Observational Study

Medicine®

OPEN

In silico screening using bulk and single-cell RNA-seq data identifies *RIMS2* as a prognostic marker in basal-like breast cancer

A retrospective study

Lingyun Zhang, MD[a,b], Zheng Liu, BS[c], Jingqiang Zhu, MD[a,b,*]

Abstract

Single-cell RNA-seq has become a powerful tool to understand tumor cell heterogenicity. This study tried to screen prognosis-related genes in basal-like breast tumors and evaluate their correlations with cellular states at the single-cell level.

Bulk RNA-seq data of basal-like tumor cases from The Cancer Genome Atlas-Breast Cancer (TCGA-BRCA) and single-cell RNA-seq from GSE75688 were retrospectively reviewed. Kaplan–Meier survival curves, univariate and multivariate analysis based on Cox regression model were conducted for survival analysis. Gene set enrichment analysis (GSEA) and single-cell cellular functional state analysis were performed.

Twenty thousand five hundred thirty genes with bulk RNA-seq data in TCGA were subjected to screening. Preliminary screening identified 10 candidate progression-related genes, including *CDH19, AQP5, SDR16C5, NCAN, TTYH1, XAGE2, RIMS2, GZMB, LY6D,* and *FAM3B*. By checking their profiles using single-cell RNA-seq data, only *CDH19, SDR16C5, TTYH1,* and *RIMS2* had expression in primary triple-negative breast cancer (TNBC) cells. Prognostic analysis only confirmed that *RIMS2* expression was an independent prognostic indicator of favorable progression free survival (PFS) (HR: 0.78, 95%: 0.64–0.95, $P = .015$). GSEA analysis showed that low *RIMS2* group expression had genes significantly enriched in DNA Repair, and MYC Targets V2. Among the 89 basal-like cells, *RIMS2* expression was negatively correlated with DNA repair and epithelial-to-mesenchymal transition (EMT).

RIMS2 expression was negatively associated with DNA repair capability of basal-like breast tumor cells and might serve as an independent indicator of favorable PFS.

Abbreviations: BRCA = breast cancer, DSS = disease specific survival, EMT = epithelial-to-mesenchymal, ER = estrogen receptor, GSEA = gene set enrichment analysis, HER2 = human epidermal growth factor receptor 2, ICGC = International Cancer Genome Consortium, OS = overall survival, PFS = progression-free survival, PR = progesterone receptor, TCGA = The Cancer Genome Atlas, TNBC = triple-negative breast cancer.

Keywords: basal-like breast cancer, DNA repair, prognosis, *RIMS2*, single-cell sequencing

1. Introduction

The heterogeneous nature of breast cancer has been characterized during the past decades. Based on the molecular profiles by a 50-gene qPCR assay (PAM50), breast cancers include five major subtypes, including luminal A, luminal B, human epidermal growth factor receptor 2-enriched (HER2+), basal-like, and normal-like.[1] Over 70% of basal-like cases are triple-negative breast cancer (TNBC), which means the lack expression of estrogen receptor (ER) and progesterone receptor (PR), in combination with an absence of human epidermal growth factor receptor 2 (HER2) gene overexpression or amplification.[2] Therefore, most of the basal-like cases are unresponsive to targeted hormonal therapies and other targeting therapies.[3] Current therapeutic strategies limit to physical surgery, nonselective chemotherapeutic agents and radiotherapy.[4,5] Compared to non-basal-like cases, basal-like cases have a higher level of genetic heterogenicity,[6–8] and significantly shorter 3-year progression-free survival (PFS) and overall survival (OS).[5] Therefore, it is necessary to explore reliable biomarkers that help identify patients with a high risk of disease progression.

In the past decades, multiple large projects, such as the Cancer Genome Atlas (TCGA) and International Cancer Genome Consortium (ICGC) have made great efforts to explore the critical cancer-causing genomic alterations and to generate a comprehen-

Editor: Geraldine Vidhya Raja.

LZ and ZL contributed equally to this work.

No funding was received for this study.

The authors have no conflict of interest to disclose.

Supplemental digital content is available for this article.

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

[a] Department of Thyroid Surgery, [b] Laboratory of Thyroid and Parathyroid Disease, Frontiers Science Center for Disease-Related Molecular Network, [c] Nursing Department, West China School of Nursing, West China Hospital, Sichuan University, Chengdu, Sichuan, China.

[*] Correspondence: Jingqiang Zhu, Department of Thyroid Surgery, West China Hospital, Sichuan University, Chengdu, 610041, Sichuan, China (e-mail: zjq-wkys@163.com).

How to cite this article: Zhang L, Liu Z, Zhu J. In silico screening using bulk and single-cell RNA-seq data identifies RIMS2 as a prognostic marker in basal-like breast cancer: A retrospective study. Medicine 2021;100:16(e25414).

Received: 11 July 2020 / Received in final form: 18 February 2021 / Accepted: 10 March 2021

http://dx.doi.org/10.1097/MD.0000000000025414

sive landscape of cancer genomic profiles.[9,10] However, cancer cells are embedded in the tumor microenvironment (TME), which is constituted by complex cellular and non-cellular components. RNA-seq data of these projects is based on the average expression signals of bulk tumor cells, which comprise at least tumor cells, stromal cells and tumor-infiltrating immune cells.[11] Therefore, it is impossible to analyze intratumor heterogeneity and identify tumor cell-associated genes.[12] In the past years, single-cell sequencing has become a powerful new tool to solve these issues.[4,12] Via preparing single-cell samples, capturing gene transcripts and generating individual cell sequencing libraries, single-cell RNA-seq makes it possible to analyze the transcriptome of a single cell.[13] Thus, this technique enables us to explore the tumor cell-specific prognostic marker genes, which have extensive-expression across different basal-like tumor samples.

This study aimed to screen tumor cell-associated prognostic genes in basal-like breast tumors via combining bulk and single-cell RNA-seq data from large online databases. Then, their correlations with tumor cell states were assessed at the single-cell level.

2. Patients and methods

2.1. Ethics declarations

Ethical approval was not required since this retrospective study was based on online open databases.

2.2. Bioinformatic data mining in the Cancer Genome Atlas-Breast Cancer (TCGA-BRCA)

The normalized level-3 data of the breast cancer cohort in TCGA database was downloaded using the UCSC Xena (http://xena.ucsc.edu/).[14] PAM50 subtypes were determined by TCGA Analysis-working groups according to RNA-seq data. RNA-seq data were presented as log2(norm_count+1), in which norm_count refers to RNA-Seq expression estimation by Expectation-Maximization (RSEM) normalized count. The following criteria were applied for screening samples included for analysis:

1. primary basal-like breast tumor;
2. with PFS data; and
3. in the censored group, follow-up time was longer than 365 days.

The data of basal-like cases were extracted for further analysis, including clinicopathological parameters, genomic and survival data. Briefly, the sample type, age at initial pathologic diagnosis, pathological stages, nodal status, margin status, radiation therapy, targeted molecular therapy, PFS and disease-specific survival (DSS) were collected. In this dataset, PFS refers to the period from the date of diagnosis until the date of the first occurrence of a new tumor event (NTE), which includes a progression of the disease, locoregional recurrence, distant metastasis, new primary tumor, or death with the tumor. Patients who were alive without these event types, or died without tumor were censored. DSS event is defined as death from the disease.[15]

2.3. Single-cell transcriptional data in basal-like breast tumor cells

To investigate the expression profile of progression-related genes in basal-like breast tumor cells and their correlation with cellular states at the single-cell level, single-cell RNA-sequencing data were retrieved from CancerSEA, which is an online platform for analyzing available RNA-seq datasets in GEO dataset.[16] Fourteen functional state models were constructed via estimating the expression of state signature genes from databases and published literature.[16]

The basal-like (TNBC) data from a previous single-cell RNA-seq dataset (GSE75688) were extracted for reanalysis. In this dataset, 317 tumor cells were prepared from 11 breast tumor samples covering four subtypes of breast cancer.[17] Six out of the 11 samples were basal-like (TNBC) from 5 patients (BC07, BC07LN, BC08, BC09, BC10, BC11). The current study typically focused on primary tumor cases. BC07LN is a metastatic lymph node tissue and thus was excluded. In terms of BC09, two runs of single-cell RNA sequencing were performed. This case was also excluded, to ensure the consistency of data preparation and analysis. The t-Distributed Stochastic Neighbor Embedding (t-SNE) based on protein-coding gene (PCG) expression was used to generate a two-dimensional map for showing the cells' distances and cluster identities.

2.4. Meta-analysis-based prognostic analysis using the Kaplan–Meier plotter

The prognostic significance of the candidate progression-related genes in patients with basal-like tumors was examined by performing Kaplan–Meier survival analysis using the Kaplan–Meier plotter (http://kmplot.com/analysis/), which is an online tool for survival analysis based on over 1800 patients collected from different GEO datasets.[18] Both PFS and OS were analyzed by using the autoselect best cutoff of gene expression.

2.5. Gene set enrichment analysis

Gene set enrichment analysis (GSEA)[19] was conducted to examine the difference within the Hallmark gene sets (H) between the basal-like cases with low and high *RIMS2* expression. GSEA software v4.0.1 on a JAVA 8.0 platform was used. The number of permutations was set to 1000. The gene set with NOM P-value $< .05$ and adjusted q-values (FDR) $< .25$ were considered significant.

2.6. Statistical analysis

Data were reported as the mean $\pm$ standard deviation (SD). Welch's t-test (unequal variance t-test) was applied to check the statistical difference. PFS, OS, and DSS curves were generated by the Kaplan–Meier method, with the Log-rank test to check the statistical difference. Patients were separated by the best cutoff of gene expression. The independent prognostic significance of *RIMS2* expression was calculated by the univariate and multivariate Cox regression models. $P < .05$ was considered statistically significant.

3. Results

3.1. Systematic screening of progression-related PCGs in basal-like breast cancer

To screen progression-related PCGs in basal-like breast cancer, PFS and RNA-seq data were extracted from TCGA-BRCA. Eight hundred forty-two primary tumor cases had PFS data, including

420 luminal A, 192 luminal B, 140 basal-like, 67 HER2+, and 23 normal-like cases (Fig. 1A). Among the 140 basal-like cases, there were 23 progression and 117 censored cases. Twenty out of the 117 censored cases had follow-up time shorter than 365 days and thus were excluded. Therefore, a total of 120 cases were subjected to the comparison of gene expression (20,530 genes with RNA-seq data) between the groups with or without progression after primary therapy (Fig. 1B).

The PFS related genes were identified by using the following two criteria:

1. |log2 fold change (FC)|≥1.5;
2. Welch's *P* value < .05.

After the screening process, 10 candidate genes were identified, including *CDH19*, *AQP5*, *SDR16C5*, *NCAN*, *TTYH1*, *XAGE2*, *RIMS2*, *GZMB*, *LY6D*, and *FAM3B* (Fig. 1C, Supplementary Table 1, http://links.lww.com/MD2/A55). Among these candidates, the first 8 were downregulated, while the last 2 were upregulated in the progression group compared to the non-progression group (Fig. 1D). Then, the expression of the 10 progression-related genes in basal-like tumor cells was checked at the single-level level (Fig. 1E and F). In GSE75688, 317 tumor cells from 11 breast cancer patients, including four basal-like cases (BC07, BC08, BC10, and BC11) were subjected to single-cell RNA seq (Fig. 1E and F). Only the primary tumor cells were included in the study, while lymph node cases, tumor-associated immune cells, and non-carcinoma stromal cells were excluded. Therefore, a total of 89 cells were included for analysis. The cells' distance in the reduced 2D space was visualized by the t-SNE plots of PCGs (Fig. 1E). By setting TPM value >1 in at least 10 cells as the cutoff, 6 genes (*AQP5, NCAN, XAGE2*, *GZMB*, *LY6D*, and *FAM3B*) had limited expression (Fig. 1F). Therefore, these genes were excluded for analysis.

3.2. Meta-analysis-based exploration of potential prognostic genes

To explore the prognostic significance of the candidate genes (*CDH19*, *SDR16C5*, *TTYH1*, and *RIMS2*), K-M survival analysis was conducted using the Kaplan–Meier plotter. Patients with basal-like tumors were separated into two groups according to the best cutoff of gene expression. Log-rank test showed that only *CDH19* and *RIMS2* had potential prognostic value. High *CDH19* expression was associated with favorable PFS (HR: 0.68, 95%: 0.49–0.94, *P* = .02) (Fig. 2A), while high *RIMS2* expression was associated with both favorable PFS (HR: 0.70, 95%: 0.49–

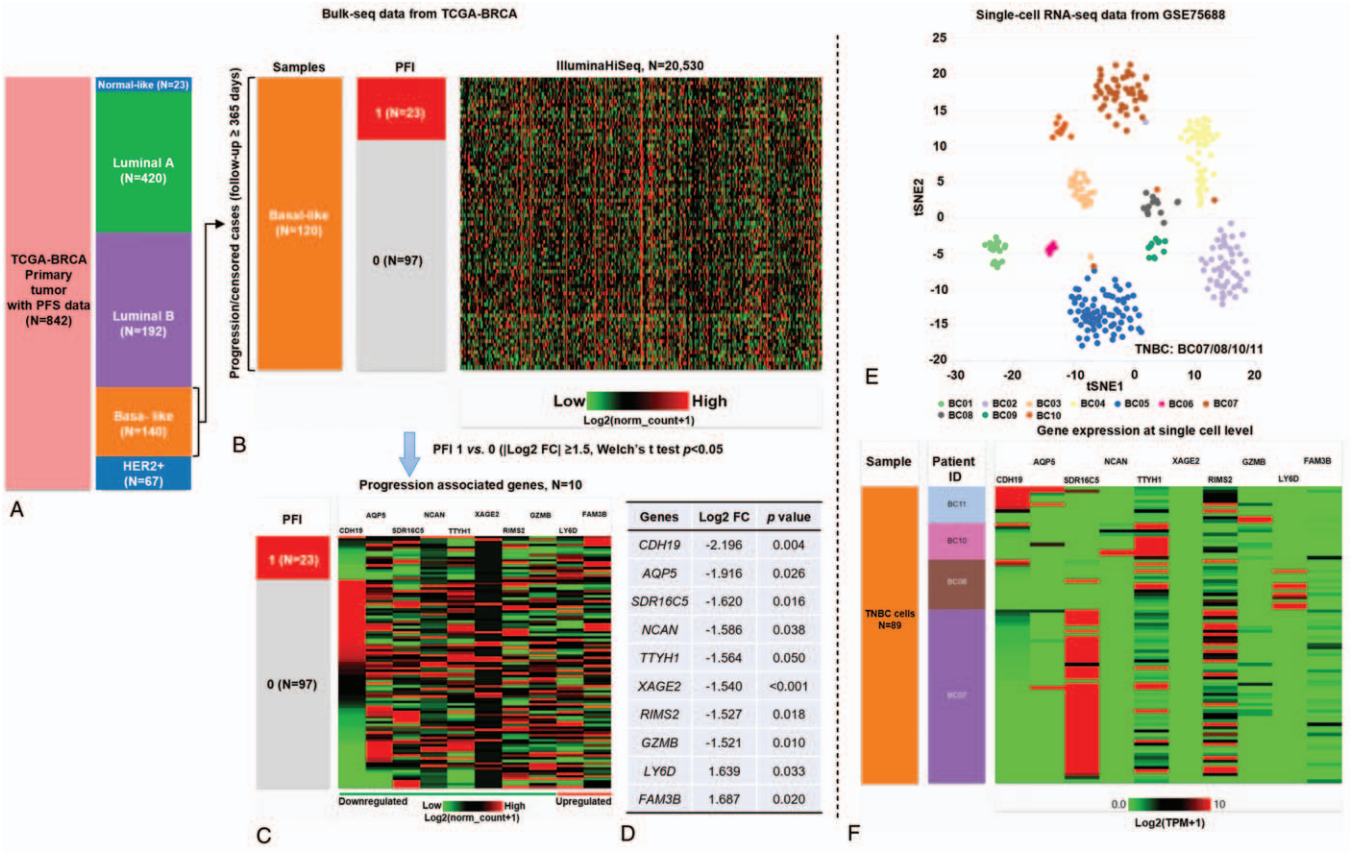

Figure 1. Systematic screening of progression related PCGs in basal-like cases. (A) A bar chart showing the composition of PAM50 cases in TCGA-BRCA. (B) A heatmap showing the expression profile of 20,530 genes in 140 basal-like cases in TCGA. The criteria applied for screening were briefly introduced at the bottom of the heatmap. (C) A heatmap showing the expression profile of 10 candidate genes between the cases with or without progression, including *CDH19, AQP5, SDR16C5, NCAN, TTYH1, XAGE2, RIMS2, GZMB, LY6D,* and *FAM3B*. (D) A table showing the log2 fold change of the candidate genes in progression group compared to the non-progression group. (E) Visualization of cells in GSE75688. Cell distances were showed in the reduced 2D space by the t-SNE method, based on the expression of PCGs. Clusters are colored by the patient of the cell source. (F) A heatmap showing the expression profile of 10 candidate genes in 89 TNBC cells from 4 patients.

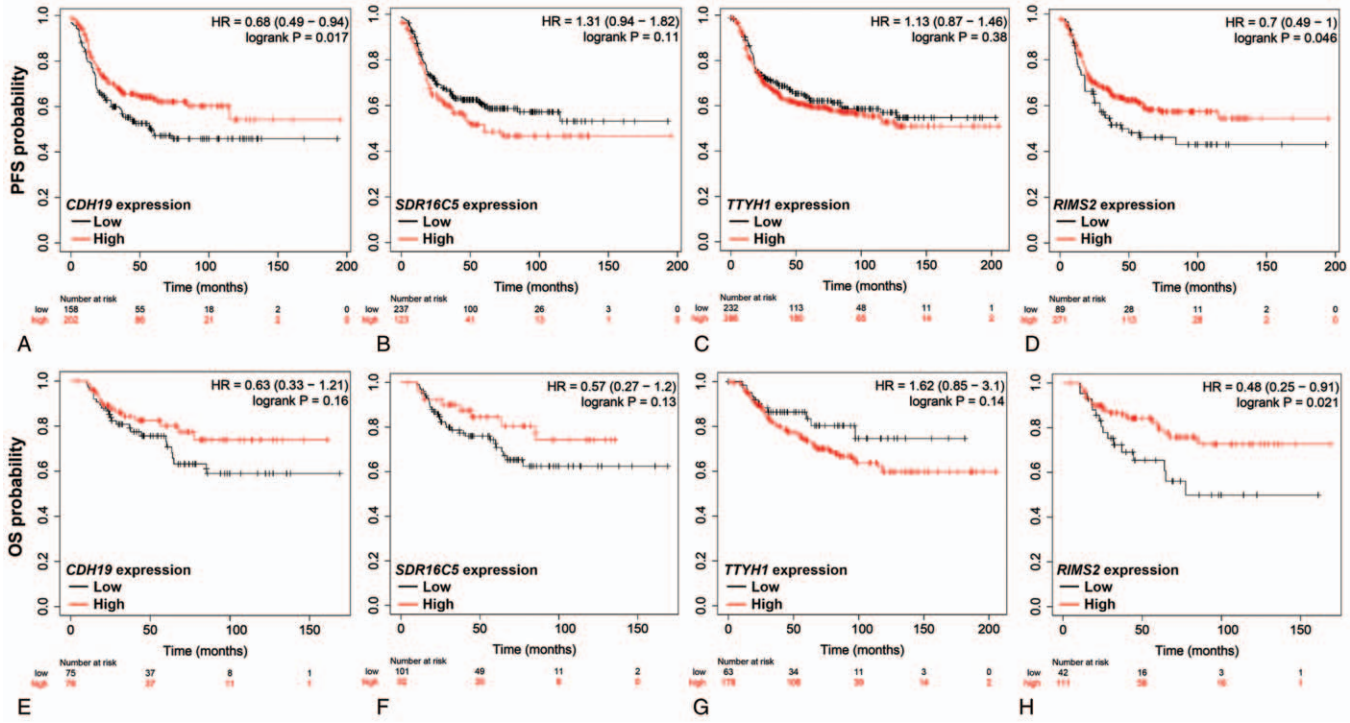

Figure 2. Meta-analysis-based exploration of the prognostic value of progression related genes. (A–H) Kaplan–Meier survival analysis to assess the difference in PFS (A–D) and OS (E–H) in patients with basal-like tumors. Patients were separated into two groups according to the best cutoff of individual gene expression, including *CDH19* (A and E), *SDR16C5* (B and F), *TTYH1* (C and G), and *RIMS2* (D and H).

1.00, $P = .02$) (Fig. 2D) and OS (HR: 0.48, 95%: 0.25–0.91, $P = .02$) (Fig. 2H). However, no significant association was observed in other gene groups (Fig. 2B and C and E and G).

3.3. RIMS2 might be an independent prognostic biomarker for PFS in patients with basal-like tumors

To validate the prognostic value of these two genes, survival analysis was also performed using the basal-like tumor data in TCGA-BRCA. Using the same cutoff in Figure 2, it was noticed that the high *RIMS2* expression group had significantly better PFS and DSS ($P < .01$ and $P = .04$, respectively, Fig. 3A and B). In comparison, the high *CDH19* expression group also had significantly longer PFS (Fig. 3C), compared to the respective low expression groups.

In the following univariate analysis, advanced pathological stages, without radiotherapy, lower *CDH19* expression and lower *RIMS2* expression were potential risk factors of shorter PFS (Table 1). In multivariate analysis, higher *RIMS2* expression, but not *CDH19* expression, was independently associated with longer PFS (HR: 0.778, 95%: 0.642–0.942, $P = .02$), after adjustment for pathological stages, radiotherapy, and *CDH19* expression (Table 1). Currently, chemotherapy and radiotherapy are still among the most effective therapeutic strategies for basal-like breast cancer. In this study, it was found the group without radiotherapy is associated with significantly shorter PFS (Fig. 3D). Multivariate analysis indicated that without radiotherapy was also an independent indicator of shorter PFS (HR: 3.414, 95%: 1.286–9.065, $P = .01$) (Table 1). However, the univariate analysis failed to confirm the independent prognostic value of *RIMS2* expression in DSS (Table 1).

3.4. GSEA and single-cell cellular functional state analysis

To explore the potential mechanisms underlying the association between *RIMS2* expression and favorable PFS in basal-like cases, GSEA was performed using bulk RNA-seq data from TCGA and assessed single cell cellular state using data from GSE75688. GSEA was conducted between the basal-like cases with low (N = 30) and high (n = 90) *RIMS2* expression (Fig. 4A). Results showed that the low *RIMS2* expression group had higher expression of genes significantly enriched in DNA Repair, and MYC Targets V2 (Fig. 4A–C).

Then, the potential contribution of *RIMS2* expression to cellular states was characterized, by calculating the correlation between its expression and 14 functional states of 89 basal-like cells, using data provided by CancerSEA platform. The brief workflow of the platform was as presented in Figure 4D. Fourteen cellular states were estimated according to the gene-expression profile at the single-cell level, using the signatures from Gene Ontology, MSigDB, Cyclebase, HCMDB, and StemMapper. The state activity scores were calculated using the Gene Set Variation Analysis (GSVA)[16] (Fig. 4D). By setting |Pearson's $r \geq 0.30$| as the cutoff, *RIMS2* expression was negatively correlated with DNA repair and EMT (Fig. 4E).

4. Discussion

The current study preliminarily identified ten potential progression-related genes between the progression and no progression groups of basal-like breast tumors. However, single-cell RNA-seq data from 89 TNBC cells only confirmed the expression of four cancer cell-specific genes (*CDH19, SDR16C5, TTYH1,* and *RIMS2*). The inconsistency suggests that the bulk RNA-seq data

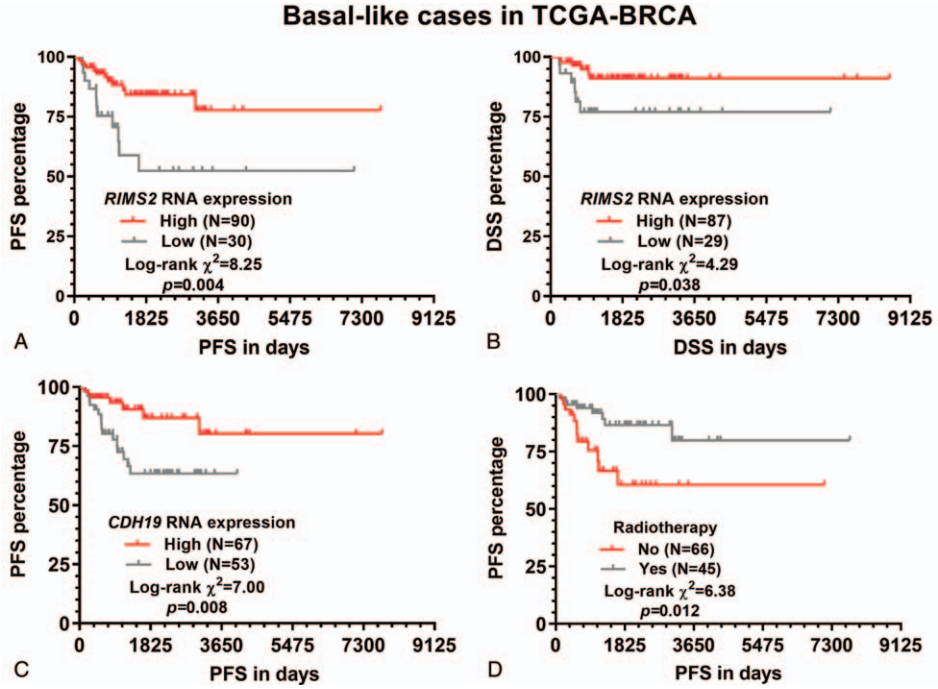

Figure 3. Kaplan–Meier survival analysis using data from TCGA-BRCA. (A–D) Kaplan–Meier survival analysis to assess the difference in PFS (A, C, and D) and DSS (D) in patients with basal-like tumors in TCGA-BRCA. Patients were separated into two groups according to the cutoff of *RIMS2* (A and B) and *CDH19* (C) expression in Figure 2, or by radiotherapy status (D).

is heterogeneous and shows the value of single-cell sequencing in dealing with heterogenicity. Following analysis only assessed the prognostic value of the four tumor-cell associated genes. Using pooled data from Kaplan–Meier plotter and TCGA-BRCA, it was validated that *RIMS2*, but not *CDH19* expression was an independent prognostic indicator of favorable PFS (HR: 0.78, 95%: 0.64–0.95, $P=.015$). *RIMS2* encodes Rab-3-Interacting Molecule 2 in human genome. It is a presynaptic protein that can interact with multiple proteins such as RAB3A, RAB3B, RAB3C, RAB3D, RAB26, ERC1, TSPOAP1, RIMBP2, PPFIA3, PPFIA4, and UNC13B.[20–22] Since it acts as an important modulator of synaptic membrane exocytosis, previous studies typically focused

Table 1

Univariate and multivariate analysis of PFS in basal-like cases in TCGA.

	Univariate analysis				Multivariate analysis			
Parameters	***P***	**HR**	**95%CI (lower/upper)**		***P***	**HR**	**95%CI (lower/upper)**	
RFS								
Age (continuous)	.76	1.005	0.973	1.038				
Pathological stages								
III/IV (N = 17)		1.000						
I/II (N = 1100)	<**.01**	0.236	0.100	0.557	<**.01**	0.201	0.074	0.545
Nodal status								
N1/N2/N3 (N = 46)		1.000						
N0 (N = 74)	.054	0.445	0.195	1.014				
Margin status								
Negative (N = 105)		1.000						
Close/Positive (N = 10)	.08	2.634	0.891	7.790				
Radiation therapy								
Yes (N = 66)		1.000						
No (N = 45)	**.02**	3.034	1.232	7.473	**.01**	3.414	1.286	9.065
Targeted molecular therapy								
Yes (N = 87)		1.000						
No (N = 10)	.62	1.458	0.335	6.348				
CDH19 expression	**.01**	0.834	0.722	0.963	.08	0.868	0.742	1.015
RIMS2 expression	**.01**	0.778	0.642	0.943	**.02**	0.781	0.640	0.953
OS								
RIMS2 expression	.88	0.986	0.817	1.188				

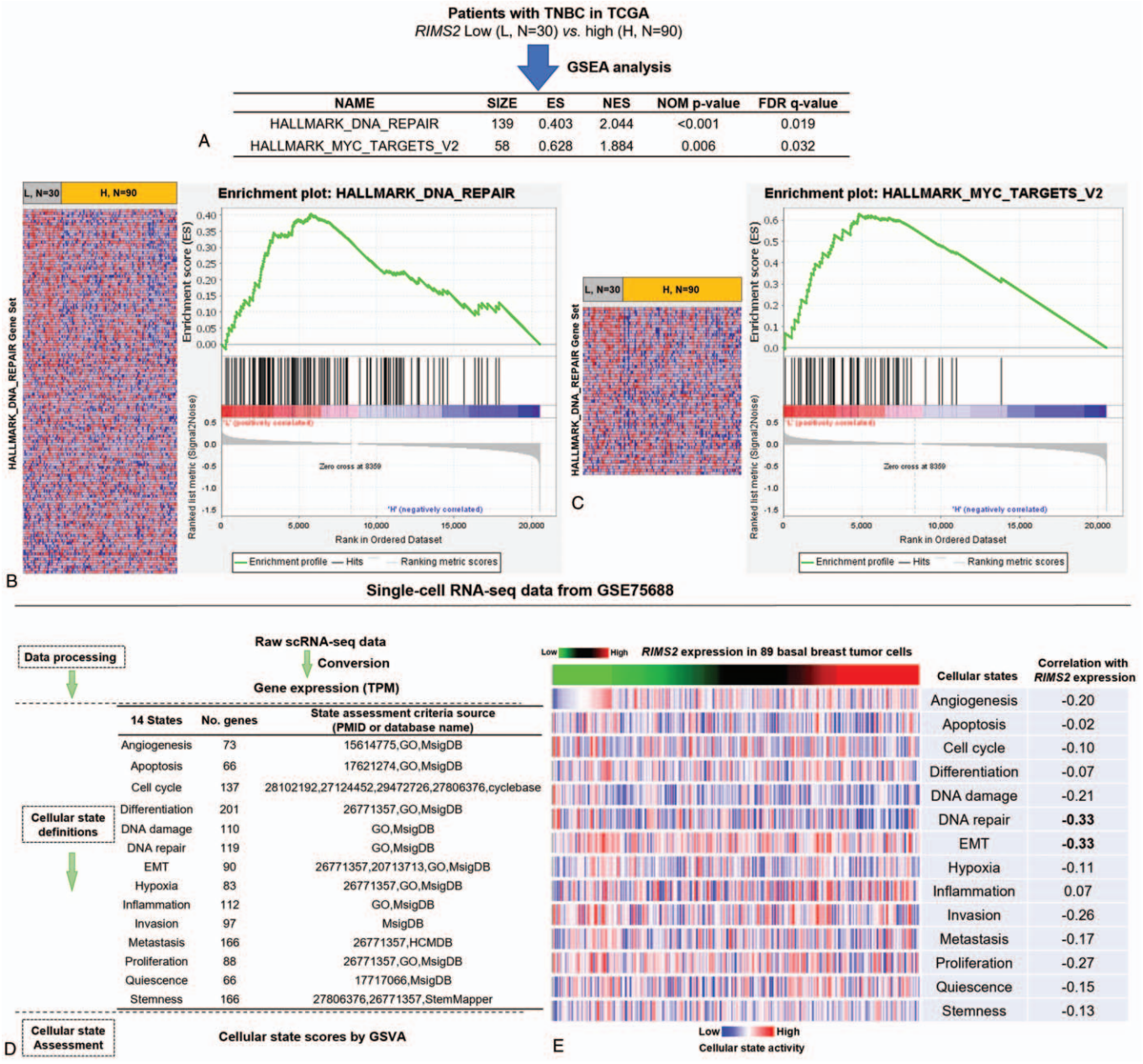

Figure 4. Gene set enrichment analysis (GSEA) and single-cell cellular functional state analysis. (A–C) Group stratification for single gene GSEA in basal-like tumor cases from TCGA-BRCA (A) and summary and individual gene set enrichment plots (B and C) of gene set enrichment in low *RIMS2* expression group. (D) A workflow to show brief GSE75688 data processing, conversion and cellular state assessment. (E) Heatmap and Pearson's correlation analysis showing the correlation between *RIMS2* expression and the 14 cellular functional states in 89 basal-like tumor cells. Green: low expression; red: high expression. Blue: low activity; red: high activity.

on its functional role in brain-associated diseases.[23–25] A few studies reported its potential involvement in tumor biology.[26,27] However, some of its interacting proteins have well-characterized regulatory effects on breast cancer. Via interacting with partner proteins, RIMS2 might participate in quite complex signaling pathways in breast cancer. For example, RAB3A is involved in vesicle trafficking and protein secretion in mammary epithelial cells[28] and is upregulated in breast cancer cells.[29] ERC1 is an important modulator of DNA damage-responsive kinase ATM-dependent NF-kappaB activation in MDA-MB-231 cells (a typically TNBC cell line).[30] ERC1 can also regulate breast cancer cell migration via influencing lamellipodia stability and integrin-mediated focal adhesion.[31]

Both GSEA analysis based on bulk RNA-seq data and cellular state analysis based on single-cell RNA-seq data indicated that *RIMS2* expression was negatively correlated with DNA repair in basal-like tumors. DNA repair is a vital cellular function to maintain genomic integrity and tumor cell survival under stressful conditions.[32,33] BRCA1/2, together with some other genes in the Fanconi Anemia/BRCA repair pathway, play essential roles in DNA repair. They participate in re-establishing DNA replication following double-strand DNA breaks (DSBs)

via the coordination of nucleotide excision repair (NER), translesional synthesis, and homologous recombination (HR).[34] Around 60% to 69% TNBC cases have hereditary mutations in BRCA1/2.[33,35] Therefore, this subtype of tumor is supposed to be sensitive to the DNA-damaging agents and therapies. Targeting DNA repair has been considered as a potential strategy to improve the therapeutic effect of radiotherapy in TNBC.[33] In this study, it was found that radiotherapy could significantly improve PFS in patients with basal-like tumors, suggesting that radiotherapy is an effective therapy. The negative correlation between *RIMS2* expression and DNA repair activity of the cells might explain the favorable survival outcomes of the high expression group.

This study also has some limitations. Although this study identified the involvement of *RIMS2* in DNA repair, the exact mechanisms were not explored. GSEA data suggest that the low *RIMS2* expression group had higher expression of MYC target genes. A series of MYC target genes are involved in breast cancer pathological development and therapeutic responses, such as BRCA1,[36] CDK18,[37] PARP1, and PARP2.[38] Therefore, it would be interesting to explore their potential associations in future studies.

5. Conclusions

RIMS2 expression was negatively associated with DNA repair capability of basal-like breast tumor cells and might serve as an independent indicator of favorable PFS.

Author contributions

Conceptualization: Lingyun Zhang.
Data curation: Lingyun Zhang, Zheng Liu.
Formal analysis: Lingyun Zhang, Zheng Liu.
Methodology: Lingyun Zhang, Zheng Liu.
Project administration: Jingqiang Zhu.
Resources: Jingqiang Zhu.
Software: Lingyun Zhang.
Supervision: Jingqiang Zhu.
Validation: Lingyun Zhang, Zheng Liu.
Visualization: Lingyun Zhang.
Writing – original draft: Lingyun Zhang, Zheng Liu.
Writing – review & editing: Lingyun Zhang, Jingqiang Zhu.

References

[1] Nielsen T, Wallden B, Schaper C, et al. Analytical validation of the PAM50-based Prosigna Breast Cancer Prognostic Gene Signature Assay and nCounter Analysis System using formalin-fixed paraffin-embedded breast tumor specimens. BMC Cancer 2014;14:177.
[2] Lee KL, Kuo YC, Ho YS, et al. Triple-negative breast cancer: current understanding and future therapeutic breakthrough targeting cancer stemness. Cancers (Basel) 2019;11: doi: 10.3390/cancers11091334.
[3] Shah SP, Roth A, Goya R, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. Nature 2012;486:395–9.
[4] Navin NE. The first five years of single-cell cancer genomics and beyond. Genome Res 2015;25:1499–507.
[5] Liedtke C, Mazouni C, Hess KR, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. J Clin Oncol 2008;26:1275–81.
[6] Choo JR, Nielsen TO. Biomarkers for basal-like breast cancer. Cancers (Basel) 2010;2:1040–65.
[7] Choi W, Porten S, Kim S, et al. Identification of distinct basal and luminal subtypes of muscle-invasive bladder cancer with different sensitivities to frontline chemotherapy. Cancer Cell 2014;25:152–65.
[8] Gracio F, Burford B, Gazinska P, et al. Splicing imbalances in basal-like breast cancer underpin perturbation of cell surface and oncogenic pathways and are associated with patients' survival. Sci Rep 2017;7:40177.
[9] Cancer Genome Atlas N. Comprehensive molecular portraits of human breast tumours. Nature 2012;490:61–70.
[10] Hudson TJ, Anderson W, Artez A, et al. International Cancer Genome CInternational network of cancer genome projects. Nature 2010;464:993–8.
[11] Shalek AK, Satija R, Shuga J, et al. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature 2014;510:363–9.
[12] Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. Exp Mol Med 2018;50:96.
[13] Olsen TK, Baryawno N. Introduction to single-cell RNA sequencing. Curr Protoc Mol Biol 2018;122:e57.
[14] Goldman M, Craft B, Hastie M, et al. The UCSC Xena platform for public and private cancer genomics data visualization and interpretation 2019;326470.
[15] Liu J, Lichtenberg T, Hoadley KA, et al. An integrated TCGA Pan-cancer clinical data resource to drive high-quality survival outcome analytics. Cell 2018;173:400–16.e11.
[16] Yuan H, Yan M, Zhang G, et al. CancerSEA: a cancer single-cell state atlas. Nucleic Acids Res 2019;47(D1):D900–8.
[17] Chung W, Eum HH, Lee HO, et al. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. Nat Commun 2017;8:15081.
[18] Gyorffy B, Lanczky A, Eklund AC, et al. An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. Breast Cancer Res Treat 2010;123:725–31.
[19] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005;102:15545–50.
[20] Fukuda M. Distinct Rab binding specificity of Rim1, Rim2, rabphilin, and Noc2. Identification of a critical determinant of Rab3A/Rab27A recognition by Rim2. J Biol Chem 2003;278:15373–80.
[21] Bello OD, Zanetti MN, Mayorga LS, et al. and Rab3A interplay in acrosomal exocytosis. Exp Cell Res 2012;318:478–88.
[22] Shin H, Wyszynski M, Huh KH, et al. Association of the kinesin motor KIF1A with the multimodular protein liprin-alpha. J Biol Chem 2003;278:11393–401.
[23] Kim KT, Lee JS, Lee BW, et al. Association between regulating synaptic membrane exocytosis 2 gene polymorphisms and degenerative lumbar scoliosis. Biomed Rep 2013;1:619–23.
[24] Weidenhofer J, Scott RJ, Tooney PA. Investigation of the expression of genes affecting cytomatrix active zone function in the amygdala in schizophrenia: effects of antipsychotic drugs. J Psychiatr Res 2009;43:282–90.
[25] Fan Y, Du X, Liu X, et al. Rare copy number variations in a Chinese cohort of autism spectrum disorder. Front Genet 2018;9:665.
[26] Mukasa A, Ueki K, Ge X, et al. Selective expression of a subset of neuronal genes in oligodendroglioma with chromosome 1p loss. Brain Pathol 2004;14:34–42.
[27] Shanmugam C, Katkoori VR, Jhala NC, et al. Immunohistochemical expression of rabphilin-3A-like (Noc2) in normal and tumor tissues of human endocrine pancreas. Biotech Histochem 2009;84:39–45.
[28] Vadlamudi RK, Wang RA, Talukder AH, et al. Evidence of Rab3A expression, regulation of vesicle trafficking, and cellular secretion in response to heregulin in mammary epithelial cells. Mol Cell Biol 2000;20:9092–101.
[29] Lodhi SS, Farmer R, Singh AK, et al. 3D structure generation, virtual screening and docking of human Ras-associated binding (Rab3A) protein involved in tumourigenesis. Mol Biol Rep 2014;41:3951–9.
[30] Alpay M, Backman LR, Cheng X, et al. Oxidative stress shapes breast cancer phenotype through chronic activation of ATM-dependent signaling. Breast Cancer Res Treat 2015;151:75–87.
[31] Chiaretti S, de Curtis I. Role of liprins in the regulation of tumor cell motility and invasion. Curr Cancer Drug Targets 2016;16:238–48.
[32] Zhang Y, He Q, Hu Z, et al. Long noncoding RNA LINP1 regulates repair of DNA double-strand breaks in triple-negative breast cancer. Nat Struct Mol Biol 2016;23:522–30.
[33] Paluch-Shimon S, Evron E. Targeting DNA repair in breast cancer. Breast 2019;47:33–42.
[34] Venkitaraman AR. Linking the cellular functions of BRCA genes to cancer pathogenesis and treatment. Annu Rev Pathol 2009;4:461–87.

[35] Farmer H, McCabe N, Lord CJ, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature 2005;434: 917–21.
[36] Chen Y, Xu J, Borowicz S, et al. c-Myc activates BRCA1 gene expression through distal promoter elements in breast cancer cells. BMC Cancer 2011;11:246.
[37] Ning JF, Stanciu M, Humphrey MR, et al. Myc targeted CDK18 promotes ATR and homologous recombination to mediate PARP inhibitor resistance in glioblastoma. Nat Commun 2019;10:2910.
[38] Zhang W, Liu B, Wu W, et al. Targeting the MYCN-PARP-DNA damage response pathway in neuroendocrine prostate cancer. Clin Cancer Res 2018;24:696–707.