, Watson PY, Cottrell JW, Fedor MJ. *mRNA secondary structures fold sequentially but exchange rapidly in vivo*. PLoS Biol 2010; 8 (2): e1000307.

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

34. Long D, Lee R, Williams P, Chan CY, Ambros V, Ding Y. *Potent effect of target structure on microRNA function*. Nat Struct Mol Biol 2007; 14 (4): 287-294.

35. Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP. *Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs*. Nat Struct Mol Biol 2011; 18 (10): 1139-1146.

36. Friedman RC, Farh KK, Burge CB, Bartel DP. *Most mammalian mRNAs are conserved targets of microRNAs*. Genome Res 2009; 19 (1): 92-105.

37. Betel D, Koppal A, Agius P, Sander C, Leslie C. *Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites*. Genome Biol 2010; 11 (8): R90.

38. Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. *MicroRNA targeting specificity in mammals: determinants beyond seed pairing*. Mol Cell 2007; 27 (1): 91-105.

39. Kiriakidou M, Nelson PT, Kouranov A, et al. *A combined computational-experimental approach predicts human microRNA targets*. Genes Dev 2004; 18 (10): 1165-1178.

40. Dweep H, Sticht C, Pandey P, Gretz N. *miRWalk--database: prediction of possible miRNA binding sites by "walking" the genes of three genomes*. J Biomed Inform 2011; 44 (5): 839-847.

41. Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R. *Fast and effective prediction of microRNA/target duplexes*. Rna 2004; 10 (10): 1507-1517.

42. Liu W, Wang X. *Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data*. Genome Biol 2019; 20 (1): 18.

43. Koraei M, Ahmadi S. *Comparing MicroRNA Target Gene Predictions Related to Alzheimer's Disease Using Online Bioinformatics Tools*. J Health Biomed Inform. 2021; 7 (4) :376-389 [In Persian].

44. Chen L, Heikkinen L, Wang C, Yang Y, Sun H, Wong G. *Trends in the development of miRNA bioinformatics tools*. Brief Bioinform 2019; 20 (5): 1836-1852.

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]

Biotechnology-Tarbiat Modares University

A review on microRNAs target prediction with bioinformatics tools in biomedical research

Maryam Koraei[1], Shamseddin Ahmadi [2*]

1. M.Sc. Department of Biological Science, Faculty of Science, University of Kurdistan, Sanandaj, Iran
2. Ph.D. Department of Biological Science, Faculty of Science, University of Kurdistan, Sanandaj, Iran

*Corresponding Author: sh.ahmadi@uok.ac.ir

Abstract

MicroRNAs are a group of small non-coding RNAs that regulate gene expression in eukaryotes at the post-transcriptional level. Regulating the expression of large numbers of mRNAs, MicroRNAs act as major regulators of various biological processes such as embryonic development, cell proliferation, differentiation, and apoptosis. Therefore, the identification of microRNAs and their target genes is very effective in finding the mechanisms of embryonic development, growth, and also the processes underlying the induction and progression of various diseases. Because of the high costs of molecular experiments, the identification of effective microRNAs through bioinformatics tools and computational biology is faster and cheaper than the experimental methods. Several online bioinformatics tools and databases have been developed and are freely available for predicting microRNAs target genes. The available online tools use a broad range of information, including sequencing data, gene expression data, and computational algorithms for predicting microRNAs target genes. Some of the most important of these online tolls are miRWalk, TargetScan, RNAhybrid, Diana-microT, miRanda, and MirTarget. The four main features of the interaction between a microRNA and an mRNA, including seed pairing, sequence conservation, free energy, and access to the binding site in a target are used in the algorithm of all of these prediction tools. This study aimed to review the latest findings on the characteristics and capabilities of microRNA target prediction tools, comparing the performance of these tools, and finally introducing the most efficient tool in the field of target gene prediction for bioinformatics, biomedicine, and molecular medicine studies.

Keywords: microRNAs, Bioinformatics, RNA databases, Target gene, Posttranscriptional RNA Processing

[Downloaded from biot.modares.ac.ir on 2024-05-03]

[DOR: 20.1001.1.23222115.1400.12.4.4.6]