



پیش بینی و تایید *in silico* ریز RNA ها در بافت های مختلف گوسفند با منشا کروموزوم ۲۰

وحیده عباسی^۱ - محمدرضا نصیری^۲ - علی جوادمنش^{۳*}

تاریخ دریافت: ۱۳۹۷/۰۵/۱۴

تاریخ پذیرش: ۱۳۹۷/۰۶/۳۱

چکیده

در سال‌های اخیر با روشن شدن اهمیت ریز RNAها در فرآیندهای حیاتی بدن موجودات زنده، شناسایی ژن‌های ریز RNA اهمیت زیادی پیدا کرده است از سویی دیگر روش‌های بیوانفورماتیکی نیز گسترش یافته و سبب شده تا روند شناسایی ریز RNAها با سرعت بیشتر و هزینه کمتری پیش برود. تحقیقات در سال‌های اخیر بر روی گونه‌های زیادی از جمله انسان و موش و همچنین حیوانات اهلی از قبیل گاو و بز و مرغ باعث شده تا در این گونه‌ها ریز RNAهای زیادی گزارش شود با این حال در گونه‌ی گوسفند هنوز اطلاعات نسبتاً کمی وجود دارد. در این مطالعه پس از پیش‌بینی ریز RNAهای کروموزوم ۲۰ گوسفند، جهت تایید با اورتولوگ‌ها در سایر گونه‌ها مقایسه و در نهایت برای بررسی بیان از داده‌های ترنسکرپتوم گوسفند استفاده شد همچنین در این تحقیق برای ارزیابی روش مورد استفاده از داده‌های شاهد مثبت استفاده شد. با استفاده از این روش در کل ۴۰۰ ریز RNA و از این تعداد ۸۱ ریز RNA جدید در کروموزوم ۲۰ گوسفند پیش‌بینی شدند. بررسی ترنسکرپتوم نشان داد که از این تعداد ۳۳ ریز RNA فقط در بافت عضله، ۱۰ ریز RNA فقط در بافت کبد، ۳۵ ریز RNA در هر دو بافت کبد و ماهیچه و ۳ ریز RNA نیز در مخلوط بافت‌های قلب، کلیه، مغز، تخمدان، پوست، چربی سفید و ریه بیان شدند. مقایسه با پایگاه داده miRbase مشخص کرد که تعداد ۶۴ ریز RNA از ریز RNAهای پیدا شده برای اولین بار گزارش شدند. حساسیت و دقت انتخاب این روش برابر ۶۷ و ۹۵ درصد ارزیابی شد و از این رو می‌توان گفت این روش می‌تواند به عنوان مکمل روش‌های پیش‌بینی ریز RNA اطلاعات ارزشمندی را فراهم کند. از این روش می‌توان برای پیش‌بینی ریز RNAها در سایر کروموزوم‌های گوسفند و یا در سایر پستانداران استفاده نمود.

واژه های کلیدی: بیوانفورماتیک، روش از آغاز، ریز RNAها، گوسفند

مقدمه

به این‌ها ریز RNAهای بین ژنی گفته می‌شود (۳). بسیاری از ریز RNAها در میان جانوران خویشاوند (مانند انسان و موش) دارای یک توالی هستند این ریز RNAهای اورتولوگ حفاظت شده هستند (۳). ثابت شده است که گروهی از ریز RNAها برای شکل‌گیری اندام‌های خاص مانند مغز (۱۱) و قلب (۲۱) و برای تمایز خون‌سازی (۵) ضروری هستند.

دو روش اصلی برای شناسایی ریز RNAها وجود دارد شامل روش‌های آزمایشگاهی و روش‌های محاسباتی (بیوانفورماتیکی). مرسوم‌ترین روش آزمایشگاهی روش توالی‌یابی کتابخانه cDNA از روی ریز RNAهای کوچک است این روش یک روش موثر است اما دارای محدودیت‌هایی است از جمله رایج‌ترین آن‌ها، عدم شناسایی ریز RNAهایی است که در دوران خاصی از زندگی موجود زنده و یا در یک بافت خاص بیان می‌شوند و برخی از ریز RNAهایی که ممکن است به سختی کلون شوند و همچنین این روش یک روش هزینه‌بر است (۴). روش‌های محاسباتی پیش‌بینی ریز RNAها دو نوع است: روش اول یا روش مقایسه‌ای که محدود به ریز RNAهای حفاظت شده‌ای است که قبلاً در گونه‌های نزدیک کشف شده است (۸).

ریز RNAها (miRNA) متعلق به طبقه‌ای از RNAهای کوچک تکامل یافته هستند (۱۹-۲۵ نوکلئوتید) که به پروتئین تبدیل نمی‌شوند در گیاهان و حیوانات به عنوان تنظیم کننده بیان ژن در سطح پس از رونویسی شناخته شده‌اند. این ژن‌ها پس از رونویسی به ریز RNAهای اولیه (pri-miRNA) تبدیل و پس از پردازش به ریز RNAهای پیش‌ساز (Pre-miRNA) تبدیل می‌شود و سپس وارد سیتوپلاسم می‌شود و در آنجا به ریز RNA بالغ تبدیل می‌شود (۱). تعداد بسیار اندکی از ریز RNAها در داخل اگزون ژن‌های کد کننده پروتئین قرار می‌گیرند این ژن‌ها یا درون اینترون‌های ژن‌های کد کننده پروتئین قرار گرفته‌اند (۱۲) و یا خارج از ژن‌ها قرار دارند که

۱، ۲ و ۳- به ترتیب دانش آموخته دوره کارشناس ارشد، استاد و استادیار گروه علوم دامی، دانشکده کشاورزی، دانشگاه فردوسی مشهد، مشهد، ایران.
* ایمیل نویسنده مسئول: javadmanesh@um.ac.ir
DOI: 10.22067/ijasr.v11i2.74619

پیش بینی اولیه ریز RNA

روش استفاده شده در این تحقیق روش از آغاز است اما همراه با تغییراتی در جهت تکمیل و افزایش دقت به شرح زیر اعمال شده است. در این مطالعه از نرم‌افزارهای پیش‌بینی‌کننده‌ی EMBOSS (<http://emboss.bioinformatics.nl/cgi-bin/emboss/einverted>) و Sequence-Structure Motif و Base: Pre-miRNA Prediction برای پیش‌بینی ریز RNAهای موجود در کروموزوم ۲۰ گوسفند استفاده شد.

گزینش ریز RNAهای پیش‌بینی‌شده

تمام pre-miRNAهایی که بین نرم‌افزارهای پیش‌بینی‌کننده مشترک بودند وارد نرم‌افزار آنالاین mfold شدند و با توجه به خصوصیات ریز RNAهای پیش‌ساز، گزینش شدند. توالی پیش‌بینی‌شده و فیلتر شده با نرم‌افزار mfold، در بین ریز RNAهای ثبت شده در miRBase با در نظر گرفتن حداقل E-value برابر ۱۰، بلاست شد. توالی‌های مشابه با توالی ریز RNAهای شناخته شده (با عدم تطابق کمتر از ۴ نوکلئوتید)، به عنوان ریز RNA حفظ شده معرفی شد (۱۶).

ریز RNAهایی که در بخش آگزون ژن کد کننده پروتئین هستند کنار گذاشته شدند. برای این منظور ریز RNAهای پیشنهادی مرحله‌ی قبل، با توالی کد کننده پروتئین در گوسفند بلاست شد (در قسمت BLASTx وب سایت NCBI)، ریز RNAهای پیشنهادی که در توالی کد کننده پروتئین نبودند به عنوان ریز RNA معرفی شدند.

تأیید بیان به روش in silico

برای اطمینان از بیان شدن ریز RNAهای پیش‌بینی شده، همه ریز RNAهای گزینش شده با استفاده از داده‌های بیان SRA (Sequence Read Archive) بافت عضله (Accession: PRJNA223213)، بافت کبد (Accession: GSM1366318) و همچنین در مخلوط داده‌های بیان بافت‌های قلب، کلیه، مغز، کبد، تخمدان، پوست، چربی سفید و ریه گوسفند (Accession: GSE56643) بلاست شد.

محاسبه‌ی دقت پیش‌بینی

برای محاسبه دقت این روش، کلیه مراحل ذکر شده در بالا روی داده‌های مثبت (ناحیه‌ای از کروموزوم ۱۸ گوسفند حاوی ۵۴ ریز RNA ثبت شده) اعمال شد و سپس حساسیت و دقت انتخاب با استفاده از دو فرمول زیر محاسبه شد (۱۸):

$$Sensitivity = 100 * \frac{TP}{TP + FN} \quad Selectivity = 100 * \frac{TP}{TP + FP}$$

در فرمول بالا، TP تعداد ریز RNA پیش‌ساز است که به درستی پیش‌بینی شده‌اند (مثبت درست)، FN منفی کاذب (به معنی عدم

روش دوم پیش‌بینی محاسباتی ریز RNA روش از آغاز^۲ است که بر اساس ویژگی‌های ساختار دوم RNA، ثبات ترمودینامیکی ساختار ساقه-حلقه و خصوصیات میکرو RNA پیش‌ساز است. برای بهبود این روش نرم‌افزارهای زیادی ساخته شده است از جمله MiRFinder (۹)، BosFinder (۱۳) و MiRPara (۲۰). این روش از لحاظ هزینه خیلی پایین‌تر از روش آزمایشگاهی است و می‌تواند miRNA کل ژنوم را پیش‌بینی کند که در هر بافت یا سلولی و همچنین در هر دوره‌ای از زندگی موجود زنده بیان می‌شوند. همچنین محدود به ریز RNAهای حفظ شده نیست. اما در این روش احتمال خطای مثبت کاذب به معنی معرفی اشتباه قطعه‌ای از ژنوم به عنوان ریز RNA پیش‌ساز بالاست (۱۸). برای اطمینان از صحت پیش‌بینی و کنار گذاشتن داده‌های مثبت کاذب دو رویکرد وجود دارد ۱ بررسی بیان miRNAها ۲ انتخاب ویژگی ساختار مناسب (۷).

تاکنون هزاران ریز RNA در بیش از صد گونه‌ی مختلف شناخته و با روش‌های آزمایشگاهی تأیید شده اما در گوسفند به این موضوع کمتر پرداخته شده است و اطلاعات محدود تری در مورد ریز RNAهای گوسفند نسبت به سایر گونه‌های حیوانات اهلی وجود دارد. در این تحقیق روش تکمیل شده‌ی بیوانفورماتیکی طراحی شد و استفاده از این به دلیل مقایسه ریز RNAهای پیش‌بینی شده و بررسی بیان in silico تا حد زیادی خطاهای روش‌های متداول شناسایی را کاهش می‌دهد. در این تحقیق از کروموزوم شماره ۲۰ گوسفند به علت اندازه‌ی نسبتاً کوچک و نیز حضور یک کلاستر نسبتاً بزرگ از ریز RNAهای تأیید شده استفاده گردید.

مواد و روشها

اخذ توالی‌های مرجع

توالی کروموزوم ۲۰ از سایت NCBI اخذ شد (Oar_v4.0 Chr (NC_019477.2):1 - 51,049,468). سپس برای اطمینان از صحیح بودن روش کار و همچنین کاربردی بودن نرم‌افزارهای به کار گرفته شده، توالی بین نوکلئوتید ۶۴۴۶۵۷۶۰ تا ۶۴۴۶۶۰۸۸۰ از کروموزوم ۱۸ گوسفند، به عنوان شاهد مثبت انتخاب و از NCBI (Accession number: NW_004080181.1) اخذ شد. در این ناحیه از کروموزوم ۱۸، تعداد ۵۴ ریز RNA ثبت شده در پایگاه miRBase قرار دارد. در مرحله بعد ریز RNAهای پیش‌ساز مربوط به گونه‌ی گوسفند از miRBase دریافت شد و با استفاده از نرم‌افزار mfold ساختار دوم آن رسم و برای شناخت خصوصیات ریز RNAهای پیش‌ساز، مورد بررسی قرار گرفت (۱۴).

شناسایی کرده است miR-219 و miR-2284 بوده است که در مطالعه‌ی حاضر نیز وجود دارند.

تأیید بیان به روش *in silico*

نتیجه بلاست کردن ریز RNA های پیشنهادی در داده‌های SRA نشان داد که تعداد ۸۱ ریز RNA در داده‌های بیان بافت‌های مختلف گوسفند وجود دارند که از این تعداد، ۱۷ تا ریز RNA حفظ شده و ۶۴ تا ریز RNA جدید بودند. در بافت ماهیچه ۲۵ ریز RNA جدید و ۸ ریز RNA حفظ شده از جمله miR-206، miR-133b، miR-22، miR-30 و miR-376 بیان شده‌اند (جدول ۲).

۳۰ ریز RNA جدید و ۵ ریز RNA حفظ شده نیز هم در بافت عضله و هم در کبد بیان شده‌اند. در مخلوط بافت‌های قلب، کلیه، مغز، تخمدان، پوست، چربی سفید و ریه نیز ۳ ریز RNA جدید به نام‌های oar-miR-9773، oar-miR-9777 و oar-miR-9848 بیان شده است (جدول ۲).

ژن‌های ریز RNA ها پیش‌بینی شده ممکن است در قسمت‌های مختلف ژنوم وجود داشته باشد، برخی واقعی هستند و بیان می‌شوند و برخی دیگر بیان نمی‌شوند اما فقط ریز RNA هایی که بیان شوند (در واقع رونویسی از آنها اتفاق بیافتد) کاربردی هستند و می‌توان در مراحل بعدی عمل آنها را در آزمایشگاه تأیید کرد. در این تحقیق با تأیید بیان به روش *in silico* در وقت و هزینه در مراحل بعدی صرفه‌جویی می‌شود.

شناسایی برخی miRNA های موجود در ژنوم) و FP مثبت کاذب (به معنی معرفی اشتباه قطعه‌ای از ژنوم به عنوان ریز RNA پیش-ساز) را نشان می‌دهد. حساسیت^۳ یعنی درصدی از ریز RNA پیش-ساز موجود را می‌توان با این روش شناسایی کرد و دقت انتخاب^۴ یعنی درصدی از ریز RNA پیش‌ساز پیش‌بینی شده که واقعا ریز RNA هستند (۱۸).

نتایج

در تحقیقات انجام شده مهم‌ترین ویژگی‌های معرفی شده برای ریز RNA های پیش‌ساز شامل: دارا بودن ساختار سنجاق سری با اندازه‌ای مشخص بدون حباب بزرگ، انرژی آزاد کمتر از -20 kcal/mol ، ثبات ترمودینامیکی آن، درصد نوکلئوتیدهای C G و درصد هر کدام از بازها نسبت به کل توالی است (۱۳، ۱۵ و ۲۰). ریز RNA های پیش‌ساز در گونه‌ی گوسفند دارای ساختار سنجاق-سری با شکل مناسب، بدون حباب بزرگ و ویژگی‌های موجود در جدول ۱ بودند.

پیش بینی اولیه و گزینش ریز RNA ها

پس از پیش‌بینی توسط نرم‌افزارهای پیش‌بینی کننده و بررسی ویژگی‌های ریز RNA پیش‌ساز به وسیله 400 mfold ساختار ساقه حلقه به عنوان ریز RNA پیش‌ساز پیشنهاد شد.

در تحقیقات انجام شده به روش از آغاز مانند سور و همکاران (۱۵) و صادقی و همکاران (۱۳) و همچنین لی و همکاران (۱۰) فقط از یک نرم‌افزار استفاده شده است در صورتی که در این تحقیق از مجموعه‌ای از نرم‌افزارهای پیش‌بینی کننده استفاده شد که عامل موثری در کاهش خطای مثبت کاذب است و همچنین در این تحقیق برای اطمینان از وجود خصوصیات pre-miRNA در توالی‌های پیشنهادی توسط نرم‌افزارهای پیش‌بینی کننده، تمام توالی‌ها توسط 400 mfold بررسی شدند و این مرحله در کاهش خطای مثبت کاذب موثر بود.

با بلاست کردن (BLASTx) این ۴۰۰ ریز RNA ها با RNA های کد کننده پروتئین گوسفند، مشخص شد که هیچ کدام از آنها در توالی کدکننده پروتئین نبودند. با بلاست کردن این ۴۰۰ ریز RNA در miRBase مشخص شد که از این بین، ۵۰ ریز RNA حفاظت شده و ۳۵۰ عدد ریز RNA جدید بودند. در تحقیقی که توسط پروزای (۲) انجام شده توالی EST گونه‌ی گوسفند را مورد آنالیز قرار داده است تا ریز RNA های حفظ شده را شناسایی کند در نتیجه ۱۷۲ ریز RNA شناخته شده است از جمله ریز RNA هایی که

³ - Sensitivity

⁴ - Selectivity

جدول ۱- ویژگی ریزRNAهای پیش‌ساز که از ریزRNAهای پیش‌ساز گوسفندی استخراج شده است

Table 1- Pre-miRNAs features that extracted from ovine registered pre-miRNAs

ویژگی ریزRNAهای پیش‌ساز Pre-miRNAs features	دامنه مورد قبول در گوسفند Acceptable range in ovine	دامنه معمول در گوسفند Common range in ovine
انرژی آزاد Free energy (kcal/mol)	Less than -20	-40 – -30
طول ریزRNAهای پیش‌ساز Length of Pre-miRNAs (Nt)	65 – 130	65-90
تعداد بازها در هر ساقه Nucleotides in stem (Nt)	23 – 37	23 – 37
تعداد بازهای حلقه Nucleotide in terminal loop Nt	5 – 30	20
درصد باز A+U A+U%	30 – 70	40 - 60

جدول ۲- ریزRNAهای حفاظت شده و جدید بیان شده در بافت‌های مختلف گوسفند

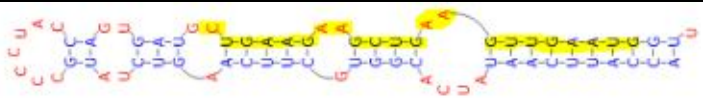
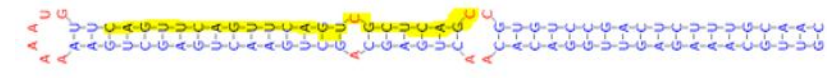
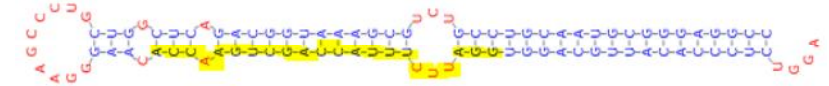
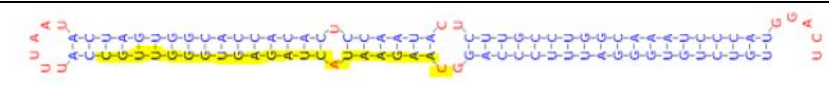

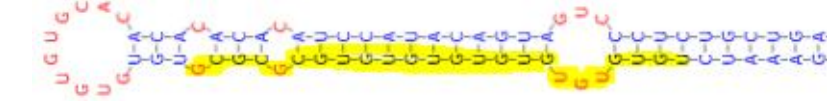
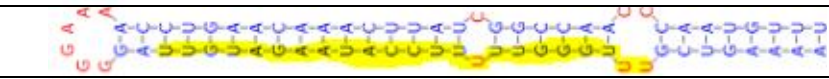
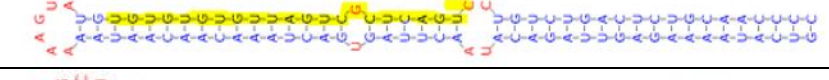
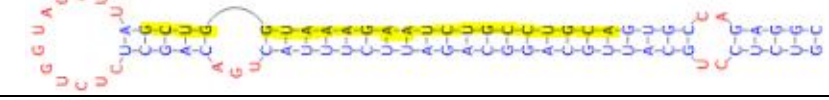
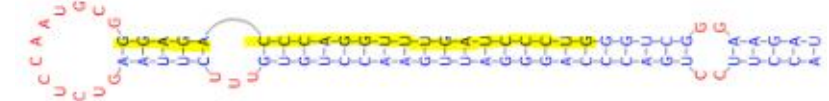
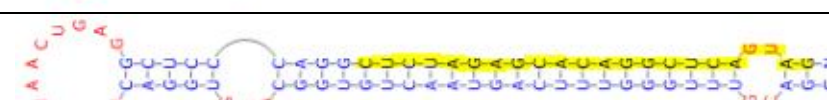
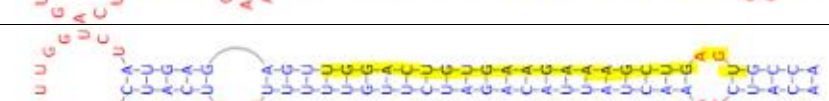
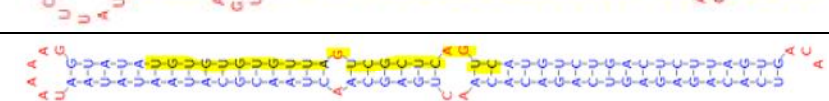
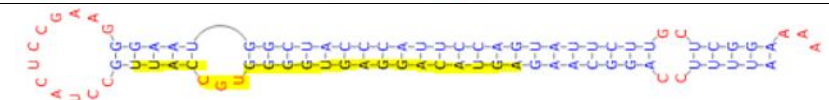

Table 2- Conserved and new miRNAs that were expressed in different ovine tissues

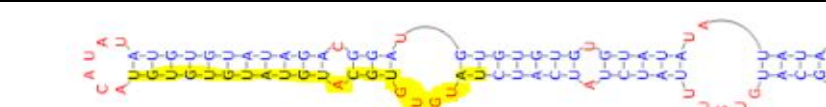
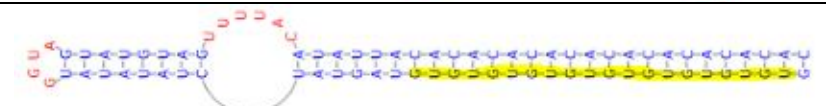
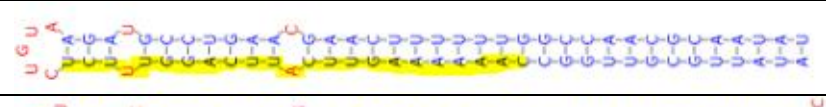


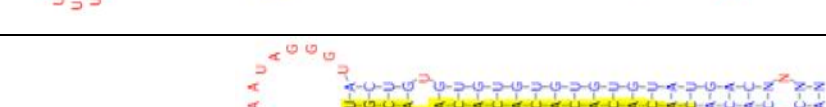

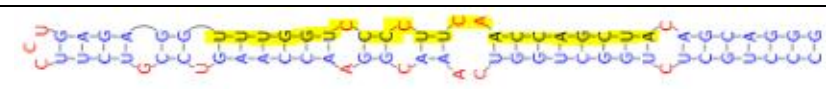
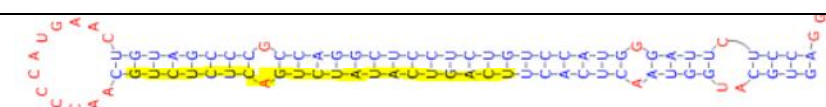
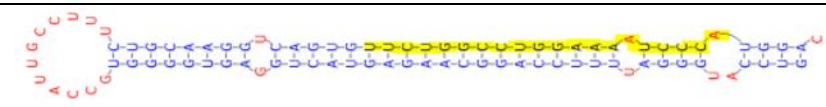
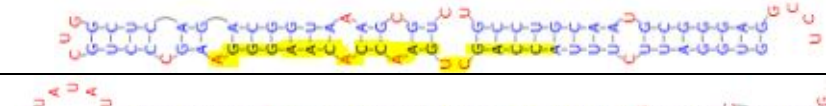

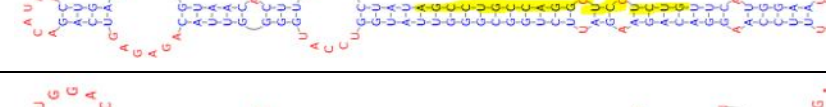


نام ریزRNA MiRNA name	محل توالی روی کروموزوم ۲۰ Coordinates on chromosome # 20	بافت Tissue	ساختار ساقه و حلقه توالی ریزRNA پیش‌ساز و بالغ (رنگ زرد) Stem and loop structure of miRNA precursor and mature (yellow highlight)
Oar-miR-9770-3'	- ۹۵۰۹۱۲۷ ۹۵۰۹۲۲۵	کبد Liver	
Oar-miR-9770-5'	- ۹۵۰۹۱۲۷ ۹۵۰۹۲۲۵	عضله Muscle	
Oar-miR-9771	- ۹۸۱۸۶۴۷ ۹۸۱۸۷۴۶	کبد و عضله Liver & muscle	
Oar-miR-9772	- ۷۰۴۵۲۴۵ ۷۰۴۵۳۱۸	کبد و عضله Liver & muscle	
Oar-miR-9773	- ۱۱۹۰۰۲۰۰ ۱۱۹۰۰۲۹۴	مخلوط بافتها Mixture of tissues	
Oar-miR-9774	- ۱۳۹۱۷۷۱۴ ۱۳۹۱۷۸۰۷	کبد و عضله Liver & muscle	

Oar-miR-9775	- ۱۵۸۶۱۳۸۴ ۱۵۸۶۱۴۶۹	کبد Liver	
Oar-miR-9776	- ۱۶۰۰۸۰۷۹ ۱۶۰۰۸۱۷۴	کبد Liver	
Oar-miR-9777	- ۱۶۳۳۳۱۸۶ ۱۶۳۳۳۲۸۳	مخلوط بافتها Mixture of tissues	
Oar-miR-9778	- ۲۰۶۸۵۶۹۲ ۲۰۶۸۵۷۹۰	کبد و عضله Liver & muscle	
Oar-miR-9779	- ۲۱۶۶۷۱۶۸ ۲۱۶۶۷۲۶۵	کبد و عضله Liver & muscle	
Oar-miR-9780	- ۲۱۸۴۲۷۲۷ ۲۱۸۴۲۸۲۶	کبد و عضله Liver & muscle	
Oar-miR-9781	- ۲۱۸۶۴۴۴۶ ۲۱۸۶۴۴۴۴	کبد و عضله Liver & muscle	
Oar-miR-9782	- ۲۱۹۶۹۵۸۳ ۲۱۹۶۹۶۷۹	کبد و عضله Liver & muscle	
Oar-miR-9783	- ۲۵۲۳۶۴۲۵ ۲۵۲۳۶۵۲۳	کبد و عضله Liver & muscle	
Oar-miR-9784	- ۲۷۴۷۵۶۱۹ ۲۷۴۷۵۷۱۷	کبد و عضله Liver & muscle	
Oar-miR-9785	- ۲۷۹۹۰۵۵۵ ۲۷۹۹۰۶۵۰	کبد و عضله Liver & muscle	
Oar-miR-9786	- ۳۲۶۸۶۴۹۰ ۳۲۶۸۶۵۸۷	کبد و عضله Liver & muscle	

Oar-miR-9787	- ۳۴۰۲۳۲۰۱ ۳۴۰۲۳۲۹۸	کبد و عضله Liver & muscle	
Oar-miR-9788	- ۳۴۰۹۳۷۲۴ ۳۴۰۹۳۸۲۱	کبد Liver	
Oar-miR-9789	- ۳۴۰۲۳۲۱۵ ۳۴۰۲۳۳۱۰	کبد Liver	
Oar-miR-9790	- ۳۵۷۹۶۸۹۱ ۳۵۷۹۶۹۸۹	کبد و عضله Liver & muscle	
Oar-miR-9791-5'	- ۳۷۴۸۹۱۲۶ ۳۷۴۸۹۲۱۶	کبد و عضله Liver & muscle	
Oar-miR-9791-3'	- ۳۷۴۸۹۱۳۱ ۳۷۴۸۹۲۲۹	کبد و عضله Liver & muscle	
Oar-miR-9793	- ۳۷۷۶۳۰۷۲ ۳۷۷۶۳۱۵۹	کبد Liver	
Oar-miR-9794	- ۲۸۷۲۸۱۱۸ ۲۸۷۲۸۲۰۶	کبد و عضله Liver & muscle	
Oar-miR-9795	- ۳۹۱۸۵۳۸۷ ۳۹۱۸۵۴۸۶	کبد و عضله Liver & muscle	
Oar-miR-9796	- ۳۹۸۸۵۸۷۱ ۳۹۸۸۵۹۶۰	کبد و عضله Liver & muscle	
Oar-miR-9797-5' Oar-miR-9797-3'	- ۴۱۹۰۴۶۶۷ ۴۱۹۰۴۷۶۲	کبد و عضله Liver & muscle	
Oar-miR-9798	- ۴۳۳۴۵۳۶۳ ۴۳۳۴۵۴۶۱	کبد و عضله Liver & muscle	
Oar-miR-9799	- ۴۳۷۷۴۹۹۸ ۴۳۷۷۵۰۹۶	کبد و عضله Liver & muscle	

Oar-miR-9800	- ۴۴۰۳۱۰۱۳ ۴۴۰۳۱۰۹۸	کبد و عضله Liver & muscle	
Oar-miR-9801-5' Oar-miR-9801-3'	- ۴۴۱۲۶۷۵۶ ۴۴۱۲۶۸۵۴	کبد و عضله Liver & muscle	
Oar-miR-9802	- ۴۴۱۷۵۰۱۲ ۴۴۱۷۵۰۹۹	کبد و عضله Liver & muscle	
Oar-miR-9803	- ۴۵۰۴۳۳۸۲ ۴۵۰۴۳۴۷۷	کبد و عضله Liver & muscle	
Oar-miR-9804	- ۴۸۹۵۳۳۴۰ ۴۸۹۵۳۴۲۲	مخلوط بافتها Mixture of tissues	
Oar-miR-9805	- ۴۹۱۴۷۳۴۴ ۴۹۱۴۷۴۳۱	کبد و عضله Liver & muscle	
Oar-miR-9806	- ۵۰۷۸۶۷۲۴ ۵۰۷۸۶۸۱۳	کبد و عضله Liver & muscle	
Oar-miR-9807	- ۱۲۵۸۸۱۸۱ ۱۲۵۸۸۲۷۲	عضله Muscle	
Oar-miR-9808	- ۱۴۹۲۲۲۸۵ ۱۴۹۲۲۳۷۸	عضله Muscle	
Oar-miR-9809	- ۱۵۰۴۱۳۸۹ ۱۵۰۴۱۴۸۸	عضله Muscle	
Oar-miR-9810	- ۲۰۶۸۵۷۰۹ ۲۰۶۸۵۸۰۳	عضله Muscle	
Oar-miR-9811	- ۲۱۷۵۸۵۰۹ ۲۱۷۵۸۶۰۷	عضله Muscle	
Oar-miR-9812	- ۱۹۹۶۰۳۱۳ ۱۹۹۶۰۳۹۱	عضله Muscle	
Oar-miR-9813	- ۲۵۳۶۸۲۷۰ ۲۵۳۶۸۴۷۰	عضله Muscle	

Oar-miR-9814	- ۲۶۵۰۴۸۵۷ ۲۶۵۰۴۹۲۸	عضله Muscle	
Oar-miR-9815	- ۳۰۰۸۲۴۵۴ ۳۰۰۸۲۵۴۷	عضله Muscle	
Oar-miR-9816	- ۳۰۱۵۰۲۱۳ ۳۰۱۵۰۳۱۰	عضله Muscle	
Oar-miR-9817	- ۳۱۹۲۵۵۳۰ ۳۱۹۲۵۶۲۸	عضله Muscle	
Oar-miR-9818	- ۳۶۱۸۷۱۷۱ ۳۶۱۸۷۹۶۱	عضله Muscle	
Oar-miR-9819	- ۳۷۴۰۲۶۷۱ ۳۷۴۰۲۷۵۹	عضله Muscle	
Oar-miR-9820	- ۳۸۰۶۵۸۶۶ ۳۸۰۶۵۹۴۸	عضله Muscle	
Oar-miR-9821	- ۳۸۲۷۸۷۱۵ ۳۸۲۷۸۸۰۸	عضله Muscle	
Oar-miR-9822	- ۳۸۶۳۶۲۸۳ ۳۸۶۳۶۳۶۵	عضله Muscle	
Oar-miR-9823	- ۴۱۲۹۹۶۴۰ ۴۱۲۹۹۷۲۲	عضله Muscle	
Oar-miR-9824	- ۴۴۷۰۰۱۵۷ ۴۴۷۰۰۲۳۶	عضله Muscle	
Oar-miR-9825	- ۴۵۹۶۶۷۲۰ ۴۵۹۶۶۸۰۲	عضله Muscle	
Oar-miR-9826	- ۴۶۱۲۰۳۹۱ ۴۶۱۲۰۴۸۷	عضله Muscle	
Oar-miR-9827	- ۴۷۳۲۸۸۹۲ ۴۷۳۲۸۹۷۷	عضله Muscle	
Oar-miR-83	- ۸۲۵۸۶۰ ۸۲۵۹۶۰	کبد و ماهیچه Liver & muscle	

Oar - miR-1187	- ۴۱۱۵۵۷۰ ۴۱۱۵۶۶۹	ماهیچه Muscle	
Oar - miR-466i	- ۲۱۶۰۳۵۶۲ ۲۱۶۰۳۶۴۱	ماهیچه Muscle	
Oar - miR-2285r	- ۷۰۲۵۳۸۴ ۷۰۲۵۴۶۰	ماهیچه Muscle	
Oar - miR-219	- ۷۴۹۶۷۷۳ ۷۴۹۶۸۵۷	کبد Liver	
Oar - miR-2285e	- ۱۴۱۵۵۳۱۴ ۱۴۱۵۵۴۰۲	ماهیچه Muscle	
Oar-- miR-2284d	- ۱۴۱۵۵۳۱۴ ۱۴۱۵۵۴۰۲	کبد و ماهیچه Liver & muscle	
Oar-miR-8485	- ۲۵۸۱۳۲۴۸ ۲۵۸۱۳۳۰۹	کبد و ماهیچه Liver & muscle	
Oar-miR-1322	- ۲۳۶۱۰۱۴۰ ۲۳۶۱۰۲۰۸	کبد Liver	
Oar-miR-133b	- ۲۴۳۵۷۹۱۵ ۲۴۳۵۷۹۹۱	کبد و ماهیچه Liver & muscle	
Oar-miR-6562-3p	- ۲۵۲۴۳۳۴۷ ۲۵۲۴۳۴۶۶	ماهیچه Muscle	
Oar-miR-193	- ۲۹۰۲۰۰۱۶ ۲۹۰۲۰۱۰۵	ماهیچه Muscle	
Oar-miR-4724-3p	- ۳۰۱۵۰۲۳۱ ۳۰۱۵۰۳۱۱	ماهیچه Muscle	
Oar-miR-1692	- ۴۶۶۸۵۱۰۵ ۴۶۶۸۵۲۱۵	کبد Liver	
Oar-miR-596	- ۳۴۳۴۳۳۲۹ ۳۴۳۴۳۴۶۲	کبد Liver	
Oar-miR-2218a-3p	- ۴۰۳۷۱۱۹۶ ۴۰۳۷۱۳۰۰	کبد و ماهیچه Liver & muscle	

منابع

- 1- Abba, M., N. Patil, and H. Allgayer. 2014. MicroRNAs in the Regulation of MMPs and Metastasis. *Cancers*, 6: 625-645.
- 2- Barozai, M. Y. Kh. 2012. The novel 172 sheep (*Ovis aries*) microRNAs and their targets. *Molecular Biology and Reproduction*, 39:6259-6266.
- 3- Bartel, D. P. 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116: 281-297.
- 4- Berezikov, E., E. Cuppen, and R. H. Plasterk. 2006. Approaches to microRNA discovery. *Nature Genetics*, 38 (Suppl: S2) 7.
- 5- Chen, C. Z., L. Li, H. F. Lodish, and D. P. Bartel. 2004. MicroRNAs modulate hematopoietic lineage differentiation. *Science*, 303: 83-86.
- 6- Grundhoff, A., C. S. Sullivan, and D. Ganem. 2006. A combined computational and microarray-based approach identifies novel microRNAs encoded by human gamma-herpesviruses. *RNA*, 12: 733-750.
- 7- Hansen, T., T. V. Morten, K. Jorgen, and K. D. Christian. 2014. miRdentify: high stringency miRNA predictor identifies several novel animal miRNAs. *Nucleic Acids Research*, 42(16): e124.
- 8- Hertel, J., and P. F. Stadler. 2006. Hairpins in a Haystack: Recognizing microRNA Precursors in Comparative Genomics Data. 22(14): e197-202.
- 9- Huang, T. H., B. Fan, M. F. Rothschild, Z. L. Hu, K. Li, K., Zhao, and H. Sh. 2007. MiRFinder: an improved approach and software implementation for genome wide fast microRNA precursor scans. *BMC Bioinformatics*, 8: 341.
- 10- Li, S. C., Pan, C. Y., Lin, and W. Ch. 2006. Bioinformatic discovery of microRNA precursors from human ESTs and introns. *BMC Genomics*, 7:164.
- 11- Miska, E. A., E. Alvarez Saavedra, M. Townsend, A. Yoshii, N. Sestan, P. Rakic, M. Constantine Paton, and H. R. Horvitz. 2004. Microarray analysis of microRNA expression in the developing mammalian brain. *Genome Biol*, 5: R68.
- 12- Rodriguez, A., S. Griffiths Jones, J. L. Ashurst, and A. Bradley. 2004. Identification of mammalian micro RNA host genes and transcription units. *Genome Research*, 14: 1902-1910.
- 13- Sadeghi, B., H. Ahmadi, S. Azimzadeh Jamalkandi, M. R. Nassiri, and A. Masoudi Nejad. 2014. BosFinder: a novel pre microRNA gene prediction algorithm in *Bos taurus*. *Animal Genetics*, 45(4): 479-84.
- 14- SantaLucia, J. Jr. 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proceedings of National Academy of Science USA*, 95: 1460-1465.
- 15- Sewer, A., N. Paul, P. Landgraf, A. Aravin, S. Pfeffer, M. J. Brownstein, Th. Tuschl, E. Nimwegen, and M. Zavolan. 2005. Identification of clustered microRNAs using an *ab initio* prediction method. *BMC Bioinformatics*, 6: 267.
- 16- Sheng, X., X. Song, Y. Yu, L. Niu, Sh. Li, H. Li, C. Wei, T. Liu, L. Zhang, and L. Du. 2011. Characterization of microRNAs from sheep (*Ovis aries*) using computational and experimental analyses. *Molecular Biology of Reproduction*, 38(5): 3161-3171.
- 17- Stark, A., P. Kheradpour, L. Parts, J. Brennecke, E. Hodges, G. J. Hannon, and M. Kellis. (2007). Systematic discovery and characterization of fly microRNAs using 12 *Drosophila* genomes. *Genome Research*, 17: 1865-1879.
- 18- Tempel, S., and F. Tahi. 2012. A fast *ab initio* method for predicting miRNA precursors in genomes. *Nucleic Acids Research*, 40(11): e80.
- 19- Tyagi, S., C. Vaz, V. Gupta, R. Bhatia, S. Maheshwari, A. Srinivasan, and A. Bhattacharya. 2008. CID-miRNA: a web server for prediction of novel miRNA precursors in human genome. *Biochemical and Biophysical Research Communications*, 372: 831-834.
- 20- Wu, Y., B. Wei, H. Liu, and T. Li, and S. Rayner. 2011. MiRPara: a SVM based software tool for prediction of most probable microRNA coding regions in genome scale sequences. *BMC Bioinformatics*, 12: 107.
- 21- Zhao, Y., J. F. Ransom, A. Li, V. Vedantham, M. Drehle, A. N. Muth, T. Tsuchihashi, MT. McManus, R. J. Schwartz, and D. Srivastava. 2007. Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA 1-2. *Cell*, 129: 303-317.



Prediction and *In Silico* Validation of Micro-RNAs in Different Tissues Originated from Ovine Chromosome 20

V. Abbasi¹- M. R. Nasiri²- A. Javadmanesh^{3*}

Received: 05-08-2018

Accepted: 22-09-2018

Introduction MicroRNAs (miRNAs) are small noncoding RNA molecules that are found in plants, animals and some viruses and play important role in regulation of transcription (3). Recently, importance of miRNA roles in biology of living organisms has been discovered, thus miRNAs identification became more important (1). Experimental detection of miRNAs can be obtained using different miRNA profiling methods, such as quantitative real-time PCR, microarray, and high-throughput RNA sequencing technologies. Since the identification and verification of miRNA by laboratory methods is time-consuming and costly (4), for improving miRNA identification and lowering costs, it is more reasonable that miRNA loci predicted by reliable bioinformatic approaches then experimental methods used for confirmation. This may decrease false positive results. Recently, several hundreds of miRNAs reported in variety of species including human and mouse, as well as domestic animals such as cattle, pig, chicken and goat, however there are relatively less information about sheep miRNAs. In this study, we developed a method for prediction and *in silico* validation of miRNAs located on ovine chromosome twenty.

Materials and Methods In this study, an *ab initio* approach was used. The sequence of ovine chromosome 20 was used as input for EMBOSS and Sequence-Structure Motif Base: Pre-miRNA Prediction Webserver applications, then all predicted stem and loop structures were entered into mfold software. Pre-miRNA features for them were calculated and sequences that had these features were selected. Since the probability of miRNA presence in the coding region is very low, miRNAs that were predicted in the coding regions were removed. To confirm the prediction of miRNAs, selected sequences were homology searched within all registered miRNAs in miRBase. In order to evaluate the *in silico* expression of miRNAs, predicted miRNAs were BLASTed against expression data from Sequence Read Archive (SRA) of ovine muscle tissue (Accession: PRJNA223213), liver tissue (Accession: GSM1366318) and mixture of tissues including heart, kidney, brain, liver, ovary, lung, skin, and adipose (Accession: GSE56643). In order to evaluate the accuracy of this method, a positive control region including a cluster of validated miRNAs from ovine chromosome 18 were analyzed by the same method and sensitivity and selectivity of this method were calculated based on this region from chromosome 18.

Results and Discussion After prediction by softwares and investigation of pre-miRNAs features by mfold, 400 stem and loop structures that had pre-miRNA features were chosen. Fifty miRNAs from those miRNAs contained conserved mature miRNAs sequence and 350 of them were recognized as novel miRNAs based on registered miRNAs in the miRBase. None of the novel miRNAs were located in the coding regions. *In silico* validation of these novel miRNAs in SRA data was indicated that 81 miRNAs are expressed in different ovine tissues including 33 in muscle and 10 in liver. Results on the positive control data showed that 40 miRNAs were predicted which majority of them (36 miRNAs) were already validated by experiments. This indicates a high reliability for this method. With putting in sensitivity and selectivity formulas, both of two factors were calculated and it was observed that the sensitivity and selectivity values for our method were 67% and 90%, respectively. Fewer studies accomplished in detection of ovine miRNAs in compare to other farm animals. In previous studies to identify miRNAs in ovine species, mostly laboratory-based or comparative methods were used. This was the first study that used SRA database to check miRNA expression in RNAseq data in order to decrease the discovery of false positive results. Comparing this method with others including CID-miRNA (19), miRPara (20), VMir (6) and miRNAFold (18) methods, we may conclude that this method's sensitivity is less than CID-miRNA, miRPara, miRNAFold and srnaloop. Although selectivity for this method is higher than all above

1, 2 and 3- MSc Graduated, Professor and Assistant Professor of Animal Science department, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran.

(*- Corresponding Author Email: javadmanesh@um.ac.ir)

methods because the false positive in this method is less than other method. This method showed high selectivity and low FP that due to improved prediction method for identify miRNAs.

Conclusion In the current study, predicted ovine miRNAs were validated by an *in silico* method using SRA database that resulted in a higher reliability than other *ab initio* approaches. Although this method is not very fast and fully automated. With running this method on ovine chromosome 20, 81 novel miRNA were predicted which were expressed in different tissues of sheep. This method could be applied on other ovine chromosomes as well as other mammalian species, although future validation by experimental approaches is required.

Keywords: *Ab initio*, Bioinformatics, miRNA, Sheep.