به نام خدا

پیش بینی و تایید *in silico* بیان ژن‌های ریزRNA کروموزوم ۲۰ گوسفند

وحیده عباسی[۱]، علی جوادمنش[۲]، محمدرضا نصیری[۳]
[۱] دانشجوی کارشناسی ارشد ژنتیک و اصلاح‌نژاد دام، گروه علوم دامی دانشگاه فردوسی مشهد
[۲] استادیار ژنتیک و اصلاح‌نژاد دام، گروه علوم دامی دانشگاه فردوسی مشهد
[۳] استاد ژنتیک و اصلاح‌نژاد دام، گروه علوم دامی دانشگاه فردوسی مشهد
آدرس پست الکترونیکی: Vahidehabbasi97@yahoo.com

چکیده

ریزRNAها گروهی از RNAهای غیر کد کننده به طول ۲۲ نوکلئوتیدی هستند که از طریق اتصال به mRNA بیان ژن را تنظیم می‌کنند. این گروه از RNAها دارای وظایف گسترده‌ای هستند از جمله وظایف آنها می‌توان به نقشی که در مراحل رشد و نمو و همچنین در ایمنی موجود زنده ایفا می‌کنند، اشاره کرد. با وجود مطالعات زیادی که در مورد ریزRNAها انجام شده است اما هنوز در مورد ریزRNAهای گونه‌ی گوسفند اطلاعات موجود بسیار کم است. در این مطالعه پس از پیش بینی اولیه ریز RNAو فیلتر کردن نتایج بر اساس معیار های مناسب، نتایج حاصل با اورتولوگها در موجودات مختلف مقایسه و در نهایت بررسی *in silico* بیان ریزRNA های نهایی در بافتهای مختلف بررسی شد. در کل بیش از ۴۰۰ ریزRNA در کروموزوم شماره ۲۰ گوسفند پیش بینی شد که تعداد ۹۷ ریزRNA که بیان آنها در داده‌های ترنسکریپتوم بافتهای مختلف گوسفند بالغ از جمله عضله، کبد، قلب، ریه، بافت چربی، کلیه، مغز، پوست تخمدان با روش *in silico* نیز تأیید شد. روش مورد استفاده در این مطالعه می تواند در پیدا کردن ریزRNA های جدید و تکمیل کردن بانک اطلاعاتی ریزRNA های گوسفند موثر باشد.

کلمات کلیدی: ریزRNAها، بیوانفورماتیکی، گوسفند، تأیید *in silico*

مقدمه

ریزRNAها، RNAهای کوچک غیر کد کننده (ncRNAs) به طول حدود ۲۰ نوکلئوتید هستند که تنظیم بیان ژن پس از رونویسی را به عهده دارند. پیش بینی شده است که حدود ۳۰٪ از ژنهای کد کننده پروتئین انسانی به وسیله‌ی ریزRNAها تنظیم می‌شود. ریزRNAها در کنترل بسیاری از فرآیندهای بیولوژیک بدن مانند آپوپتوز، تکثیر، تمایز و متاستاز نقش مهمی را ایفا می‌کنند (۱،۳،۵). Lin-4 اولین ریزRNA کشف شده در سال ۱۹۹۳ توسط لی و همکاران در کرم نماتود (الگانس) بود آنها کشف کردند که این ژن در مراحل رشد و نمو نقش مهمی دارد (۶).

روش‌های شناسایی ریزRNAها بر دو اساس تقسیم می‌شود اول بر اساس بررسی بیان ریزRNAها با استفاده از روش آزمایشگاهی، که این روش در شناسایی ریزRNA بسیار موثر است اما محدودیت‌هایی دارد از جمله اینکه برخی از ریزRNAها که در سطوح پایین بیان می‌شود و همچنین برخی از ریزRNAها فقط در بافتهای خاص و در سن خاصی بیان می شوند (۸). دوم براساس پیش بینی ساختار ژنی ریزRNA بر اساس روش‌های بیوانفورماتیک (۴)، روش بیوانفورماتیک محدودیت‌های روش آزمایشگاهی را ندارد، این روش ژنوم را مورد تحلیل قرار می‌دهد و جایگاه ژنی ریزRNAها را پیش بینی می‌کند. روش های محاسباتی پیش بینی ریزRNAها خود دو نوع است: الف) با مقایسه‌ی توالی ریزRNAها شناخته شده

در گونه‌ها با توالی ژنوم گونه مورد نظر(۷). این روش محدود به ریزRNAهای حفظ شده است ب) پیش بینی محاسباتی ریزRNA بر اساس ویژگی های ساختار دوم RNA و ثبات ترمودینامیکی ساختار ساقه–حلقه است (۱۰). این روش از لحاظ هزینه خیلی پایین‌تر از روش آزمایشگاهی است، ولی این روش فقط لیستی از ریزRNAهای احتمالی می‌دهد و این سوال مطرح است که آیا ریزRNAهای پیشنهاد شده به این روش واقعا ریزRNA است یا خیر؟

در این تحقیق روش تکمیل شده‌ی بیوانفورماتیکی طراحی شد و استفاده از این به دلیل مقایسه ریزRNAهای پیش بینی شده و بررسی بیان *in silico*، تا حد زیادی خطاهای روش های متداول شناسایی را کاهش می‌دهد. در این تحقیق از کروموزوم شماره ۲۰ گوسفند استفاده گردید.

مواد و روشها:

استخراج خصوصیات:

چون اساس این روش شناسایی ریزRNAهای پیش‌ساز از بین ساختارهای سنجاق‌سری موجود در ژنوم است ابتدا باید دید ریزRNAهای پیش‌ساز چه خصوصیاتی دارند در تحقیقات انجام شده مهمترین ویژگی‌های معرفی شده ریزRNAهای پیش‌ساز شامل ساختار سنجاق سری، انرژی آزاد کمتر از ۲۰- ، اندازه توالی سنجاق سری، ثبات ترمودینامیکی آن، همچنین درصد نوکلئوتیدهای C G و ویژگی‌هایی مانند نداشتن حباب بزرگ و درصد هر کدام از بازها نسبت به کل توالی است. (۹).

جمع‌آوری داده‌ها:

کروموزوم ۲۰ گوسفند به عنوان داده‌های مورد آنالیز انتخاب شد و توالی آن از سایت NCBI اخذ شد (Accession number: NW_004080183.1).

برای اطمینان از صحیح بودن مراحل روش کار و همچنین کاربردی بودن نرم‌افزارهای به کار گرفته شده از داده‌های شاهد مثبت استفاده شد. در این مطالعه ناحیه‌ای بین ۶۴۴۶۲۳۰۰ تا ۶۴۶۶۵۰۲۰ از کروموزوم ۱۸ گوسفند که دارای ۵۴ ریزRNA شناخته شده و ثبت شده در miRBase است، به عنوان شاهد مثبت انتخاب شد و توالی آن از NCBI (Accession number: NW_004080181.1) دانلود شد.

طراحی روش مناسب:

روش استفاده شده در این تحقیق مراحل روش بر پایه ساختار را دارد اما با یک تغیراتی که در جهت تکمیل و کاهش خطاهای موجود در آن اعمال شده است.

پیش‌بینی ریزRNAهای پیش‌ساز و بالغ:

۱- با استفاده از نرم‌افزار EMBOSS e-inverted ژنوم به قطعاتی به طول حدود ۲۰۰-۱۰۰ نوکلئوتیدی تقسیم شدو برای هر قطعه ساختار دوم رسم شد؛ ۲- توالی‌های خروجی از نرم‌افزار EMBOSS e-inverted، وارد نرم‌افزار mfold شد با استفاده از این نرم‌افزار امکان وجود ویژگی‌های ریزRNAهای پیش‌ساز را در توالی مورد نظر محاسبه شد در نهایت قطعات دارای این ویژگی ها به عنوان ریزRNAهای پیش‌ساز پیشنهاد شد. چون این قسمت از کار نیاز به حساسیت بالا دارد و امکان خطا در آن وجود دارد، داده‌های گزینش شده از فیلتر نرم‌افزارهای پیش بینی کننده‌ی دیگری عبور داده شد؛ وب سرور پیش بینی

ریزRNAهای بر پایه‌ی توالی موتیف (Sequence-Structure Motif Base: Pre-miRNA Prediction Webserver) استفاده شد

۳- سپس بخشی از ساقه‌ی ریزRNAهای پیش‌ساز پیشنهاد شده، توسط سرور نامبرده به عنوان ریزRNAهای بالغ معرفی شد.

حذف توالی کد کننده پروتئین:

ریزRNAهایی که در بخش اگزون ژن کد کننده‌ی پروتئین هستند کنار گذاشته می‌شوند؛ برای این منظور ریزRNAهای بیان شده، با توالی کد کننده پروتئین بلاست شد و توای‌هایی که در بین توالی کد کننده پروتئین نبودند به عنوان ریزRNA معرفی کرد.

شناسایی ریزRNAهای حفظ شده:

توالی ریزRNAهای پیش‌ساز پیشنهاد شده در قسمت قبل در قسمت جستجوی miRBase با E-value cutoff برابر ۱۰، در تمامی گونه‌ها بلاست شد. در صورتی که در توالی وارد شده مشابه با توالی ریزRNA های شناخته شده وجود داشته باشد به عنوان ریزRNA شناخته شده معرفی می‌شود.

تأیید *in silico*:

برای اطمینان از بیان شدن ریزRNAها، همه ریزRNAهای پیش بینی شده در داده‌های ترنسکریپتوم بافت عضله (با Accession: PRJNA223213)، بافت کبد (با Accession: GSM1366318) و همچنین در مخلوط ترنسکریپتوم‌های بافتهای قلب، کلیه، مغز، کبد، تخمدان، پوست، چربی سفید و ریه (با Accession: GSE56643) بلاست شد.

نتایج و بحث

استخراج خصوصیات:

ریزRNAهای پیش‌ساز مربوط به گونه‌ی گوسفند از miRBase دانلود شد با استفاده از نرم‌افزار mfold مشخص شد ریزRNAهای پیشساز دارای ساختار سنجاق‌سری با شکل مناسب، بدون حباب بزرگ و ویژگی‌های موجود در جدول هستند؛

خصوصیات ریزRNAهای پیشساز	دامنه مورد قبول در این تحقیق	دامنه معمول
انرژی آزاد	کمتر یا مساوی ۲۰-	۴۰- تا ۳۰-
طول	۶۵-۱۳۰ نوکلئوتید	۶۵-۹۰ نوکلئوتید
تعداد بازها در هر ساقه	۲۳-۳۷	۲۳-۳۷
تعداد بازهای حلقه	۵-۳۰	۲۰
درصد بازA+U	۳۰-۷۰	۴۰-۶۰

پیش بینی:

با اعمال روش طراحی شده در این مطالعه در مرحله اول حدود ۴۰۰ ریزRNA در کروموزوم ۲۰ گوسفند پیشنهاد داده شده است که از این بین ۱۰۵ تا از آنها در بین گونه‌های مختلف جانوری حفظ شده بودند و ۲۹۵ تا ریزRNA جدید بودند.

با بلاست کردن این ۴۰۰ ریزRNAها با RNAهای کد کننده پروتئین، مشخص شد که هیچ کدام از آنها مشابه با RNAهای کد کننده پروتئین نیستند

تایید *in silico*:

از ۴۰۰ ریزRNA پیشنهادی تعداد ۹۷ تا در داده‌های ترنسکریپتوم بیان شده بودند که ۳۰ تا ریزRNA حفظ شده و ۶۷تا ریز RNA جدید بودند.

در بافت ماهیچه: در ریزRNA های جدید ۲۷ ریزRNA فقط در بافت ماهیچه بیان شده‌اند و ۳۱ ریزRNA هم در بافت عضله و هم در کبد بیان شده‌اند. در ریزRNAهای حفظ شده ۱۰ ریزRNA فقط در بافت عضله و ۱۲ ریزRNA هم در عضله و هم در کبد موجود است. در بافت کبد: علاوه بر ۳۱ ریزRNA مشترک بین کبد و عضله ۶ ریزRNA فقط در بافت کبد بیان شده است و ۷ ریزRNA حفظ شده نیز فقط در بافت کبد وجود دارند. در مخلوط بافتهای قلب، کلیه، مغز، تخمدان، پوست، چربی سفید و ریه: ۴ ریزRNA جدید بیان شده است.

بروزای در سال ۲۰۱۲، ۱۷۲ ریزRNA به روش مقایسه‌ای در گوسفند پیش بینی کردند که همگی جزو ریزRNAهای حفظ شده بودند که از قبل در گونه‌های دیگر کشف و تایید شده بودند (۲). در واقع در گوسفند به روش بیوانفورماتیکی بر پایه ساختار کار نشده است اما در سایر گونه‌های کار شده است که ابتدا نرم‌افزاری بر مبنای ژنوم یک گونه‌ی خاص جهت پیش بینی ریزRNA برنامه نویسی شد که ممکن است در صورت استفاده برای گونه‌های دیگر دقت کافی نداشته باشد و از طرفی ریزRNAهای پیش بینی شده فقط خروجی یک نرم‌افزار است و فقط در صورت تأیید آزمایشگاهی می‌تواند به عنوان ریزRNA واقعی معرفی گردد (۷، ۹).

با اعمال روش طراحی شده در این مطالعه در ناحیه‌ی بین نوکلئوتید شماره‌ی ۶۴۴۶۲۳۰۰ تا ۶۴۶۶۵۰۲۰ کروموزوم ۱۸گوسفند، اکثر ریزRNAهای ثبت شده در پایگاه miRBase شناسایی شدند و نکته مهمتر این است که هیچ RNA دیگری به عنوان ریزRNA پیدا نشد و این نشان دهنده‌ی دقت روش طراحی شده در این تحقیق بود.

فهرست منابع:

(1) **Ambros V. (2004)** The functions of animal microRNAs. Nature.;431(7006):350-5

(2) **Barozai, M.Y.Kh**. (**2012**). The novel 172 sheep (Ovis aries) microRNAs and their targets. Mol Biol Rep 39:6259–6266

(3) **Bartel DP. (2004)** MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. ;116(2):281-97

(4) **Berezikov, E., Cuppen, E., Plasterk, R. (2006)** Approaches to microRNA discovery. NATURE GENETICS 38

(5) Filipowicz W, Bhattacharyya SN, Sonenberg N. (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?. Nat Rev Genet.;9(2):102-14

(6) **Lee, R.C., Feinbaum, R.L. and Ambros, V. (1993)** The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75, 843–854.

(7) **Hertel, J., Stadler, P. (2006)** Hairpins in a Haystack: Recognizing microRNA Precursors in Comparative Genomics Data. ISMB'06

(8) **Sadeghi, B., Nassiri, M., Masoudi-Nejad, A., Tahmoorespour, M., Dehghani, H., Ahmadi, H. (2013)** Computational Analyses for Identification Novel MicroRNAs from Cattle and Sheep. *Cell and Molecular Research; 4 (2), 62-67*

(9) **Sadeghi, B., Ahmadi, H., Azimzadeh-Jamalkandi, S., Nassiri, M. R., Masoudi-Nejad, A. (2014)** .BosFinder: a novell pre-microRNA gene prediction algorithm in Bos Taurus. Animal Genetics

(10) **Sheng, X., Song, X., Yu, Y., Niu, L., Li, Sh., Li, H., Wei, C., Liu, T., Zhang, L., Du, L. (2011)** . Characterization of microRNAs from sheep (Ovis aries) using computational and experimental analyses . Mol Biol Rep 38:3161–3171

Prediction and *in silico* validation of expressed micro-RNAs in ovine chromosomes 20

Vahideh Abbasi[1], Ali Javadmanesh[2], Mohammadreza Nassiri[3]
Master of Animal Genetics Student, Assistant Professor and Professor, Department of Animal Science, Faculty of Agriculture, Ferdowsi University of Mashhad
Email: Vahidehabbasi97@yahoo.com

Abstract
Micro-RNAs are a group of non-coding RNAs of 21-23 nucleotides in length that regulate gene expression by binding to mRNA. This group of RNAs has extensive functions, for example in the organism's growth and development and also immune system. Despite the vast number of studies that have been conducted on micro-RNAs, there is a big gap of knowledge in ovine micro-RNA data base. In this study, after prediction of miRNAs and filtration based on features, results compared with orthologues in other organisms and finally the expression of selected miRNAs approved by in silico validation. In total, more than 400 miRNAs were predicted in chromosome 20 of sheep which of those, expression of 97 miRNA were validated by *in silico* method in muscle, liver, heart, lung, adipose tissue, kidney, brain, skin and ovary tissues. This new method of validation can be used to complete the data set of sheep miRNA.
Key words: micro RNA, bioinformatics, ovine, *in silico* validation