

بررسی جامع ژن‌ها و eRNAهای درگیر در سرطان معده و ساخت مدل پیش‌آگهی برای بقای کلی بیماران

بصیره بهرامی^۱، پروانه نیک‌پور^۲

مقاله پژوهشی

چکیده

مقدمه: سرطان معده، یکی از شایع‌ترین سرطان‌ها و اولین علت مرگ ناشی از سرطان‌ها در ایران است. به دلیل دانش ناکافی در مورد عناصر مولکولی دخیل در بیماری برای معرفی زیست‌نشانگرهای تشخیصی / پیش‌آگهی‌دهنده، بیماران مبتلا در مراحل بالاتر تشخیص داده می‌شوند که این امر منجر به پیش‌آگهی ضعیف این بیماران می‌شود. هدف مطالعه‌ی حاضر، بررسی جامع enhancer RNAs (eRNAs) به منظور تعیین ارتباطات مولکولی این عناصر با سایر ژن‌ها و ساخت یک مدل مرتبط با بقای بیماران بود.

روش‌ها: اطلاعات خام توالی‌یابی رونوشت‌ها در نمونه‌های سرطان معده از پایگاه داده‌ی اطلس ژنوم سرطان اخذ و eRNAها و ژن‌های با بیان متمایز در نمونه‌های توموری نسبت به نمونه‌های غیرتوموری استخراج شد. سپس ژن‌های هدف eRNA با بیان متمایز بر اساس فاصله‌ی فیزیکی و همبستگی تعیین شدند و شبکه‌ی تنظیمی با استفاده از این عناصر ترسیم شد. در نهایت با انجام تحلیل رگرسیون کاکس تک‌متغیره و چندمتغیره، یک مدل پیش‌آگهی‌دهنده برای پیش‌بینی بقای کلی بیماران ساخته و کارایی مدل بررسی شد.

یافته‌ها: با انجام تحلیل تمایز بیانی، ۶۹ eRNA و ۲۶۰۶ ژن دارای بیان متمایز استخراج شدند و با تعیین ارتباطات بین آن‌ها، یک شبکه‌ی تنظیمی متشکل از ۸۴ گره و ۱۱۹ یال ترسیم شد. در نهایت یک مدل سه جزئی مرتبط با بقا ساخته شد که قادر به پیش‌بینی بقای بیماران بود.

نتیجه‌گیری: نتایج این مطالعه نشان داد، مدل سه جزئی ساخته شده شامل یک eRNA و دو ژن دیگر می‌تواند به عنوان ابزار پیش‌آگهی‌دهنده در سرطان معده مورد استفاده قرار گیرد؛ اگرچه مطالعات گسترده‌تری برای کاربرد بالینی این مدل لازم است.

واژگان کلیدی: رونوشت‌های نسخه‌برداری شده از افزایشنده‌ها؛ سرطان‌های معده؛ تحلیل بقا؛ زیست‌شناسی سامانه‌ای؛ توالی‌یابی RNA

ارجاع: بهرامی بصیره، نیک‌پور پروانه. بررسی جامع ژن‌ها و eRNAهای درگیر در سرطان معده و ساخت مدل پیش‌آگهی برای بقای کلی

بیماران. مجله دانشکده پزشکی اصفهان ۱۴۰۲؛ ۴۱ (۷۲۴): ۴۹۵-۴۸۸

مقدمه

سرطان معده، یکی از کشنده‌ترین سرطان‌ها می‌باشد که در ایران از نظر بروز در رتبه‌ی سوم و از نظر مرگ‌ومیر ناشی از سرطان‌ها، در رتبه‌ی اول قرار دارد (۱). دانش مولکولی در رابطه با این سرطان کم است و به دلیل نبود زیست‌نشانگرهایی برای تشخیص زودهنگام و پیش‌آگهی بیماری، اغلب سرطان در مراحل بالاتر و همراه با متاستاز تشخیص داده می‌شود و نمی‌توان بیماران را درمان کرد. با شناخت عوامل و مسیرهای مولکولی دخیل در ایجاد و پیشرفت سرطان معده و بررسی ارتباط این عوامل با میزان بقای بیماران، می‌توان بیماری را در مراحل آغازین شناسایی و یا در جهت درمان به‌موقع و مؤثر آن اقدام کرد (۲).

یک دسته از عوامل مولکولی که به‌تازگی مشخص شده با پیشرفت سرطان‌ها در ارتباط هستند، enhancer RNAs (eRNAs) می‌باشند که جزئی از RNAهای غیرکدکننده‌ی بلند (lncRNAs) هستند (۳). این دسته از RNAها که از عناصر افزایشنده (Enhancers) رونوشت‌برداری می‌شوند، مستقل از جهت و فاصله، می‌توانند با توالی‌های پیش‌بر (Promoters) برهم‌کنش داشته و با تنظیم فعالیت کمپلکس‌های رونویسی، بیان ژن‌های هدف را متأثر سازند (۳). در مطالعه‌ای در سال ۲۰۱۹، eRNAهای مرتبط با بقا در نمونه‌های سرطان‌های سر و گردن از پایگاه داده‌ی اطلس ژنوم سرطان (TCGA) (۴) استخراج شدند و از میان آن‌ها یک eRNA با

۱- کارشناسی ارشد، گروه ژنتیک و بیولوژی مولکولی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران

۲- دانشیار، گروه ژنتیک و بیولوژی مولکولی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران

نویسنده‌ی مسؤول: پروانه نیک‌پور؛ دانشیار، گروه ژنتیک و بیولوژی مولکولی، دانشکده‌ی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران

Email: pnikipour@med.mui.ac.ir

این eRNA با مسیرهای آنکوژنی همچون چرخه‌ی سلولی و رونویسی را نیز مشاهده کردند (۱۰).

اگرچه مطالعاتی بر روی eRNA های مرتبط با سرطان معده انجام گرفته است، ولی تاکنون مطالعه‌ای که با استفاده از جایگاه‌های افزایشدهی شناخته شده، به بررسی جامع بیان eRNA ها پردازد و علاوه بر شناسایی اهداف ژنی آن‌ها، مدل پیش‌آگهی مرتبط با بقای کلی بیماران مبتلا به سرطان معده را معرفی کند، انجام نگرفته است. بنابراین، در مطالعه‌ی حاضر، با بررسی اطلاعات بیان eRNA ها، بیان ژن‌های مرتبط با آن‌ها و اطلاعات بالینی بیماران، مدلی مرتبط با بقای بیماران ساخته شد که قادر به جداسازی بیماران مبتلا به سرطان معده با نرخ بقای متفاوت بود.

روش‌ها

این مطالعه در سال‌های ۱۴۰۰-۱۴۰۱ در گروه ژنتیک و بیولوژی مولکولی دانشگاه اصفهان با استفاده از داده‌های سرطان معده در پایگاه TCGA انجام گرفت. ابتدا قراردادی با مرکز NIH (The National Institutes of Health) آمریکا برای اخذ مجوز دسترسی به داده‌های خام موجود در این پایگاه منعقد گردید (۱۱). سپس داده‌های خام توالی‌یابی RNA و اطلاعات بالینی ۳۷۳ نمونه شامل ۳۴۳ نمونه‌ی توموری و ۳۰ نمونه‌ی غیرتوموری دریافت شدند. برای تحلیل داده‌های بیان eRNA ها، فایل‌های تحت عنوان فایل (General Feature Format) GFF که حاوی موقعیت‌هایی است که eRNA ها می‌توانند بیان داشته باشند، با استفاده از داده‌های مطالعه‌ی Zhang و همکاران، با کمک ابزار Galaxy ساخته شد (۶، ۱۲). فایل GFF مربوط به ژن‌های شناخته شده از پایگاه Ensembl اخذ گردید (۱۳). در ادامه، تحلیل جامع بیان eRNA ها و ژن‌ها، شامل استخراج مقادیر بیان مربوط به هر ژن، نرمال‌سازی مقادیر و استخراج eRNA ها و ژن‌های با بیان متمایز در نمونه‌های توموری نسبت به نمونه‌های غیرتوموری (با معیارهای $|\log_2 \text{fold change}| > 2$ و $\text{false discovery rate} < 0.01$)، با استفاده از ابزارهای موجود در محیط لینوکس و زبان برنامه‌نویسی R انجام گرفت (۱۴-۱۶). پس از تعیین ژن‌های هدف هر eRNA بر اساس همبستگی بیان بیشتر از ۰/۴ و حداکثر یک میلیون جفت باز فاصله از منطقه‌ی میانی eRNA با نقطه‌ی شروع رونویسی ژن هدف، شبکه‌ی تنظیمی متشکل از eRNA ها و ژن‌های هدف آن‌ها با کمک نرم‌افزار Cytoscape ترسیم گردید (۱۷). در ادامه، برای شناسایی eRNA ها و ژن‌های مرتبط با بقا، ۳۳۵ نمونه‌ی توموری سرطان معده دارای اطلاعات بقا به صورت تصادفی به مجموعه داده‌های آموزشی (Training) و تأییدی (Validation) با نسبت ۳۰:۷۰ تقسیم شدند.

نام *AP001056.1* که دارای بیشترین ارتباط با بقای بیماران بود، به عنوان کاندید مرحله‌ی آزمایشگاهی انتخاب شد که افزایش بیان *AP001056.1* به همراه کاهش بیان ژن هدف آن در نمونه‌های توموری نسبت به نمونه‌های غیرتوموری زبان مشاهده شد (۵). پس از آن Zhang و همکاران در سال ۲۰۱۹ با اخذ همه‌ی نمونه‌های موجود در پایگاه TCGA، بررسی جامعی در ارتباط با eRNA انجام دادند. آن‌ها بر اساس جایگاه افزایشدهی‌ها و تخمین مناطقی که eRNA می‌تواند بیان داشته باشند، مجموعاً نه هزار و صد و هشت eRNA بیان‌یافته در سرطان‌های مختلف را شناسایی کردند که بیشترین تعداد آن در سرطان معده (۲۲۶۷ عدد) می‌باشد. در نهایت یک eRNA که بیشترین تغییر بیان را در نمونه‌های سرطان پستان داشت با نام *ENSR00000023843* برای بررسی‌های آزمایشگاهی انتخاب شد. نتایج این مطالعه نشان داد بیان این eRNA با تکثیر سلولی در رده‌ی سلولی سرطان پستان ارتباط دارد (۶).

در یکی دیگر از مطالعات انجام‌گرفته بر روی نمونه‌های پروستات پایگاه TCGA در سال ۲۰۲۱، نشان داده شد که سی و هفت eRNA با بقای کلی بیماران در ارتباط هستند که *PARGPI* با عود مجدد بیماری در ارتباط بود (۷). در مطالعه‌ی دیگری که در سال ۲۰۲۱ بر روی ۱۹۱ نمونه‌ی گلیوبلاستوما انجام شد، هفتاد eRNA به عنوان eRNA های مرتبط با بقا شناسایی شدند که *AC003092.1* به عنوان eRNA کاندید انتخاب شد. در ادامه، نتایج بررسی بیان آن در ۱۱ نمونه گلیوما با درجه‌ی بالا و ۷ رده‌ی سلولی مرتبط با گلیوما نشان داد، همبستگی مثبت بین بیان *AC003092.1* و ژن هدفش، *TFPI2*، وجود دارد (۸).

همچنین در مطالعه‌ای در سال ۲۰۲۰، ارتباط بیان *WAKMAR2* با بقای بیماران مبتلا به سرطان معده، ویژگی‌های کلینیکوپاتولوژیک و ویژگی‌های تومور بررسی شد. نتایج نشان داد که در مقایسه با افراد با بیان بیشتر *WAKMAR2*، افراد با بیان کمتر این eRNA میزان بقای کمتری داشتند. محققان این مطالعه نشان دادند، بیان *WAKMAR2* با ویژگی‌هایی همچون مراحل تومور، درجه‌ی تومور، میزان بار جهش تومور و ناپایداری ریزماهواره‌ها ارتباط معنی‌دار دارد (۹). در مطالعه‌ی دیگر در سال ۲۰۲۲، اطلاعات ۴۰۷ نمونه‌ی توموری و ۳۲ نمونه‌ی غیرتوموری سرطان معده با باریگری شد و eRNA های مرتبط با سرطان معده که ۲۳ عدد بودند، استخراج شدند. سپس، *CDK6-ASI* به عنوان eRNA ی با معنی‌دارترین ارتباط با بقا انتخاب شد. علاوه بر افزایش بیان *CDK6-ASI* در نمونه‌های توموری نسبت به نمونه‌های غیرتوموری، نتایج نشان داد افراد با بیان بالاتر *CDK6-ASI*، نرخ بقای کمتری نسبت به گروه با بیان پایین‌تر دارند و بیان *CDK6-ASI* با سن زیر ۶۰ سال و مراحل تومور در ارتباط است. محققان ارتباط

شناسایی شدند که شبکه‌ی تنظیمی ترسیم شده برای آن‌ها متشکل از ۸۴ گره (node) و ۱۱۹ یال (edge) در شکل ۲ نمایش داده شده است (شکل ۲).

در مرحله‌ی بعد، نمونه‌های دارای اطلاعات بقا به دو مجموعه داده‌ی آموزشی با ۲۳۴ نمونه و مجموعه داده‌ی تأییدی با ۱۰۱ نمونه تقسیم شدند. تحلیل‌های کاکس تک متغیره نشان داد، دو eRNA و دو ژن از شبکه‌ی ترسیم شده دارای ارتباط معنی‌دار با بقای کلی بیماران هستند. در نهایت، با استفاده از این چهار ژن، مدل سه‌جزئی مرتبط با بقای بیماران (مطابق فرمول ۱) در مجموعه داده‌ی آموزشی ساخته و در مجموعه داده‌ی تأییدی به تأیید رسید.

فرمول ۱. مدل مرتبط با بقاء ساخته شده با استفاده از eRNA و ژن‌های موجود در شبکه‌ی تنظیمی

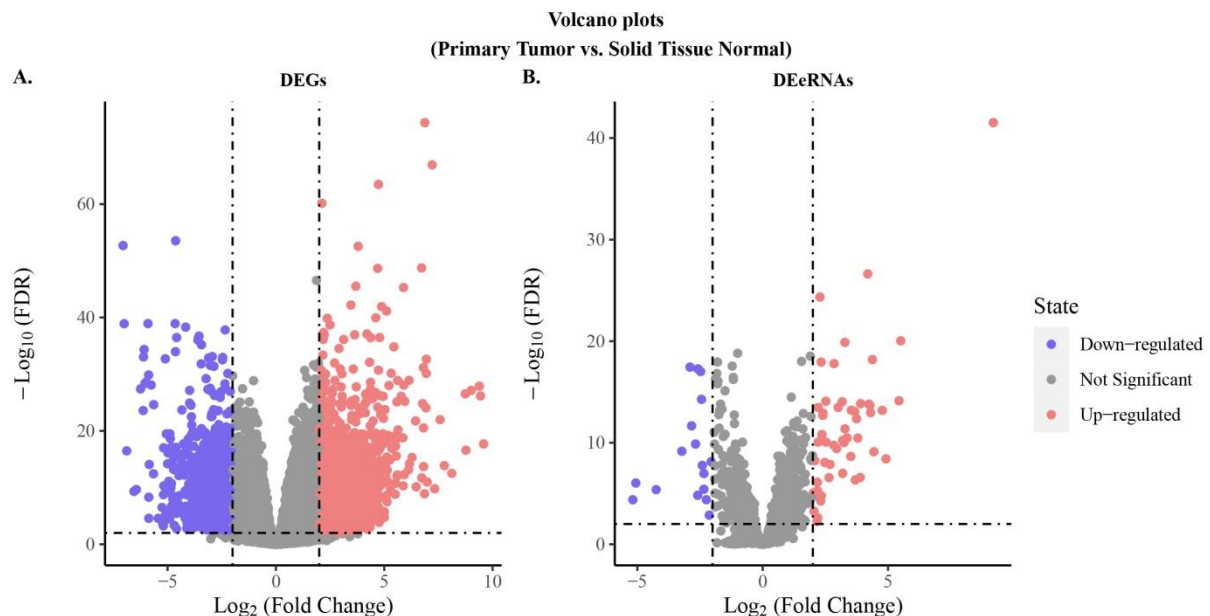
$$\text{Risk score} = (-0.6745 * \text{ENSR00000223532}) - (0.5234 * \text{MIR4435-2HG}) - (0.3877 * \text{MYO16-AS1})$$

نتایج این مرحله نشان داد افراد با خطر بیشتر (با Risk score بالاتر) دارای نرخ بقای کمتری نسبت به افراد با خطر کمتر (با Risk score کمتر) هستند و مدل ساخته شده، قادر به پیش‌بینی بقای بیماران در مجموعه داده‌ی آموزشی و تأییدی است (شکل ۳).

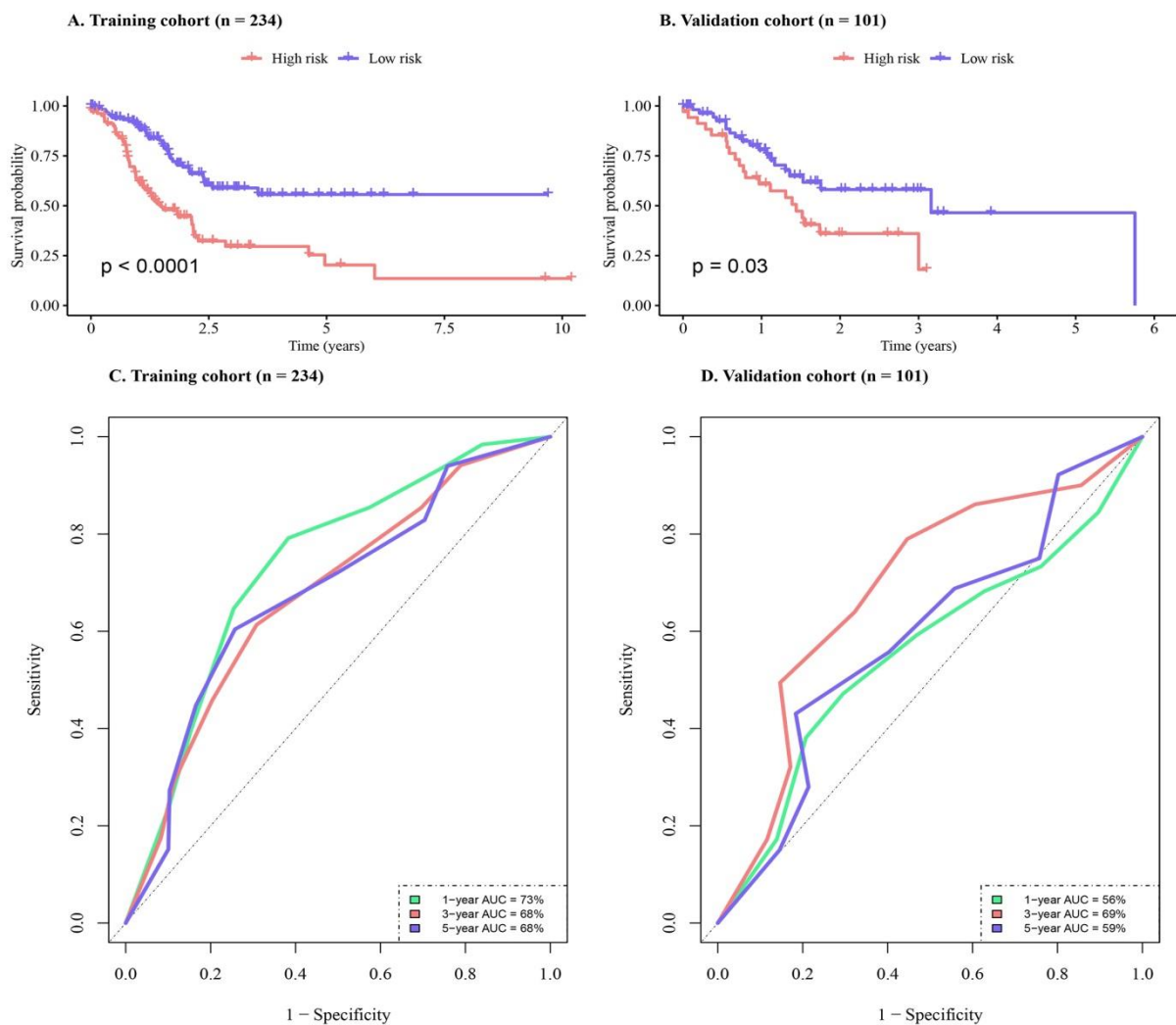
پس از انجام آزمون رگرسیون کاکس تک‌متغیره و انتخاب eRNA و ژن‌های دارای ارتباط معنی‌دار با بقا ($P < 0.05$)، مدل مرتبط با بقا برای پیش‌آگهی عاقبت بیماران در مجموعه داده‌ی آموزشی با استفاده از آزمون کاکس چند متغیره ساخته و در مجموعه داده‌ی تأییدی بررسی شد. برای سنجش کارایی مدل پیشنهادی، بیماران بر اساس مدل ساخته شده، به دو گروه با خطر زیاد و با خطر کم تقسیم شدند و نمودارهای کاپلان-مایر (Kaplan-Meier) و منحنی‌های مشخصه‌ی عملکرد سیستم (ROC curve (Receiver Operating Characteristics curve) برای مجموعه داده‌ی آموزشی و تأییدی رسم گردید (۱۸-۲۰). تمام تحلیل‌های صورت گرفته در مطالعه‌ی پیش‌رو با استفاده از نرم‌افزار R، با حداقل سطح معنی‌داری کمتر از ۰/۰۵ و ترسیم نمودارهای مربوطه با استفاده از بسته‌ی نرم‌افزاری ggplot2 انجام گرفت (۱۶، ۲۱).

یافته‌ها

در مرحله‌ی اول مطالعه، نتایج نشان داد از میان ۲۲۶۶۷ eRNA و ۳۶۷۵۸ ژن دارای بیان، ۶۹ eRNA و ۲۶۰۶ ژن (شامل mRNA و lncRNA) دارای بیان متمایز در نمونه‌های توموری نسبت به نمونه‌های غیرتوموری معده بودند (شکل ۱). از این میان، تعداد ۴۴ ژن به‌عنوان اهداف ژنی eRNA ۴۰



شکل ۱. نمودارهای آتشفشانی نشان‌دهنده‌ی eRNA و ژن‌های با بیان متمایز در نمونه‌های توموری نسبت به نمونه‌های غیرتوموری معده. در سمت چپ، نمودار ژن‌های با بیان متمایز (A) DEGs (Differentially-expressed genes) و در سمت راست، نمودار eRNAهای با بیان متمایز (B) DEeRNAs (Differentially-expressed eRNAs) با محورهای افقی که نشان‌دهنده‌ی میزان تغییر بیان و محورهای عمودی که بیانگر میزان FDR (False discovery rate) هستند، به تصویر کشیده شده‌اند. نقاط قرمز و آبی به ترتیب eRNA و ژن‌های افزایش و کاهش بیان یافته را نشان می‌دهند.



شکل ۳. بررسی کارایی مدل سه‌جزئی ساخته شده در مجموعه داده‌های آموزشی و تأییدی. در ابتدا نمونه‌ها بر اساس فرمول ۱ به دو گروه با خطر زیاد (با Risk score بالاتر) و با خطر کم (با Risk score پایین‌تر) تقسیم شدند. سپس نمودارهای کاپلان-مایر برای آن‌ها در مجموعه داده‌های آموزشی (A) و مجموعه داده‌های تأییدی (B) ترسیم شد که در هر دو مجموعه داده، افراد با خطر زیاد، نرخ مرگ‌ومیر بیشتری داشته‌اند. همچنین منحنی مشخصه عملکرد سیستم برای مجموعه داده‌ها در یک-سه- و پنج سال ترسیم شد که سطح زیر نمودارها در مجموعه داده‌های آموزشی (C) و تأییدی (D) بیانگر عملکرد قابل قبول مدل در پیش‌آگهی عاقبت بیماران مبتلا به سرطان معده است.

در نهایت، با انجام آزمون رگرسیون کاکس چند متغیره، یک مدل سه‌جزئی شامل *MIR4435-2HG*, *ENSR00000223532* و *MYO16-AS1* برای پیش‌آگهی بیماران مبتلا به سرطان معده ساخته شد. عنصر eRNA موجود در این مدل پیشنهادی، *ENSR00000223532*، دارای سه ژن هدف (*LINC01605*, *ENSG00000254290* و *ENSG00000254111*) در شبکه‌ی تنظیمی است که همگی در نمونه‌های توموری معده نسبت به نمونه‌های غیرتوموری افزایش بیان نشان داده‌اند. در مطالعه‌ی بر روی نمونه‌های سرطان معده در پایگاه TCGA نشان داده شده است که بیان ژن *ENSG00000254290* دچار افزایش بیان شده که تأییدکننده‌ی نتایج مطالعه‌ی حاضر بود (۲۴).

در مرحله‌ی بعد، با استفاده از eRNA و ژن‌های هدف آن‌ها، شبکه‌ی تنظیمی متشکل از ۸۴ گره و ۱۱۹ یال ترسیم شد و تحلیل رگرسیون تک متغیره برای عناصر موجود در شبکه انجام گرفت. همانطور که پیش از این ذکر شد، قبلاً سه مطالعه به‌صورت مجزا، eRNAهای *EMX2OS* و *WAKMAR2*، *CDK6-AS1* را به‌عنوان eRNAهای مرتبط با بقای بیماران مبتلا به سرطان معده معرفی کرده‌اند (۹، ۱۰، ۲۳). اگرچه این مطالعات به بررسی ارتباط eRNAها با بقای بیماران پرداخته‌اند، ولی به‌دلیل اینکه روش انتخاب eRNAهای احتمالی در هر مطالعه متفاوت است و هنوز سیستم ارزیابی و نام‌گذاری مشخصی برای eRNAها وجود ندارد، امکان مقایسه و جمع‌بندی نتایج مطالعات دشوار بود.

پیش‌آگهی و درمان هدفمند بیماران کمک کرد.

نتیجه‌گیری

مطالعه‌ی حاضر با هدف بررسی جامع eRNAهای درگیر در سرطان معده با رویکرد زیست‌شناسی سامانه‌ای، تعیین eRNAها و ژن‌های با بیان متمایز، رسم شبکه و ساخت مدل پیش‌آگهی‌دهنده با استفاده از عناصر موجود در شبکه، انجام گرفت و نشان داده شد، این عناصر با بقای کلی بیماران مبتلا به سرطان معده در ارتباط هستند و می‌توانند به‌عنوان زیست‌نشانه‌های احتمالی کاربرد داشته باشند. اگرچه استفاده از این عناصر به‌عنوان زیست‌نشانه‌گر تشخیصی/پیش‌آگهی‌دهنده نیازمند مطالعات بیشتر و بررسی آن‌ها در دیگر مجموعه داده‌ها است.

تشکر و قدردانی

مطالعه‌ی حاضر برگرفته از پایان‌نامه‌ی مقطع کارشناسی ارشد رشته‌ی ژنتیک انسانی با شماره‌ی ۳۹۸۹۸۱ مصوب دانشگاه علوم پزشکی اصفهان می‌باشد.

در مطالعه‌ی دیگری نشان داده شد که بیان ژن *LINC01605* در سرطان پستان سه‌گانه‌ی منفی نیز افزایش یافته و کاهش دادن بیان آن منجر به مهار تکثیر، مهاجم و مهاجرت در رده‌ی سلولی سرطان پستان سه‌گانه‌ی منفی و مهار تشکیل تومور و مهاجرت در مدل موشی می‌گردد (۲۵).

در خصوص ژن *MIR4435-2HG*، مطالعات قبلی نشان داده‌اند که بیان این ژن در نمونه‌های توموری معده افزایش یافته و بیان آن مرتبط با بقای بیماران است و کاهش دادن بیان آن منجر به کاهش تکثیر سلولی، کاهش مهاجم و مهاجرت و افزایش مرگ سلولی می‌گردد (۲۶، ۲۷).

در مطالعه‌ی دیگر، نقش *MYO16-ASI* در سرطان مثانه‌ی مهاجمی عضلانی بررسی و مشخص شد، بیان آن با بقای بیماران و مهاجم و مهاجرت در ارتباط است (۲۸). نتایج مطالعه‌ی ما نیز هم‌راستا با این مطالعات می‌باشد. به‌طورکلی می‌توان گفت نتایج مطالعه‌ی حاضر در کنار سایر مطالعات می‌تواند به شناخت بهتر اساس مولکولی سرطان معده منجر شود تا با شناخت این عوامل بتوان به تشخیص زودهنگام، تعیین

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; 71(3): 209-49.
- Sexton RE, Al Hallak MN, Diab M, Azmi AS. Gastric cancer: a comprehensive review of current and future treatment strategies. *Cancer Metastasis Rev* 2020; 39(4): 1179-203.
- Ding M, Liu Y, Liao X, Zhan H, Liu Y, Huang W. Enhancer RNAs (eRNAs): New insights into gene transcription and disease treatment. *J Cancer* 2018; 9(13): 2334-40.
- Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol (Pozn)* 2015; 19(1A): A68-77.
- Gu X, Wang L, Boldrup L, Coates PJ, Fahraeus R, Sgaramella N, et al. Ap001056.1, a prognosis-related enhancer rna in squamous cell carcinoma of the head and neck. *Cancers (Basel)* 2019; 11(3): 347.
- Zhang Z, Lee JH, Ruan H, Ye Y, Krakowiak J, Hu Q, et al. Transcriptional landscape and clinical utility of enhancer RNAs for eRNA-targeted therapy in cancer. *Nat Commun* 2019; 10(1): 4562.
- Ang X, Xu Z, Zhou Q, Zhang Z, Ma L, Zhang X, et al. PARGP1, a specific enhancer RNA associated with biochemical recurrence of prostate cancer. *All Life* 2021; 14(1): 774-81.
- Guo XY, Zhong S, Wang ZN, Xie T, Duan H, Zhang JY, et al. Immunogenomic profiling demonstrate AC003092. 1 as an immune-related eRNA in glioblastoma multiforme. *Front Genet* 2021; 12: 633812.
- Zhang Y, Yan Y, Ning N, Shen Z, Ye Y. WAKMAR2, a prognosis-related enhancer RNA in gastric cancer; 2020.
- Yang S, Zou X, Yang H, Li J, Zhang A, Zhang L, et al. Identification of enhancer RNA CDK6-AS1 as a potential novel prognostic biomarker in gastric cancer. *Front Genet* 2022; 13: 854211.
- Khalafiyani A, Emadi-Baygi M, Wolfien M, Salehzadeh-Yazdi A, Nikpour P. Construction of a three-component regulatory network of transcribed ultraconserved regions for the identification of prognostic biomarkers in gastric cancer. *J Cell Biochem* 2023; 124(3): 396-408.
- Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, Čech M, et al. The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 2018; 46(W1): W537-44.
- Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, et al. Ensembl 2020. *Nucleic Acids Res* 2020; 48(D1): D682-8.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014; 15(12): 550.
- Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2014; 30(7): 923-30.
- R Development Core Team. R: A language and environment for statistical computing; 2013.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular

- interaction networks. *Genome Res* 2003; 13(11): 2498-504.
18. Therneau TM, Lumley T. Package 'survival'. *R Top Doc* 2015; 128(10): 28-33.
 19. Heagerty PJ, Saha-Chaudhuri P, Saha-Chaudhuri MP. Package 'survivalROC'. San Fr GitHub; 2013.
 20. Ripley B, Venables B, Bates DM, Hornik K, Gebhardt A, Firth D, et al. Package 'mass'. *Cran r* 2013; 538: 113-20.
 21. Wickham H. *ggplot2*. Wiley Interdiscip Rev Comput Stat 2011; 3(2): 180-5.
 22. Layke JC, Lopez PP. Gastric cancer: diagnosis and treatment options. *Am Fam Physician* 2004; 69(5): 1133-40.
 23. Liu GX, Tan YZ, He GC, Zhang QL, Liu P. EMX2OS plays a prognosis-associated enhancer RNA role in gastric cancer. *Medicine (Baltimore)* 2021; 100(41): e27535.
 24. Ren W, Zhang J, Li W, Li Z, Hu S, Suo J, et al. A tumor-specific prognostic long non-coding RNA signature in gastric cancer. *Med Sci Monit* 2016; 22: 3647.
 25. Wang W, He X, Wang Y, Liu H, Zhang F, Wu Z, et al. LINC01605 promotes aerobic glycolysis through lactate dehydrogenase A in triple-negative breast cancer. *Cancer Sci* 2022; 113(8): 2484-95.
 26. Wang H, Wu M, Lu Y, He K, Cai X, Yu X, et al. LncRNA MIR4435-2HG targets desmoplakin and promotes growth and metastasis of gastric cancer by activating Wnt/ β -catenin signaling. *Aging (Albany NY)* 2019; 11(17): 6657-73.
 27. Gao LF, Li W, Liu YG, Zhang C, Gao WN, Wang L. Inhibition of MIR4435-2HG on invasion, migration, and EMT of gastric carcinoma cells by mediating MiR-138-5p/Sox4 axis. *Front Oncol* 2021; 11: 661288.
 28. Shen D, Zhang Y, Zheng Q, Yu S, Xia L, Cheng S, et al. A competing endogenous RNA network and an 8-lncRNA prognostic signature identify MYO16-AS1 as an oncogenic lncRNA in bladder cancer. *DNA Cell Biol* 2021; 40(1): 26-35.

Comprehensive Analysis of eRNAs and Genes Involved in Gastric Cancer and Constructing a Prognostic Model Predicting the Overall Survival of Patients

Basireh Bahrami¹, Parvaneh Nikpour²

Original Article

Abstract

Background: Gastric cancer (GC) is one of the most prevalent cancers and the first cause of cancer-related deaths in Iran. These patients are diagnosed at advanced stages, which leads to poor prognosis because there is inadequate knowledge of molecular factors to provide diagnostic/prognostic biomarkers. The aim of the current study was the comprehensive analysis of enhancer RNAs (eRNAs) to determine their molecular interactions with other genes and propose a survival-related model for patients.

Methods: RNA-sequencing raw data of GC patients were downloaded from The Cancer Genome Atlas (TCGA), and differentially-expressed eRNAs and genes in tumors compared to non-tumor samples were extracted. The target genes of each eRNA were then identified based on physical distance, correlation, and a regulatory network was constructed with these elements. A prognostic model for predicting the overall survival of patients was eventually established by performing univariate and multivariate Cox regression analyses, and the performance of the model was surveyed.

Findings: By performing differential expression analysis, 69 and 2606 differentially-expressed eRNAs and genes were extracted, respectively, and by identifying the relationships between these elements, a regulatory network consisting of 84 nodes and 119 edges was constructed. A three-components survival-related model subsequently was established which could predict patients' outcomes.

Conclusion: Based on the results, the 3-component constructed model, including an eRNA and two other genes, can be considered as a possible prognostic tool; however, further research is needed to clinically implement it.

Keywords: Enhancer; RNA; Stomach neoplasms; Survival analysis; Systems biology; RNA-Seq

Citation: Bahrami B, Nikpour P. **Comprehensive Analysis of eRNAs and Genes Involved in Gastric Cancer and Constructing a Prognostic Model Predicting the Overall Survival of Patients.** J Isfahan Med Sch 2023; 41(724): 488-95.

1- MSc, Department of Genetics and Molecular Biology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran
2- Associate Professor, Department of Genetics and Molecular Biology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

Corresponding Author: Parvaneh Nikpour, Associate Professor, Department of Genetics and Molecular Biology, School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran; Email: pnikpour@med.mui.ac.ir