مقاله پژوهشی، صفحات ۲۰۱-۱۸۴
DOI: 10.22034/CMR.2024.2214

شناسایی ژن‌های کلیدی دخیل در سرطان پستان سه گانه – منفی به عنوان نشانگر زیستی مستعد با استفاده از آنالیز داده‌های ترانسکریپتومی توالی یابی RNA مبتلایان به سرطان پستان در مقایسه با افراد سالم

پگاه محبوب نیا[۱]، مصطفی رفیعی پور[۱]، رضا مهدیان[۲] و فریناز بهفرجام[۱*]

[۱] **ایران، قزوین، دانشگاه دانش البرز، دانشکده علوم، گروه ژنتیک**

[۲] **ایران، تهران، انستیتو پاستور ایران، گروه پزشکی مولکولی**

تاریخ دریافت: ۱۴۰۱/۰۳/۲۴ تاریخ پذیرش: ۱۴۰۱/۰۹/۱۶

چکیده

سرطان پستان از نوع سه گانه – منفی نوعی سرطان شایع در جهان است که در بافت پستان رشد می‌کند. این بیماری تبدیل به بزرگترین مشکل سلامتی انسان در سراسر جهان شده است. بنابراین شناسایی ژن‌های کلیدی درگیر در این بیماری می‌تواند در تشخیص، پیش‌آگهی و درمان مفید باشد. هدف تحقیق حاضر شناسایی ژن‌های با اختلاف بیان معنادار موثر بر سرطان پستان سه گانه – منفی است که با استفاده از آنالیز داده‌های ترانسکریپتومی توالی یابی RNA مبتلایان به سرطان پستان در مقایسه با افراد سالم انجام گرفت. داده‌های ترانسکریپتوم تعیین توالی شده با دستگاه Illumina مربوط به سه فرد مبتلا به سرطان پستان از نوع سه گانه– منفی و چهار فرد سالم از پایگاه داده NCBI و SRA جمع‌آوری شده و پس از کنترل کیفیت خوانش‌ها با استفاده از نرم‌افزار FastQC، همترازی توالی‌ها با ژنوم مرجع با نرم‌افزار STAR انجام شد. در نهایت، جهت بررسی بیان افتراقی ژن‌ها، از نرم‌افزارهای featureCounts و DESeq2 استفاده شد و ژن‌های با اختلاف بیان معنادار به کمک نمودار Volcano نمایه گردید. در نهایت برای بررسی نقش ژن‌ها در مسیرهای مولکولی، تجزیه و تحلیل هستی شناسی ژن انجام گرفت. در این تحقیق، ۳۱ ژن دارای بیان افتراقی با سطح معنی داری $p<0.05$ شناسایی شد. در بین این ۳۱ ژن، ۱۱ ژن کاهش بیان و ۲۰ ژن افزایش بیان نشان دادند. از جمله ژن‌های شاخص این مجموعه ژنی، می‌توان به ژن‌های *RNF223*، *ETV3L*، *CACNA1S*، *UPS34*، *NEURL3*، *CRYBG3* و *FABP4* اشاره نمود. بررسی و مطالعات انجام شده در مورد این ژن‌ها حاکی از نقش مهم آن‌ها در مسیرها و فرآیندهای مولکولی درگیر در بیماریزایی افراد مبتلا به سرطان پستان از نوع سه گانه– منفی بود. بنابراین معرفی این نشانگرهای زیستی مهم در مکانیسم‌های مولکولی حیاتی سلول‌ها می‌تواند برای اهداف تشخیصی، درمانی و پیشگیری بسیار با اهمیت و منحصر بفرد باشد.

واژه های کلیدی: سرطان پستان، بیان افتراقی ژن، ترانسکریپتوم و پایگاه داده NCBI

*نویسنده مسئول، پست الکترونیکی: Behfarjam.farinaz@gmail.com

مقدمه

سرطان پستان یکی از هتروژن‌ترین انواع سرطان‌ها و علت ۲۲٪/۹ از سرطان‌ها در زنان و دومین سرطان شایع در کشورهای درحال توسعه است [۲۲]. برطبق جدیدترین آماری که ارائه شده است، سالانه ۳۳/۲۱ زن از هر ۱۰۰۰۰۰ زن در ایران با میانگین سنی ۴۹/۸۴ سال به سرطان پستان مبتلا می‌شوند که در این میان ۱۴/۲ نفر از هر ۱۰۰۰۰۰ نفر فوت می‌کنند [۲۲]. تخمین زده می‌شود در سال ۲۰۳۰ در مجموع ۲۶/۴ میلیون مورد ابتلا به

مقاله پژوهشی، صفحات ۲۰۱-۱۸۴
DOI: 10.22034/CMR.2024.2214

سرطان پستان در جهان وجود داشته باشد که ۱۷ میلیون آن منجر به مرگ شود. در بین انواع مختلف سرطان پستان، نوع سه گانه – منفی (Triple Negative) دارای شیوع بالایی می‌باشد که این نوع، به گونه‌ای از بیماری سرطان پستان گفته می‌شود که فاقد گیرنده استروژن(ER)، گیرنده پروژسترون (PR) و HER2/neu باشد [۱۵].

از آنجایی که بافت پستان اندامی حساس به استروژن است و در ایران مصرف داروهای ضدبارداری و سایر موارد دارویی شامل استروژن و یا پیش ساز آن در کنار عواملی همچون رژیم پرچرب و کم فیبر شانس ابتلا به سرطان پستان را افزایش می‌دهد، تشخیص و درمان کارآمد دو مرحله‌ی حیاتی بشمار می‌آیند [۱۱].

نشانگرهای بافتی از اولین نشانگرهای زیستی مرتبط با سرطان پستان می‌باشند، گیرنده‌های هورمونی مانند گیرنده استروژن و گیرنده پروژسترون از جمله اولین نشانگرهایی هستند که در مراحل ابتدایی انواع سرطان پستان جهت پیش‌آگهی و انتخاب روش درمانی مناسب بکار می‌روند [۲۱]. گیرنده فاکتور رشد اپیدرمال انسانی نوع ۲ (HER2) نیز دیگر نشانگر پیش‌آگهی سرطان پستان می‌باشد که پاسخ آن تنها در صورتی قابل اتکا است که گره‌های لنفاوی نیز درگیر سرطان شده باشند. نشانگر زیستی یوروکیناز پلاسمینوژن و کاتپسین دی نیز در جهت پیش‌آگهی سرطان‌های پستان با عدم درگیری گره‌های لنفاوی کاربرد دارد [۱۹] . در نهایت، می‌توان از افزایش بیان *P53* نیز به عنوان یک نشانگر زیستی برای پیش‌آگهی سرطان پستان نام برد اما به دلیل عدم اختصاصیت این نشانگر و مشاهده-ی افزایش بیان *P53* در بسیاری از انواع سرطان‌ها به نظر می‌رسد نمی‌تواند نشانگر زیستی مفیدی واقع شود [۱۴].

تاکنون حداقل هشت مجموعه‌ی ژنی به عنوان نشانگرهای پیش آگهی پاتوژنز و درمان بیماری شناسایی شده است. در این میان، تغییرات بیانی ۵۱۲ ژن دخیل در پاسخ به آسیب سلولی، تغییرات بیانی ۹۷ ژن پیش‌بینی کننده درجه بیماری و تغییرات بیانی ۳۲ ژن بیانگر درگیری پاسخ به آسیب بواسطه *P53* و تغییرات بیانی ۱۸۶ ژن، پیش‌بینی کننده تهاجمی/غیرتهاجمی بودن سرطان پستان می‌باشند [۶]. بر اساس مواردی که ذکر شد و روش‌های مختلفی که در راستای دستیابی به راهکارهای سریع‌تر در تشخیص، درمان و پیش آگهی بیماری‌های مختلف علی الخصوص سرطان پستان نوع سه گانه – منفی وجود دارد، امروزه تمرکز بسیاری از محققان بر روی روش‌های نوین توالی یابی نسل بعدی است که در این بین روش توالی‌یابی RNA (RNA sequencing) دارای اهمیت شایان ذکری می‌باشد.

روش توالی‌یابی RNA ، با استفاده از فناوری توالی‌یابی نسل جدید، کمیت محتوای RNA یک نمونه زیستی را در یک بازه زمانی مشخص بررسی می‌کند. بنابراین به نظر می‌رسد با استفاده از این روش بتوان نشانگرهای دقیق‌تری را جهت پیش‌آگهی بیماری‌ها عرضه کرد. پیشرفت‌های اخیر در توالی‌یابی نسل بعد باعث شده که بازه پوشش توالی‌یابی توکلئوتیدهای DNA و همچنین مقدار نمونه ورودی برای توالی‌یابی افزایش پیدا کند [۱۱]. این پیشرفت‌ها توالی‌یابی رونوشت‌های RNA سلول را تسهیل می‌کند و این امکان را فراهم می‌کند که رونوشت‌های برش خورده ژن، تغییرات پس از رونویسی، همجوشی ژن‌ها، جهش‌ها و پلی‌مورفیسم تک نوکلئوتیدی و در نهایت تغییرات در بیان ژن را بررسی نماید [۵]. توالی‌یابی RNA فن آوری است که با بهره‌گیری از توانایی‌های «توالی‌یابی نسل بعدی» برای بدست آوردن تصویری کلی از حضور و مقدار RNA از ژنوم در یک بازه‌ی زمانی خاص استفاده می‌کند [۵]. همچنین این روش دارای کاربردهایی همانند شناسایی تفاوت‌های ژنتیکی، مسیرها و شبکه‌های زیست مولکولی موثر در بیماریزایی افراد است [۱۰]. از اینرو، با توجه به عدم وجود راهکارهایی جهت تشخیص زود هنگام، اقدامات پیشگیرانه و روش‌های درمانی مناسب بیماری سرطان پستان از نوع سه– گانه منفی و از طرفی به دلیل پیشرفت‌های حاصله در زمینه روش‌های نوین نسل

DOI: 10.22034/CMR.2024.2214

بعدی توالی یابی، در این مطالعه داده‌های مربوط به مطالعات ترانسکریپتوم بیماران مبتلا به سرطان پستان از نوع سه گانه منفی در مقایسه با افراد سالم با هدف شناسایی نشانگرهای زیستی احتمالی مورد مطالعه قرار گرفتند.

مواد و روشها

دریافت داده‌ها از پایگاه داده NCBI : از آنجایی که پایگاه داده‌ی NCBI یک سیستم یکپارچه بازیابی اطلاعات بوده و قادر به جستجوی همزمان در بانک‌های مختلفی نظیر GeneBank، PubMed و سایر پایگاه‌های داده‌ی مشابه می-باشد، داده‌های مورد نیاز جهت انجام این تحقیق نیز از پایگاه داده‌ای NCBI و SRA بدست آمد [۵]. در تحقیق حاضر از داده‌های ترانسکریپتوم تعیین توالی شده با دستگاه Illumina مربوط به سه فرد مبتلا به سرطان پستان از نوع سه گانه- منفی و چهار فرد سالم از پایگاه داده NCBI و SRA استفاده شد (لینک دسترسی به داده‌ها: https://trace.ncbi.nlm.nih.gov/Traces/study/?acc=SRP2356 48&o=acc_s%3Aa). جدول ۱ به اطلاعات دموگرافیک این داده‌ها اشاره می‌کند.

جدول ۱- اطلاعات دموگرافیک داده‌های ترانسکریپتومی بدست آمده از پایگاه داده NCBI و SRA

نام نمونه	بیمار/سالم	جنسیت	سن	منبع داده	Layout*
shC-1	سالم	زن	۵۱	ترانسکریپتومی	Paired
shC-2	سالم	زن	۵۱	ترانسکریپتومی	Paired
shC-3	سالم	زن	۵۱	ترانسکریپتومی	Paired
shC-4	سالم	زن	۵۱	ترانسکریپتومی	Paired
shMETTL18-1	بیمار	زن	۵۱	ترانسکریپتومی	Paired
shMETTL18-2	بیمار	زن	۵۱	ترانسکریپتومی	Paired
shMETTL18-3	بیمار	زن	۵۱	ترانسکریپتومی	Paired

* خوانش‌های دو طرفه

کنترل کیفیت خوانش‌ها: پس از گرفتن داده‌ها از پایگاه داده NCBI و SRA، کنترل کیفیت آن‌ها، به کمک نرم‌افزار Fast QC در محیط لینوکس انجام شد (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). نرم‌افزار Fast QC علاوه بر فهرست تعداد خوانش‌ها و رمزگذاری کیفیت آن‌ها، اطلاعات مربوط به کیفیت و محتوای بازها، طول خوانش و محتوای K-مر و همچنین حضور بازهای مبهم و مضاعف شدگی‌ها را گزارش داده و مصورسازی می‌کند.

همردیفی خوانش‌ها با ژنوم مرجع: در مرحله‌ی همردیف کردن خوانش‌ها با ژنوم مرجع، داده‌های تعیین کیفیت شده با ژنوم مرجع (GRCh38.p13 (GCA_000001405.28)) http://ftp.ensembl.org/pub/release-106/fasta/homo_sapiens//dna) همردیف شدند. در تحقیق حاضر جهت همردیفی توالی‌ها از نرم‌افزار STAR استفاده شد (https://github.com/alexdobin/STAR/archive) STAR یک برنامه‌ی همردیف‌ساز حساس و سریع است که برای همردیفی خوانش‌های توالی‌یابی NGS در ژنوم‌های انسانی مورد استفاده قرار می‌گیرد.

تبدیل فایل‌های SAM به BAM : فایل‌های SAM که معمولاً توسط همردیف‌سازها ایجاد می‌شوند نیازمند پردازش‌هایی نظیر تبدیل SAM/BAM، مرتب‌سازی، نمایه‌سازی یا ادغام هستند. در تحقیق حاضر از نرم‌افزار SAMtools و Picard استفاده شد. (https://sourceforge.net/)

DOI: 10.22034/CMR.2024.2214 مقاله پژوهشی، صفحات ۲۰۱-۱۸۴

/projects/samtools) ابتدا فایل‌های SAM به فایل‌های BAM تبدیل شد. این کار با استفاده از مجموعه‌ی Samtools از ابزارهای commandline صورت گرفت. از گزینه‌ی n– مرتب‌سازی Samtools استفاده شد تا از این طریق اطمینان حاصل شود که فایل SAM بر اساس مشخصه‌ی خوانش مرتب گردد.

مرتب‌سازی و نمایه‌سازی کردن فایل‌های BAM : جهت انجام این دو مرحله از نرم افزار Samtools در محیط لینوکس استفاده شد. هدف از مرتب‌سازی دسترسی کارآمدتر به داده‌ها در این فایل‌ها می‌باشد. فایل BAM مرتب شده دارای داده‌هایی است که بر اساس کروموزوم-ها/ کانتیگ‌ها و داربست‌ها در ژنوم مرجع مرتب شده است.

بررسی و شمارش خوانش‌های همردیف با ژن‌ها: جهت بررسی و شمارش خوانش‌های همردیف شده به ژن‌ها از نرم‌افزار FeatureCounts استفاده شد. (https://sourceforge.net/projects/subread/files/subread-2.0.3/) داده ورودی به نرم‌افزار FeatureCounts شامل یک یا چند فایل از خوانش‌های همردیف شده با فرمت SAM یا BAM است. با استفاده از این نرم افزار تعداد خوانش-هایی که با هر ژن همردیف شده‌اند، مشخص می‌شود و از این‌رو ویژگی خوانش‌ها تعیین خواهد شد.

بررسی بیان افتراقی ژن‌ها: آنالیز بیان افتراقی شامل شناسایی ژن‌ها یا سایر انواع ترکیبات ژنومی مانند رونوشت‌ها یا اگزون‌ها است. در تحقیق حاضر جهت بررسی بیان افتراقی، از نرم‌افزار DESeq2 در محیط R استفاده شد.

(https://bioconductor.org/packages/release/bioc/html/DESeq2.html) که فایل ورودی این نرم‌افزار در واقع همان فایل خروجی نرم‌افزار FeatureCounts است که به صورت یک ماتریس می‌باشد. این نرم‌افزار خوانش‌هایی که با هر بخش از یک ویژگی همپوشانی داشتند شمارش کرده و هر خوانشی را که با چندین ویژگی همپوشانی داشت، حذف می‌نماید. در نتایج حاصل از این نرم افزار، ژن‌های دارای بیان افتراقی در افراد مبتلا به سرطان پستان از نوع سه گانه منفی در مقایسه با افراد سالم بدست می‌آیند. پس از یافتن Ensembl Gene ID برای ژن‌های دارای بیان افتراقی، در پایگاه داده BioDBnet (https://biodbnet-abcc.ncifcrf.gov/db/db2db.php) ، Gene Symbol ها به دست آمده، در نهایت با اعمال فیلتر *p-value* <0.05 و Log2 <= -0.05 تا >= +0.05 در محیط پایتون ژن‌های دارای بیان افتراقی معنادار مشخص شدند.

هستی شناسی ژن و و غنی سازی مسیر ژن‌های با بیان متفاوت: برای انعکاس عملکردهای ژن، آنالیزهای هستی شناسی ژن انجام گرفت. ToppGene (ToppFun) (https://toppgene.cchmc.org/enrichment.jsp) یک پایگاه داده آنلاین است که مجموعه‌ای جامع از منابع را برای حاشیه نویسی عملکردی ارائه می‌دهد تا اهمیت بیولوژیکی فهرست گسترده‌ای از ژن‌ها را تشخیص دهد [۴]. آنالیزهای غنی‌سازی عملکردی ژن‌های با بیان افتراقی، از جمله آنالیزهای هستی شناسی ژن و تحلیل غنی‌سازی مسیرهای درگیر آن‌ها، در تحقیق حاضر، با استفاده از ToppGene با معیار p<0.05 cut-off و تعداد غنی‌سازی ژن بیشتر از ۲ انجام شد.

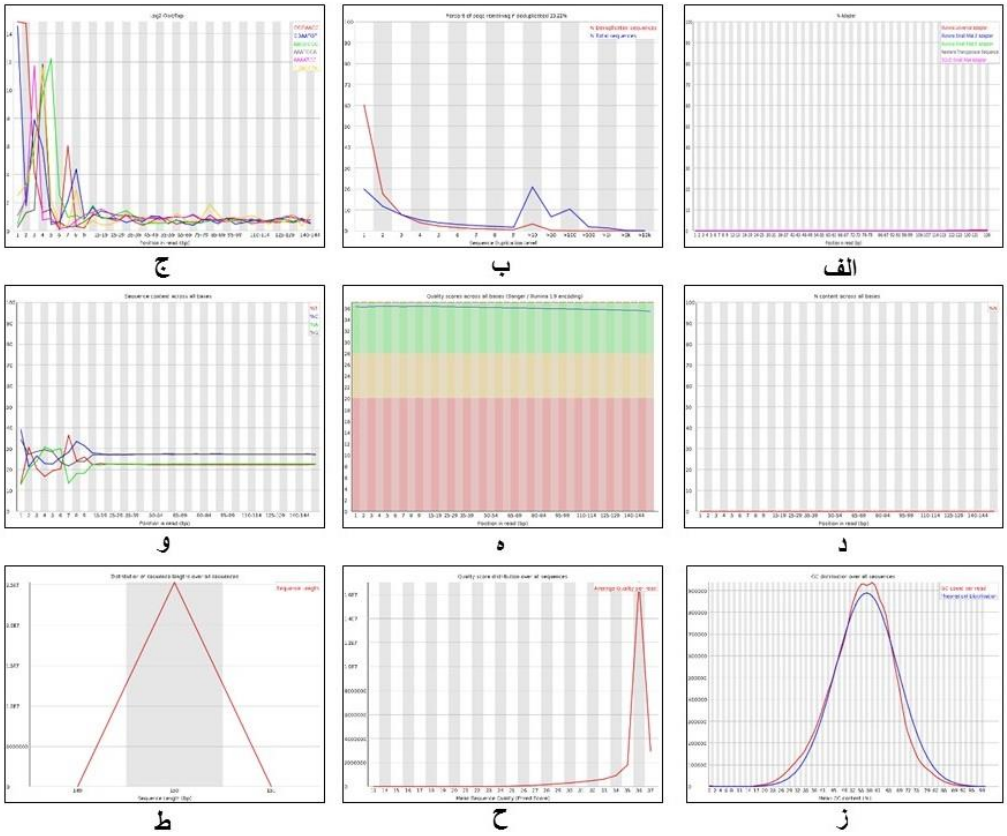
نتایج

نتایج کنترل کیفیت داده‌ها: در تحقیق حاضر داده‌های ترانسکریپتومی ۷ نمونه انسانی شامل ۳ نمونه سرطانی و ۴ نمونه سالم مورد بررسی قرار گرفت که نتایج حاصل از کنترل کیفیت داده‌های مربوطه، در شکل ۱ (فایل پیوست شکل های ۱–۶) نشان داده شده است.

نتایج مربوط به همردیفی خوانش‌ها با ژنوم مرجع: نتایج مربوط به همردیفی خوانش‌ها با ژنوم مرجع که با استفاده از نرم‌افزار STAR انجام گردید در جدول ۲ قابل مشاهده

مقاله پژوهشی، صفحات ۲۰۱-۱۸۴
DOI: 10.22034/CMR.2024.2214

است، درصد Mapping ، عمق خوانش و طول خوانش- های حاصله برای هر کدام از داده‌ها حاکی از همردیفی مناسب آن‌ها می‌باشد.



شکل ۱- نمودارهای مربوط به کنترل کیفیت خوانش‌ها در نمونه shC-1. الف) وجود و یا عدم وجود آداپتور؛ ب) مضاعف شدگی؛ ج) توالی‌های با پیچیدگی پایین؛ د) بازهای مبهم؛ ه) کیفیت خوانش؛ و) درصد بازهای فراخوانده شده؛ ز) تعداد خوانش‌ها در مقابل درصد GC؛ ح) تعداد کل خوانش‌ها در مقابل متوسط نمره کیفیت و ط) توزیع طول خوانش‌ها.

جدول ۲- نتایج Mapping با استفاده از نرم‌افزار STAR

نام نمونه	بیمار/سالم	درصد Mapping*	طول خوانش	عمق خوانش**
shC-1	سالم	92.26%	59664908	7.80
shC-2	سالم	91.94%	55775450	7.33
shC-3	سالم	92.04%	44287158	6.17
shC-4	سالم	92.64%	59096936	7.75
shMETTL18-1	بیمار	91.25%	61992298	7.59
shMETTL18-2	بیمار	92.68%	53128374	7.36
shMETTL18-3	بیمار	93.54%	61429524	7.76

* درصد خوانش های هم ردیف شده با ژنوم مرجع

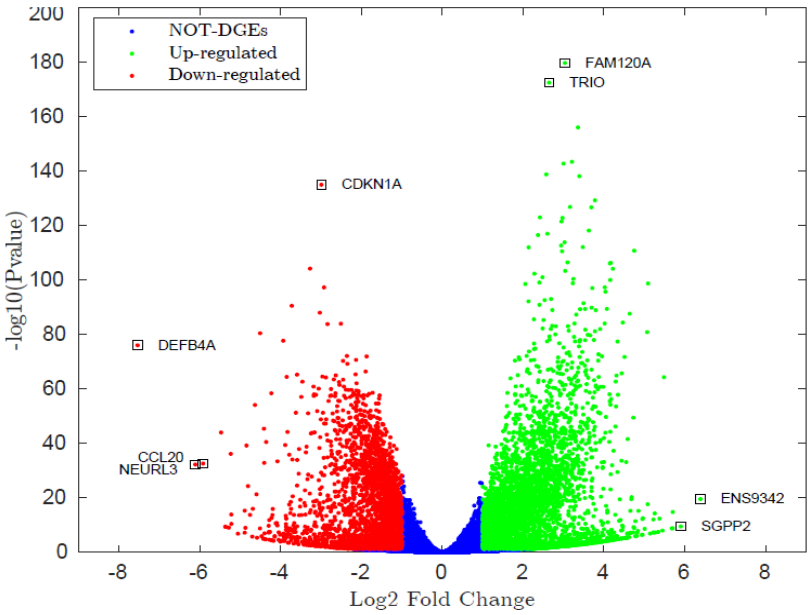
** تعداد خوانش های کوتاهی که در یک منطقه ژنومی خاص روی یکدیگر همپوشانی دارند.

مقاله پژوهشی، صفحات ۲۰۱-۱۸۴
DOI: 10.22034/CMR.2024.2214

نتایج حاصل از آنالیز بیان افتراقی: فایل‌های BAM، پس از مرتب‌سازی، با استفاده از نرم‌افزار FeatureCount ارزیابی شدند. بعد از مرتب‌سازی نتایج اولین مرحله از تجزیه و تحلیل داده‌ها در نرم افزار FeatureCount، این نتایج با استفاده از نرم افزار DESeq2 جهت شناسایی ژن‌های دارای بیان افتراقی، آنالیز شدند.

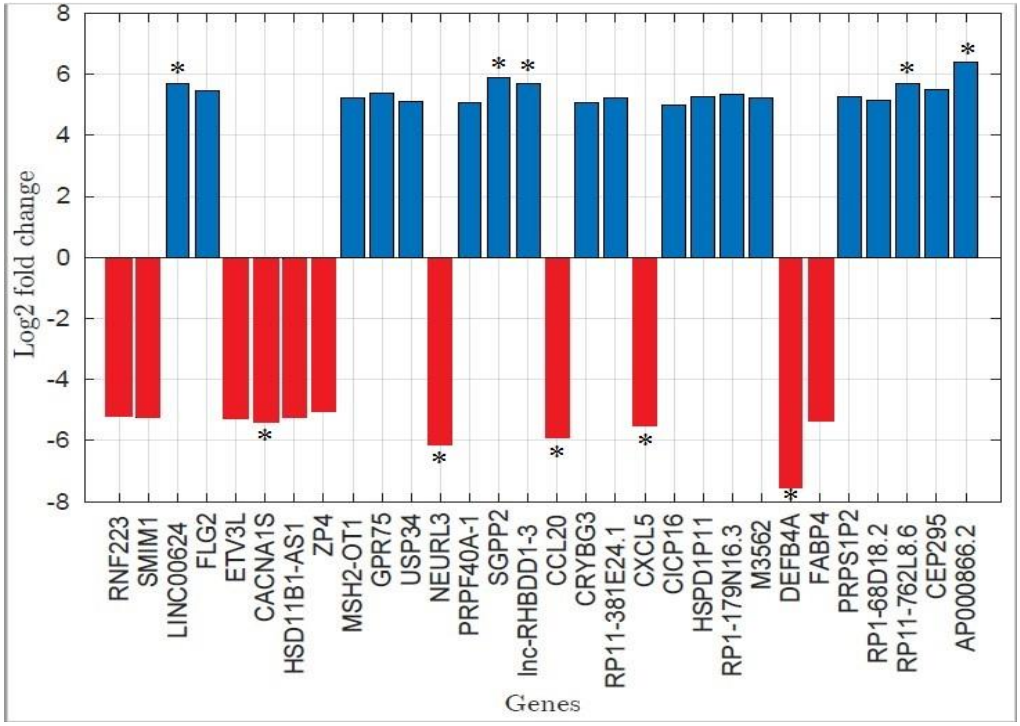
در مجموع نتایج حاصل از آنالیز بیان افتراقی ژن، ۹۲۰۱ ژن با بیان متفاوت را نشان داد که به کمک پایگاه داده BioDBnet ، Gene Symbol آنها یافت شد. برای نمایش بیان افتراقی ژن‌ها از نمودار Volcano استفاده شد (شکل ۲). در این نمودار برای نمایش ژن‌های با بیان متفاوت از سه طیف رنگی استفاده شد. نقاط سبز رنگ نمایانگر ژن‌های با افزایش بیان معنادار با اعمال فیلتر *p-value* <0.05 و Log2 > 0.01 ، نقاط قرمز رنگ، ژن‌های با کاهش بیان معنادار با اعمال فیلتر *p-value* <0.05 و Log2 > -0.01 است. در نهایت، نقاط آبی رنگ مربوط به ژن‌هایی با *p-value* نامناسب می‌باشند (*p-value* >0.05). به عبارتی دیگر، ژن-هایی که در این طیف رنگی قرار گرفته‌اند (رنگ آبی) به طور معنادار دچار افزایش یا کاهش بیان نشده‌اند.

بررسی میزان log2 fold Change و *p-value* (*p-value* <0.05 و Log2 <-0.05 تا Log2 >+0.05) به عنوان دو معیار اصلی در تفکیک ژن‌های با بیان متفاوت نشان داد که از میان ۹۲۰۱ ژن، ۳۱ ژن سطح بیان متفاوت‌تری را بر اساس دامنه‌ای که تعیین شد، داشته‌اند. قابل ذکر است که با تغییر این دامنه که در برخی مقالات هم اشاره شده است، می-توان ژن‌های با تفاوت بیان بیشتری را شناسایی کرد. در این تحقیق، از دامنه‌ی سختگیرانه‌ای برای این آنالیز استفاده شد (Log2 <-0.05 تا Log2 >+0.05) و نتیجه حاصل به صورت نمودار ستونی نمایش داده شد (شکل ۳). از این ۳۱ ژن، ۱۱ ژن کاهش بیان معنادار نشان دادند که با رنگ قرمز و ۲۰ ژن هم افزایش بیان معنادار نشان دادند که با رنگ آبی در نمودار نمایش داده شدند.



شکل ۲- نمایش بیان افتراقی ژن‌ها با استفاده از نمودار Volcano. در این نمودار بیان افتراقی ژن‌ها با استفاده از دو پارامتر *p-value* و Log2 نمایش داده شد.

مقاله پژوهشی، صفحات ۲۰۱-۱۸۴
DOI: 10.22034/CMR.2024.2214



شکل ۳- شناسایی ۳۱ ژن با بیان متفاوت با اعمال دو فیلتر log2 fold Change و *p*-value . ۱۱ ژن به رنگ قرمز دچار کاهش بیان معنادار و ۲۰ ژن به رنگ آبی دچار افزایش بیان معنادار . * نمایانگر ۱۰ ژن با تفاوت بیان متفاوت‌تری نسبت به سایر ژن‌ها

در بین این ۳۱ ژن، با اعمال فیلتر سختگیرانه‌ای، ۱۰ ژن تفاوت بیان متفاوت‌تری نسبت به سایر ژن‌ها داشتند. به طوری که پنج ژن RP11-762L8.6، lnc-RHBDD1-3، LINC00624، SGPP2 و AP000866.2 بیشترین میزان بیان یا به عبارتی بالاترین سطح بیان معناداری و پنج ژن CACNA1S، CXCL5، CCL20، NEURL3 و DEFB4A کمترین میزان بیان یا به عبارتی پایین‌ترین سطح بیان معناداری را نشان دادند (شکل ۳).

نتایج مربوط به تجزیه و تحلیل های هستی شناسی ژن:

تجزیه و تحلیل هستی شناسی برای ۳۱ ژن تمایز یافته برتر مشخص کرد که ژن های تمایز یافته به طور قابل توجهی در فرآیندهای بیولوژیکی و مکانیسم‌های مولکولی موثر در سرطان پستان از جمله فرآیندهای مرتبط با پاسخ سلولی، رشد و تکثیر سلولی، آپوپتوز، فرآیندهای متابولیسمی، فرآیندهای مرتبط با چرخه سلولی، تولیدمثل و پاسخ‌های ایمنی درگیر هستند (جدول ۳).

بحث و نتیجه‌گیری

سرطان پستان از نوع منفی – سه گانه نوع پیشرونده‌ای از این بیماری است که امروزه فاقد نشانگرهای زیستی موثر و اهداف درمانی مورد نیاز برای مطالعات تشخیصی و درمانی می‌باشد. سرطان پستان از نوع منفی سه گانه، مسئول تقریبا ۱۵ درصد تمامی سرطان‌های پستان تهاجمی است. امروزه به دلیل عدم شناخت کافی در مورد مکانیسم‌های مولکولی و عملکردی و اهداف درمانی، شیمی درمانی عمده‌ترین روش برای درمان این نوع از سرطان است. بنابراین، نیاز جدی به شناسایی نشانگرهای درمانی و تشخیصی می‌باشد [۱،۲،۱۳]. روش‌های نوینی همانند ریزآرایه و تکنیک‌های توالی یابی عمیق، تغییرات مولکولی گسترده‌ای را در این بیماری ایجاد نموده و منجر به پیشرفت قابل توجهی در کارهای بالینی شده است. در سال‌های اخیر، توالی یابی RNA به عنوان یک فن آوری جدید برای تجزیه و تحلیل ترانسکریپتوم با توان بالا مطرح شده است. بسیاری از

مقاله پژوهشی، صفحات ۲۰۱-۱۸۴
DOI: 10.22034/CMR.2024.2214

مطالعات، ترانسکریپت‌های دارای بیان افتراقی را شناسایی نموده و به عنوان نشانگرهای زیستی مستعد در طبقه بندی زیرگروه‌های سرطان معرفی نموده‌اند که می‌تواند برای شناسایی و درمان این بیماری تهاجمی و پیشرونده بسیار مفید و ارزشمند باشد [۲۶]. در این مطالعه، یک تجزیه و تحلیل افتراقی جامعی بر روی داده‌های ترانسکریپتوم افراد مبتلا به سرطان پستان از نوع سه گانه – منفی در مقایسه با افراد سالم انجام گرفت. نتایج حاصل از مطالعه ۷ داده شامل ۳ فرد مبتلا به سرطان پستان و ۴ فرد سالم، در نهایت منجر به شناسایی ۳۱ ژن دارای بیان افتراقی با سطح معنی داری $p<0.05$ شد. در بین این ۳۱ ژن، ۱۱ ژن کاهش بیان و ۲۰ ژن افزایش بیان نشان دادند. ارزیابی‌های بیشتر این ژن‌ها حاکی از نقش مهم و کلیدی آن‌ها در مکانیسم-های مولکولی درگیر در بیماریزایی سرطان پستان بود.

جدول ۳- گزیده‌ای از تجزیه و تحلیل هستی شناسی برای ۳۱ ژن تمایز یافته برتر حاصل از آنالیز بیان افتراقی

ژن‌های منتخب	مسیر مولکولی	تعداد ژن‌های مسیر مولکولی	تعداد ژن درگیر	شماره هستی شناسی ژن
CACNA1S	فعالیت کانال کلسیم با ولتاژ بالا	11	1	**0.051499296**
FABP4	اتصال اسیدهای چرب با زنجیره بلند	14	1	**0.056749282**
FABP4	فعالیت ناقل اسیدهای چرب با زنجیره بلند	18	1	**0.064075348**
CRYBG3 ZP4 FLG2	فعالیت مولکول ساختاری	784	3	**0.065610489**
CRYBG3	اجزای ساختاری عدسی چشم	26	1	**0.065610489**
FABP4	اتصال به گیرنده های هورمونی	35	1	**0.084538781**
FABP4	اتصال اسید چرب	39	1	**0.090458458**
CCL20 CXCL5 FABP4 DEFB4A	اتصال گیرنده سیگنال	1722	4	**0.091297792**
CACNA1S	فعالیت کانال کلسیمی وابسته به ولتاژ	48	1	**0.103076404**
CRYBG3	اتصال پروتئین کیناز A	53	1	**0.106052014**
FABP4	اتصال اسید مونو کربوکسیلیک	74	1	**0.142314177**

در مطالعه اخیر، ژن *USP34* در افراد مبتلا به سرطان پستان از نوع سه گانه – منفی در مقایسه با افراد سالم افزایش ۵ برابری نشان داد. بررسی و مطالعات انجام شده حاکی از این است که این ژن پپتیداز اختصاصی یوبی کوئیتین ۳۴ است که بر روی کروموزوم ۲ قرار گرفته است. یوبی کوئیتین یک پروتئین ۷۶ اسیدآمینه‌ای است که با اتصال به یکسری از پروتئین‌های هدف و تجزیه آن‌ها نقش مهمی در تنظیم فرآیندهای سلولی دارد. حال یوبی کوئیتین پروتئازها می‌توانند در حذف این عامل از پروتئین‌های هدف نقش داشته باشند. خانواده ژنی پروتئازهای اختصاصی یوبی کوئیتین، یکی از کلاس‌های مهمی از پروتئازها در انسان هستند. پروتئاز اختصاصی یوبی کوئیتین UPS34 می‌تواند نوع خاصی از آنزیم جدا کننده یوبی کوئیتین را رمزگذاری کند [۱۲]. مطالعات زیادی به نقش حیاتی پپتیدازهای اختصاصی یوبی کوئیتین در پیشرفت سرطان اشاره نموده‌اند [۲۴]. همچنین، بر اساس مطالعات قبلی چندین مسیر مهم مرتبط با سرطان توسط اعضای مختلف UPS تنظیم می‌شوند که از جمله آن‌ها می‌توان به مسیرهای JAKs-STATs ، مسیر NF-kB ، مسیر TGFB و مسیر Wnt اشاره نمود. مطالعات حاکی از نقش مهم UPS34 در مسیر پیام رسانی Wnt است [۲۴]. از این رو می توان ژن *UPS34* را به عنوان نشانگر زیستی احتمالی مهم در سرطان پستان

مقاله پژوهشی، صفحات ۲۰۱-۱۸۴
DOI: 10.22034/CMR.2024.2214

معرفی نمود و با طراحی مهارکننده‌های موثر علیه آن از پیشرفت سرطان جلوگیری نمود. ژن دیگری که در مطالعه کنونی به عنوان یک نشانگر زیستی احتمالی شناسایی گردید، ژن *CRYBG3* بود. این ژن به صورت معنی داری در افراد مبتلا به سرطان پستان در مقایسه با افراد کنترل افزایش معناداری نشان داد. تجزیه و تحلیل‌های هستی شناسی ژن و بررسی منابع حاکی از نقش مهم و تعیین کننده این ژن در مکانیسم‌های مولکولی درگیر در بیماریزایی افراد مبتلا به سرطان پستان بود که تاییدی بر یافته‌های این مطالعه بود. نقش مهم ژن *CRYBG3* به این صورت است که این ژن می‌تواند با یک آنزیم مهم فرآیند گلیکولیز به نام لاکتات دهیدروژناز برهمکنش داشته باشد. مطالعات انجام شده توسط چن و همکاران در سال ۲۰۱۸، افزایش بیان هماهنگ این دو ژن را در افراد مبتلا به سرطان ریه گزارش نمودند [۳]. ژن *CRYBG3* به عنوان تنظیم کننده گلیکولیز است و افزایش بیان آن باعث جذب گلوکز و تولید لاکتات می‌شود و این در حالی است که ناک داوون آن نتایج معکوسی را نشان داده و باعث مهار تکثیر سلولی می‌گردد. بنابراین می‌توان گفت که ژن *CRYBG3* می‌تواند به عنوان یک نشانگر زیستی احتمالی مهم در تشخیص و پیش آگهی افراد مبتلا به سرطان بوده و هدف گیری آن می‌تواند گامی در جهت درمان این بیماری پیچیده باشد [۲۳].

یکی از ژن‌هایی که در این مطالعه دچار کاهش بیان شد، ژن *FABP4* بود. کاهش ۵ برابری این ژن با سطح معنی داری $p<0.05$ در افراد مبتلا به سرطان پستان در مقایسه با افراد سالم مشاهده شد. پروتئین متصل شونده به اسید چرب آدیپوسیت (FABP4) به وفور در ماکروفاژها و سلول‌های آدیپوسیت وجود دارد. با توجه به فراوانی این پروتئین در این دسته از سلول‌ها، نظر بر این است که FABP4 در متابولیسم لیپید نقش دارد. طبق مطالعات انجام شده توسط ژانک و همکاران (۲۰۱۸)، نقش پروتئین FABP4 در بسیاری از سرطان‌ها از جمله سرطان ریه، سرطان پستان، سرطان تخمدان و سرطان پروستات گزارش شده است. این پروتئین می‌تواند به لیگاندهای لیپیدی هیدروفوب متصل گردد و در تنظیم متابولیسم گلوکز و لیپید نقش داشته باشد. همچنین نقش مهمی در پیام رسانی درون سلولی ایفا می‌کند [۲۵]. نتایج حاصل از تجزیه و تحلیل‌های هستی شناسی ژن و یکسری از مطالعات حاکی از تاثیر تعیین کننده این پروتئین بر روی فرآیند تکثیر سلولی و آپوپتوز می‌باشد. علاوه براین، گزارش شده که هر گونه کاهش بیان در این پروتئین کاملا مرتبط با سایز توموری هست که مورد مطالعه قرار گرفته است و افزایش بیان آن با القای ژن خارجی منجر به مهار تکثیر و مهاجرت سلول‌های توموری شده است [۲۵]. با توجه به اینکه تشخیص و شناسایی بیماری سرطان یکی از معظلات امروزه پزشکان و متخصصان می‌باشد، شناسایی ژن مهم و کلیدی *FABP4* می‌تواند بسیار ارزشمند باشد. یکی دیگر از نشانگرهای کلیدی و مهم گزارش شده در این مطالعه، ژن *RNF223* بود. نتیجه حاصله برای این ژن کاهش بیان معنی دار آن در افراد مبتلا به سرطان پستان در مقایسه با افراد سالم بود. ژن *RNF223* بر روی کروموزوم ۱ واقع شده است و پروتئین حاصل از آن با داشتن یک دومین انگشت حلقه نقشی کلیدی در مسیر یوبی کویتینه شدن دارد. بسیاری از پروتئین‌های دارای این دومین، می‌توانند به اهداف مولکولی خود متصل گردند و عملکرد آن‌ها را تنظیم کنند. از جمله اهداف این پروتئین‌ها می‌توان به مولکول‌های DNA، RNA، پروتئین و یا ساختارهای لیپیدی اشاره نمود. زوآ و همکاران (۲۰۱۹) در مطالعه‌ای که با هدف بهبود تشخیص در بیماری سارکوما انجام دادند، به نقش مهم و تعیین کننده ژن‌های *FGF23*، *TLX2*، *TIFAB*، *RNF223*، *HIST1H3A* و *AADACL4* پی بردند. آنالیزهای عملکردی با کیفیت بالایی که این محققین انجام دادند، در نهایت منجر به معرفی این ۶ ژن شد که دارای نقش حیاتی در فرآیندهای بیولوژیکی مرتبط با سرطان بوده و در بهبود تشخیص این بیماری پیچیده می‌توانند بسیار کمک کننده باشند [۲۱]. با توجه به نقش مهم این پروتئین در فرآیند

مقاله پژوهشی، صفحات ۲۰۱-۱۸۴
DOI: 10.22034/CMR.2024.2214

مهم یوبی کوتئینه کردن پروتئین‌های هدف، کاهش بیان آن می‌تواند منجر به بهم ریختن تنظیم مسیرهای مولکولی مختلف درگیر در داخل بدن شود و اثرات جبران ناپذیری داشته باشد. از این رو شناسایی چگونگی عملکرد این پروتئین و تشخیص زودهنگام آن می‌تواند در امر تشخیص و پیش آگهی مفید باشد.

ژن دیگری که در نتایج حاصل از این مطالعه به عنوان نشانگر زیستی احتمالی معرفی شد، ژن *ETV3L* بود. این ژن در تجزیه و تحلیل‌های انجام شده کاهش بیان معنی داری (۵ برابری) نشان داد. بررسی‌های انجام شده در بانک‌های اطلاعاتی، حاکی از وجود این ژن در موقعیت 1q23.1 با ۲۰۷۹۰ جفت باز بود که از نظر عملکردی نقش مهمی در بسیاری از فرآیندهای درون سلولی داشت. فاکتورهای نسخه برداری ETS ، بیان ژن‌های درگیر در تکوین طبیعی سلولی ، تکثیر، تمایز، رگ زایی و آپوپتوز را تنظیم می‌کنند. این دسته از فاکتورها دارای ۲۸ عضو هستند که ژن *ETV3L* به عنوان یکی از اعضای مهم این ابرخانواده ژنی می‌باشد. عدم تنظیم این فاکتورهای نسخه برداری، تکثیر سلولی در انواع مختلف سرطان‌ها را تسهیل می‌کند و چندین عضو آن‌ها با فعال کردن نسخه برداری ژن‌های خاصی در تهاجم و متاستاز نقش دارند. ۲۸ عضو شناسایی شده از این خانواده ژنی تا به امروز دارای یک توالی اسیدآمینه‌ای حفاظت شده به نام دومین ETS هستند که تقریبا دارای ۸۵ اسید آمینه بوده و یک دومین متصل شونده به DNA در انتهای کربوکسیلی ایجاد می‌کنند که برای شناسایی توالی هدف GGAA/T (جایگاه متصل شونده به ETS) ضروری می‌باشد. فاکتورهای نسخه برداری ETS از جمله ETV3L با اتصال به نواحی متصل شونده به DNA، می‌توانند در فرآیندهای دیگری مثل تکوین سلول‌های بنیادی، تکثیر سلولی، تمایز، رشد سلولی، ترانسفورماسیون، رگ زایی و آپوپتوز نقش داشته باشند. این فاکتورها پروتئین‌های افکتوری پایین دستی مسیرهای مولکولی مهم درون سلولی از جمله مسیرهای RAS/ RAF/ ERK هستند که مسیرهایی مهم در فرآیندهای فوق می‌باشند. به گونه‌ای که ارتباط سطح بیان ژن‌های ETS با پیشرفت تومور در بسیاری از نئوپلازیایی انسانی شامل لوکمیا، تیروئید، پانکراس، کبد، پروستات، کلون، ریه و سرطان پستان مشاهده شده است که هر گونه تغییر بیان در آن‌ها می‌تواند باعث تغییر در ژن‌های هدف پاسخ دهنده شود. [۷]. *NEURL3*، از دیگر ژن‌های شاخص معرفی شده در این مطالعه است. این ژن کاهش بیان ۶ برابری در افراد مبتلا به سرطان در مقایسه با افراد کنترل نشان داد. مطالعات حاکی از نقش مهم پروتئین حاصل از این ژن در تکوین غدد پستانی و به تبع سرطان پستان بود. خانواده ژنی نوروگلین، لیگاندهای شامل EGF را رمزگردانی می‌کنند که قادرند با اتصال به پذیرنده‌های تیروزین کینازی در شبکه‌های پیام رسانی مهم در مسیر تکوین غدد پستانی و سرطان پستان بسیار مهم و تاثیرگذار باشند [۱۶]. نتایج حاصل از تجزیه و تحلیل‌های هستی شناسی ژن و مطالعات مختلف حاکی از نقش مهم پروتئین NEURL3 در مراحل مختلف ریخت زایی و اختصاصی شدن سلول‌های پستانی در مراحل جنینی هست. مطالعات اخیر نشان داده‌اند که هر گونه تغییر در میزان تولید این پروتئین مهم می‌تواند نقش بسیار تعیین کننده‌ای در شکل گیری این سلول‌های مهم در بافت پستانی داشته باشد. مطالعات لی و همکاران (۲۰۱۷) و گاتناتی و همکاران (۲۰۲۰) موید این نتایج و اطلاعات است [۱۶،۹]. نشانگر زیستی احتمالی مهم دیگری که در مطالعه کنونی گزارش شد، ژن *CACNA1S* بود [۲۰]. نتایج حاصله کاهش معنی دار این ژن در افراد مبتلا به سرطان پستان در مقایسه با افراد سالم را نشان داد. مطالعه صورت گرفته توسط فان و همکاران در سال ۲۰۱۷ به اهمیت این ژن اشاره نموده و آن را به عنوان یکی از اهداف جدیدی برای درمان سرطان معرفی کردند [۱۷]. ژن *CACNA1S* یکی از اعضای مجموعه ژنی کانال‌های کلسیمی وابسته به ولتاژ می‌باشد که این کانال‌ها دارای نقش‌های تایید شده‌ای در فرآیندهای سلولی مثل میتوژنز، تکثیر سلولی، تمایز، آپوپتوز و متاستاز

DOI: 10.22034/CMR.2024.2214 مقاله پژوهشی، صفحات ۲۰۱-۱۸۴

هستند [۱۸]. در گزارشی ارتباط بین این دسته ژنی و کاهش در تکثیر و افزایش آپوپتوز در سرطان پروستات مطرح شده است که به دلیل کاهش بیان این ژن‌ها و هدف گیری آن‌ها گامی مهم در جهت درمان بیماری سرطان برداشته شده است. بیان ترانسکریپت‌های کانال‌های یونی به عنوان نشانگرهای زیستی مستعدی در انواع مختلف سرطان مثل سرطان پروستات و سرطان پستان مطرح شده است. بنابراین مطالعه نقش‌های عملکردی این کانال‌های یونی در مراحل تکوین سرطان ممکن است اهداف ارزشمندی برای تشخیص تومور مشخص نماید که نتایج حاصل از تجزیه و تحلیل‌های هستی شناسی ژن نیز به این امر اشاره داشت. به دلیل نقش مهم کلسیم و کانال‌های یونی وابسته به این یون در فرآیندهای درون سلولی و مکانیسم‌های مولکولی تعیین کننده سلولی، معرفی این ژن مهم می‌تواند ارزش بالینی خیلی بالایی داشته باشد [۱۸].

در این مطالعه، تجزیه و تحلیل‌های انجام شده بر روی داده‌های ترانسکریپتوم افراد مبتلا به سرطان پستان از نوع سه گانه – منفی در مقایسه با افراد سالم نتایج ارزشمندی را نشان داد. نتیجه این مطالعات، شناسایی ۳۱ ژن دارای بیان افتراقی با سطح معنی داری $p<0.05$ بود. در بین این ۳۱ ژن، ۱۱ ژن کاهش بیان و ۲۰ ژن افزایش بیان نشان دادند. از جمله ژن‌های شاخص این مجموعه ژنی یافت شده می‌توان به ژن‌های *RNF223*، *ETV3L*، *CACNA1S*، *UPS34*، *NEURL3*، *CRYBG3* و *FABP4* و سایر ژن‌های مهمی مثل *CCL20*، *CXCL5*، *CEP295* و *FLG2* اشاره نمود. بررسی و مطالعات انجام شده در مورد این ژن‌ها حاکی از نقش مهم آن‌ها در مسیرها و فرآیندهای مولکولی درگیر در بیماریزایی افراد مبتلا به سرطان پستان بود. از جمله فرآیندهایی که این ژن‌ها دخیل هستند، می‌توان فرآیندهای تکثیر سلولی، تنظیم متابولیسم گلوکز و لیپید، آپوپتوز، پیام رسانی درون سلولی، تمایز، تکوین، تنظیم کانال‌های وابسته به یون‌ها، ترانسفورماسیون، رگ زایی و اتصالات سلولی را نام برد. بنابراین معرفی این نشانگرهای زیستی مهم در مکانیسم‌های مولکولی حیاتی سلول‌ها می‌تواند برای اهداف تشخیصی، درمانی و پیشگیری بسیار با اهمیت و منحصر بفرد باشد که البته نیازمند مطالعات گسترده‌تر برای تایید نتایج کنونی است.

تضاد منافع

نویسندگان اعلام می‌کنند که هیچ تضاد منافعی ندارند.

منابع

[1] دهقانی, س., سوده, سلیمی, و نیکونهادلطف آبادی. (۲۰۱۹). بررسی اثرضد توموری ترکیب شیمیاییN ((پنج-نیتروتیوفن-۲-YL)متیلن)-۲ (فنیل تیو) بنزوهیدرازید درمدل موشی سرطان پستان. پژوهش‌های سلولی و مولکولی (مجله زیست شناسی ایران), ۳۲(۲), ۲۰۵-۲۱۵.

[2] متولی باشی مجید, کوه کن فاطمه, و حجتی زهره. نقش سلولهای بنیادی سرطانی در افزایش ریسک ابتلاء به متاستاز و کاهش میزان بقای افراد مبتلا به سرطان پستان. پژوهش‌های سلولی و مولکولی (مجله زیست شناسی ایران), ۲۶(۳), ۳۶۵-۳۷۷.

[3] Chen, H., Pei, H., Hu, W., Ma, J., Zhang, J., Mao, W., & Zhou, G. (2018). Long non-coding RNA CRYBG3 regulates glycolysis of lung cancer cells by interacting with lactate dehydrogenase A. Journal of Cancer, 9(14), 2580.

[4] Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. Nucleic acids research. 2009 Jul 1;37(suppl_2):W305-11.

[5] Chu, Y., & Corey, D. R. (2012). RNA sequencing: platform selection, experimental design, and data interpretation. Nucleic acid therapeutics, 22(4), 271-274.

[6] Desmedt, C., Piette, F., Loi, S., Wang, Y., Lallemand, F., Haibe-Kains, B., & Sotiriou, C. (2007). Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clinical cancer research, 13(11), 3207-3214.

مقاله پژوهشی، صفحات ۲۰۱-۱۸۴ DOI: 10.22034/CMR.2024.2214

[7] Fry, E. A., Mallakin, A., & Inoue, K. (2018). Translocations involving ETS family proteins in human cancer. Integrative cancer science and therapeutics, 5(4).

[8] Gao, T., Han, Y., Yu, L., Ao, S., Li, Z., & Ji, J. (2014). CCNA2 is a prognostic biomarker for ER+ breast cancer and tamoxifen resistance. PloS one, 9(3), e91771.

[9] Ghatnatti, V., Vastrad, B., Patil, S., Vastrad, C., & Kottureshetti, I. (2021). Identification of potential and novel target genes in pituitary prolactinoma by bioinformatics analysis. AIMS neuroscience, 8(2), 254.

[10] Graf, A., Krebs, S., Heininen-Brown, M., Zakhartchenko, V., Blum, H., & Wolf, E. (2014). Genome activation in bovine embryos: review of the literature and new insights from RNA sequencing experiments. Animal reproduction science, 149(1-2), 46-58.

[11] Harirchi, I., Karbakhsh, M., Kashefi, A., & Momtahen, A. J. (2004). Breast cancer in Iran: results of a multi-center study. Asian pacific journal of cancer prevention, 5(1), 24-27.

[12] Lin, C., Xia, J., Gu, Z., Meng, Y., Gao, D., & Wei, S. (2020). Downregulation of USP34 inhibits the growth and migration of pancreatic cancer cells via inhibiting the PRR11. OncoTargets and therapy, 13, 1471.

[13] Modi, S., Saura, C., Yamashita, T., Park, Y. H., Kim, S. B., Tamura, K., & Krop, I. (2020). Trastuzumab deruxtecan in previously treated HER2-positive breast cancer. New England Journal of Medicine, 382(7), 610-621.

[14] Moghaddam, F. D., Mortazavi, P., Hamedi, S., Nabiuni, M., & Roodbari, N. H. (2020). Apoptotic effects of melittin on 4T1 breast cancer cell line is associated with up regulation of Mfn1 and Drp1 mRNA expression. Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Cancer Agents), 20(7), 790-799.

[15] Nafissi, N., Khayamzadeh, M., Zeinali, Z., Pazooki, D., Hosseini, M., & Akbari, M. E. (2018). Epidemiology and histopathology of breast cancer in Iran versus other Middle Eastern countries. Middle East Journal of Cancer, 9(3), 243-251.

[16] Nam, K., Lee, K. W., Chung, O., Yim, H. S., Cha, S. S., Lee, S. W., & Jeong, J. Y. (2017). Analysis of the FGF gene family provides insights into aquatic adaptation in cetaceans. Scientific reports, 7(1), 1-13.

[17] Ou, H. D., Phan, S., Deerinck, T. J., Thor, A., Ellisman, M. H., & O'shea, C. C. (2017). ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. Science, 357(6349), eaag0025.

[18] Phan, N. N., Wang, C. Y., Chen, C. F., Sun, Z., Lai, M. D., & Lin, Y. C. (2017). Voltage-gated calcium channels: Novel targets for cancer therapy. Oncology letters, 14(2), 2059-2074.

[19] Pharoah, P. D. P., Day, N. E., & Caldas, C. (1999). Somatic mutations in the p53 gene and prognosis in breast cancer: a meta-analysis. British journal of cancer, 80(12), 1968-1973.

[20] Schartner, V., Romero, N. B., Donkervoort, S., Treves, S., Munot, P., Pierson, T. M., & Laporte, J. (2017). Dihydropyridine receptor (DHPR, CACNA1S) congenital myopathy. Acta neuropathologica, 133(4), 517-533.

[21] Sun, P., Shen, Y., Wang, T., He, Y., Zhang, Y., Tian, W., ... & Hu, Y. (2021). Distinct clinical and genetic mutation characteristics in sporadic and Lynch syndrome-associated endometrial cancer in a Chinese population. Cancer Epidemiology, 73, 101934.

[22] Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., & Jemal, A. (2015). Global cancer statistics, 2012. CA: a cancer journal for clinicians, 65(2), 87-108.

[23] Wang, K. J., Zha, X., Chen, D. D., & Zhu, S. Q. (2018). Mutation analysis of families with autosomal dominant congenital cataract: a recurrent mutation in the CRYBA1/A3 gene causing congenital nuclear cataract. Current Eye Research, 43(3), 304-307.

[24] Young, M. J., Hsu, K. C., Lin, T. E., Chang, W. C., & Hung, J. J. (2019). The role of ubiquitin-specific peptidases in cancer progression. Journal of biomedical science, 26(1), 1-14.

[25] Zhong, C. Q., Zhang, X. P., Ma, N., Zhang, E. B., Li, J. J., Jiang, Y. B., & Cheng, S. Q. (2018). FABP4 suppresses proliferation and invasion of hepatocellular carcinoma cells and predicts a poor prognosis for hepatocellular carcinoma. Cancer medicine, 7(6), 2629-2640.

[26] Zou, X., Chen, K., Zou, J., Han, P., Hao, J., & Han, Z. (2020). Single-cell RNA-seq data analysis on the receptor ACE2 expression reveals the potential risk of different human organs vulnerable to 2019-nCoV infection. Frontiers of medicine, 14(2), 185-192.

مقاله پژوهشی، صفحات ۲۰۱-۱۸۴ DOI: 10.22034/CMR.2024.2214

Identification of key genes as potential biomarkers involved in triple-negative breast cancer using RNA sequencing transcriptomic data analysis of breast cancer patients compared to healthy individuals

Mahboobnia P.[1], Rafiepour M.[1], Mahdian R.[2] and BehFarjam F.[1]

[1] Dept. of Genetics, Faculty of Science, Danesh Alborz University, Qazvin, I.R. of Iran

[2] Dept. of Molecular Medicine, Pasteur Institute of Iran, Tehran, I.R. of Iran

Abstract

Triple Negative breast cancer is a common cancer in the world that grows in breast tissue. The disease has became the biggest human health problem worldwide. Therefore, identifying the key genes involved in this disease can be useful in the diagnosis, prognosis and treatment. The aim of this study was to identify genes with significant expression differences affecting triple negative breast cancer using RNA sequencing transcriptomic data analysis of breast cancer patients compared to healthy individuals. Illumina-sequenced transcriptomic data of three breast cancer patients and four healthy individuals were collected from the NCBI and SRA databases, and after controlling the quality of the readings using FastQC software, the sequences were aligned with Reference genome was performed with STAR software. Finally, featureCounts and DESeq2 softwares were used to investigate the differential expression of genes. Volcano graph was used to show genes with significant expression differences. In this study, 31 genes with differential expression were identified with a significance level of $p < 0.05$. Among these 31 genes, 11 genes showed reduced expression and 20 genes showed increased expression. Among of these gene set, we can mention to *NEURL3*, *UPS34*, *CACNA1S*, *ETV3L*, *RNF223*, *FABP4*, *CRYBG3* genes. Studies on these genes indicate their important role in the pathways and molecular processes involved in the pathogenesis of triple negative breast cancer. Therefore, the introduction of these important biomarkers in the vital molecular mechanisms of cells can be very important and unique for diagnostic, therapeutic and preventive purposes.

Keywords: Breast cancer; Differential gene expression; Transcriptome; NCBI