



OPEN

# Prognostic risk model under the immune-associated long chain non-coding ribonucleic acid and its application in survival prognosis assessment of patients with breast cancer

Shuo Yang &amp; Qing Wang

This study aimed to develop a prognostic risk model based on immune-related long non-coding RNAs (lncRNAs). By analyzing the expression profiles of specific long non-coding RNAs, the objective was to construct a predictive model to accurately assess the survival prognosis of breast cancer (BC) patients. This effort seeks to provide personalized treatment strategies for patients and improve clinical outcomes. Based on the median risk value, 300 samples of triple-negative BC (TNBC) patients were rolled into a high-risk group (HR group,  $n = 140$ ) and a low-risk group (LR group,  $n = 160$ ). Multivariate Cox (MVC) analysis was performed by combining the patient risk score and clinical information to evaluate the prognostic value of the prognostic risk (PR) model. A total of 371 immune-related lncRNAs associated with the prognosis of TNBC were obtained from 300 TNBC samples. Nine associated with prognosis were obtained by univariate Cox (UVC) analysis, and 3 (AC090181.2, LINC01235, and LINC01943) were selected by MVC analysis for the construction of TNBC PR model. Survival analysis showed a great difference in TNBC patients in different groups ( $P < 0.001$ ). The receiver operator characteristic (ROC) curve showed the model possessed a good area under ROC curve (AUC), which was 0.928. The patient RS jointing with clinical information as well as the MVC analysis revealed that RS was an independent risk factor (IRF) for prognosis of TNBC ( $P < 0.05$ , HR = 1.033286). Therefore, the lncRNAs associated with TNBC immunity can be screened by bioinformatics analysis, and the established PR model of TNBC could better predict the prognosis of patients with TNBC, exhibiting a high application value in clinic.

**Keywords** Breast cancer, Immune-related genes, Long chain non-coding ribonucleic acids, Prognosis, Prognostic risk model

Breast cancer (BC) is a commonly confirmed malignant tumors in female population, ranking first among all female cancers, with the second mortality rate<sup>1</sup>. The main type of BC is invasive ductal carcinoma. However, according to the estrogen receptor (ER), progesterone receptor, human epidermal growth factor receptor 2 are followed (HER-2), Ki-67 proliferation index, and gene expression profile, BC can be classified into Luminal A and B, over-expression of HER-2, and triple-negative BC (TNBC)<sup>2</sup>. TNBC is featured with the absence of ER and progesterone receptor and the negative expression of HER-2. Although the number of patients with BC is only 20 percent of BC patients, the number of deaths is 80 percent. This indicates that the characteristics of TNBC patients are poor prognosis, high malignancy, strong invasiveness, strong heterogeneity, early recurrence and metastasis, high drug resistance, and short overall survival (OS)<sup>3</sup>.

Despite the limitations in treatment modalities, particularly the lack of targeted therapies for TNBC patients, leading to their limited benefit from endocrine therapy or anti-HER-2 treatment, current therapy primarily relies on systemic chemotherapy. Although chemotherapy can yield certain efficacy against TNBC, this subtype of BC exhibits a higher recurrence rate and propensity for metastasis. Clinical data indicate a low efficacy rate of

Shaoxing People's Hospital, Shaoxing City 312000, China. email: nancy2023123@163.com

chemotherapy, highlighting the limitations of existing treatment strategies and the urgency to enhance treatment outcomes<sup>4</sup>. Given this context, exploring and implementing more effective treatment regimens to optimize the survival prognosis of TNBC patients is paramount. In recent years, the immunomodulatory effects within the tumor microenvironment (TME) have increasingly become a focus of research, as they not only regulate the dynamics of tumor development but also directly influence patient prognosis<sup>5</sup>. Particularly in TNBC, due to its unique biological characteristics and active immune features, immunotherapy is considered a highly promising target, demonstrating the potential to overcome the limitations of traditional treatment by activating or enhancing immune responses<sup>6</sup>.

In this scientific context, long non-coding RNAs (lncRNAs), as an important and complex class of molecules, have garnered increasing attention for their roles in regulating immune responses and influencing disease progression (Bridges, Daulagala, & Kourtidis,<sup>7</sup>. These molecules, with lengths exceeding 200 nucleotides, do not encode proteins but are extensively involved in and regulate various immune processes, including immune cell functions, antigen presentation, and immune evasion<sup>8</sup>. As a crucial component of the transcriptome, lncRNAs play indispensable roles in regulating cellular physiology and pathology, particularly in tumorigenesis and tumor progression (Hang, Guan, & Tang, 2021). They finely modulate gene expression networks by interacting with various molecules such as DNA, RNA, and proteins, thereby influencing cellular fate determinations (Zhang,<sup>11</sup>, Yang,<sup>4</sup>, Xiu,<sup>12</sup>. In the field of tumor immunoregulation, the significant role of lncRNAs is particularly evident, as their interactions with immune-related genes are gradually being elucidated, providing new perspectives for understanding tumor immune evasion and developing novel therapeutic strategies<sup>9</sup>. Given the regulatory roles of lncRNAs at various stages of TNBC development, especially their potential to promote or inhibit tumor progression by influencing the immune microenvironment, systematic exploration of lncRNA expression characteristics associated with immune regulation is of significant importance for identifying new prognostic biomarkers and guiding personalized immune therapy strategies<sup>10</sup>. However, current research on the specific mechanisms of lncRNAs in TNBC prognosis prediction, particularly their association with immune regulatory functions, remains relatively scarce, necessitating further exploration and validation. In summary, this study aimed to fill this gap by comprehensively analyzing the expression profiles of immune-related lncRNAs, constructing a prognostic risk model aimed at accurately predicting the survival prognosis of TNBC patients, providing scientific evidence for clinical practice, promoting the development of personalized treatment regimens, and ultimately improving patient prognosis and quality of life.

Therefore, the immuno-related lncRNAs in TNBC that were significantly associated with prognosis were screened in this work, and a prognostic risk (PR) model was established. It aimed to explore the predictive ability of this model for the prognosis of TNBC patients and the independent prognostic value of this model to guide the clinical immune-related lncRNAs associated with the prognosis.

## Research objects

### Data source

Data acquisition: raw transcriptomic RNA expression profiles of female TNBC tissue and para-cancer tissue samples were downloaded from The Cancer Genome Atlas (TCGA) database (all samples included complete high-throughput counts). The clinical and pathological information of TNBC patients, including survival information, age and TNM stage, was also downloaded. The clinical information was processed to obtain the clinical information of target patients. The exclusion criteria applied to the patients were as follows: (1) survival information of the patient including survival time and survival state was incomplete; (2) survival time of the patient was less than 30 days; (3) clinicopathological parameters included incomplete T, M, and N stages of tumors. In this study, all patients who fully met the inclusion criteria received the trial regulations and signed the informed consent form for the trial. The study was approved by the Medical Ethics Committee of Shaoxing People's Hospital. All methods were performed in accordance with the relevant guidelines and regulations.

Pretreatment: R language was utilized to organize the expression of expression matrix of transcriptomic RNA data of all TNBC samples and para-cancer samples. The expression matrix of TNBC samples was extracted by id transformation of transcriptome expression profile for downloaded Counts data. Perl script was employed to extract the expression matrix of TNBC lncRNAs. The repetitive genes were averaged, and the low-expressed genes were eliminated. The expression matrix of TNBC samples and adjacent tissue samples was grouped. In addition, the R language “limma” package was utilized to analyze the differences in expression matrix of lncRNAs, the multiple parameter differences were drafted in more than 2 times ( $|\log_2FC| > 2$ ), and the corrected  $P$  values  $< 0.01$  was set as the screening criteria. The significant difference matrix was extracted to obtain the differential genes. With the help of the “pheatmap” package in R software, the volcano map can more directly show the significantly different gene expression distribution in TNBC tissue and para-cancer tissue.

### Screening of lncRNAs associated with prognosis of TNBC

(1) TNBC immuno-related lncRNAs: the IRG set was downloaded from the Immport database (<https://www.immport.org/>). Perl script was utilized to obtain the TNBC IRGs, which were intersected with the differentially expressed genes of TNBC samples and para-cancer samples previously obtained. Pearson correlation analysis (PCA) and co-expression method were adopted, and the TNBC immune-related lncRNAs were obtained with  $Cor > 0.5$  and  $P < 0.001$ .

(2) Immuno-related lncRNAs associated with TNBC prognosis: univariate Cox (UVC) analysis was conducted on immuno-related lncRNAs expression matrix of 300 TNBC patients by survival package of R software. Then, the relationship between OS and differentially expressed immuno-related lncRNAs in TNBC patients was evaluated.  $P < 0.05$  was set as the filtering condition to obtain the immune-related lncRNAs associated with the prognosis.

## Construction of PR model for TNBC

### (1) Establishment of PR model.

The optimization of data was guided by the Akaike Information Criterion (AIC) value, aiming to selectively identify immune-related lncRNAs that could reliably forecast the prognosis of TNBC patients, thereby establishing the most effective predictive scoring model. Subsequently, regression coefficients were computed, yielding the model expression equation (Eq. (1)), and a PR model was formulated to assess the prognosis of TNBC patients.

$$(\text{Risk Score}) = \sum (\text{EXPi} * \text{Coefi}) \quad (1)$$

In Eq. (1), EXP is the fragments per kilobase million (FPKM) value of IRGs; and Coef refers to a MUC regression coefficient.

### (2) Visualization of the relationship between RS and survival state prognosis (SSP).

Based on the PR model, the levels of 3 lncRNAs in each patient with TNBC were substituted into the equation to calculate the RS value. The samples were grouped into a high-risk (HR) and a low-risk (LR) group according to the median of RS values. The higher the RS value, the higher the risk, and the worse the prognosis. The RS curve and the SSP relationship scatter graphs were drawn to visualize the results.

### (3) Verification of PR model.

Survival analysis was conducted on both groups to assess the disparity in OS between patients categorized into HR and LR groups. Furthermore, Kaplan–Meier (KM) analysis was employed alongside the Log-rank test to generate survival curves, facilitating the comparative analysis of OS disparities between patients in the HR and LR groups.

The area under the receiver operating characteristic (ROC) curve (AUC) was computed by plotting the ROC curve for survival prediction. The predictive power of the prediction model was judged by the AUC value (a value of < 0.5 meant the model exhibited no predictive power, that of 0.51–0.7 suggested a low accuracy, that of 0.71–0.9 indicated a medium accuracy, and the accuracy was high when it > 0.9). ROC curves were plotted using the “timeROC” package.

## Evaluation of prognostic value by PR model

The UVC analysis was conducted on the RS value, the multi-gene prediction model, and clinicopathological information (age, clinical staging, T stage, and N stage) of TNBC patients. Then, UVC analysis was integrated into the MVC analysis, with a significance level of  $P > 0.05$  serving as the exclusion criterion for variable selection. Finally, the prognostic value of RS model was clarified.

## Methods for statistics

All statistical analysis and data visualization were realized with R software (version 3.5.1 or above). The multiple analysis was employed for variance analysis. The difference ratio was defined as  $|\log_2 \text{FC}| > 2$ , and  $P$  values (FDR) < 0.01. The immune-related lncRNAs with prognostic value were obtained by PCA, UVC analysis, and MVC analysis. Meanwhile, the KM method was utilized to draw the survival curve, and survival prediction ROC curve was drawn to verify the model. Risk value of each patient was combined with clinicopathological data, and the independent prognostic value of PR model was evaluated by UVC and MVC analysis methods.

## Research results

### Data processing results

The original transcriptome RNA expression profile data and related clinical information of 300 female TBC tissue samples and 110 para-carcinoma tissue samples were downloaded using the methods described in section “Data source”. Finally, the clinical information of 300 target patients was obtained, among which 38.33% were over 300 years old. Among the pathological stages, clinical staging II was 71.67%, T2 was 76.67%, N0 was 70%, and M0 was 91.67%, accounting for the largest proportion in each stage. The specific results were listed in Table 1.

### Data preprocessing result

If multiple parameters difference in more than 2 times ( $|\log_2 \text{FC}| > 2$ ) and  $P$  value (PDF) < 0.05, it was determined as differentially expressed gene (DEG). According to this criterion, 1131 DEGs were obtained, among which 189 were highly expressed and 182 were lowly expressed, which were marked with red and green, respectively, in Fig. 1. Figure 1 showed a specific volcano map that visualizes the DEGs. The horizontal coordinate was the negative logarithm of  $P$  value after correction, and the  $\log_2 \text{FC}$  result was set as the vertical coordinate.

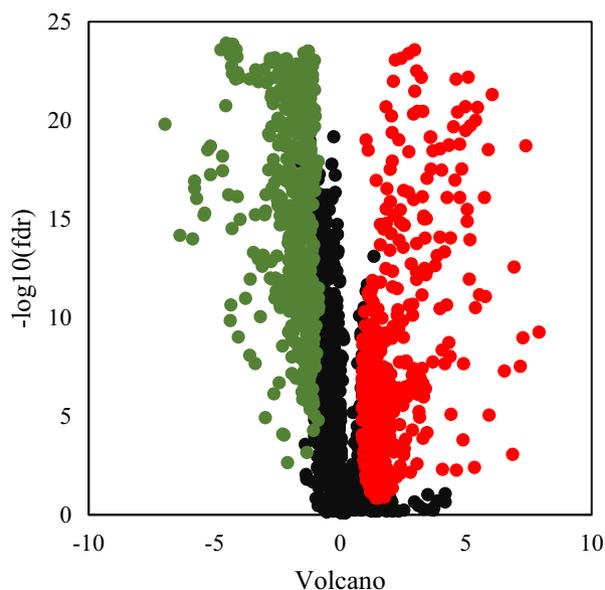
2,504 IRGs were downloaded from the Immport database. By combining the above two methods, the immune lncRNAs were extracted and identified by PCA with  $\text{Cor} > 0.5$  and  $P < 0.05$  as criteria. Finally, 371 immune-related lncRNAs were obtained.

### Screening results of immune-related lncRNAs

Relationship between immune-related lncRNAs and survival time was determined by UVC analysis. With survival time and survival state as dependent variables, 9 immune-related lncRNAs were screened out (including USP30-AS1, AC090181.2, HCP5, LINC01235, AL354920.1, MIAT, AC004847.1, TRG-AS1, and LINC01943). All of them were proved to be correlated with the OS of TNBC patients ( $P < 0.01$ ), as displayed in Table 2. Therefore, they were determined to be immuno-related lncRNAs associated with the prognosis of TNBC. Differentially expressed genes (DEGs) refer to genes whose expression levels vary significantly under different conditions, such as between normal tissues and tumor tissues, or among different prognostic groups. In cancer research,

Index	Number (n (%))	
> 60 years old	115 (38.33%)	
Clinical staging	I	55 (18.33%)
	II	215 (71.67%)
	III	30 (10%)
	IV	0 (0%)
T stage	T1	40 (13.33%)
	T2	230 (76.67%)
	T3	20 (6.67%)
	T4	10 (3.33%)
	Tx	0 (0%)
N stage	N0	210 (70%)
	N1	60 (20%)
	N2	15 (5%)
	N3	15 (5%)
	Nx	0 (0%)
M stage	M0	275 (91.67%)
	M1	0 (0%)
	Mx	25 (8.33%)

**Table 1.** Brief introduction of 300 target patients.



**Figure 1.** Differential expression of TNBC gene. Red and green dots refer to the highly and lowly DGEs, and black dots are not the DGEs.

DEG analysis can reveal key molecules involved in disease onset, progression, and prognosis. Immune-related DEGs, particularly lncRNAs, whose aberrant expression is often closely associated with tumor microenvironment remodeling, immune evasion mechanisms, and patient responsiveness to treatment. The differential expression of these genes may be driven by epigenetic alterations, transcriptional dysregulation, or other molecular events, thereby impacting cellular physiological functions and immune surveillance capabilities.

#### Construction results of PR model

After MVC analysis, three immuno-related lncRNAs, specifically AC090181.2, LINC01235, and LINC01943, which were greatly correlated to the prognosis of TNBC were finally screened out according to the optimal AIC value. Based on them, the PR model of TNBC was established. LINC01235 showed a positive coefficient ( $HR > 1$ ), indicating a risk factor whose high expression was associated with shorter survival. AC090181.2 and LINC01943 showed negative coefficients ( $HR < 1$ ), indicating that they were protective factors, and their high

ID	HR	HR.95L	HR.95H	P
USP30-AS1	0.264892341	0.074275203	0.632572169	0.004872934
AC090181.2	0.001518648	2.0701E13	0.106277434	0.003391782
HCP5	0.470632736	0.279173634	0.781736285	0.002873024
LINC01235	1.417643829	1.082974837	1.771632394	0.010732837
AL354920.1	3.560153627	1.378437672	9.201880246	0.015374985
MIAT	0.111653631	0.021836762	0.531683894	0.008621632
AC004847.1	0.047436876	0.014672943	0.408372801	0.006374939
TRG-AS1	0.047938462	0.005002188	0.469832649	0.010190367
LINC01943	0.079015732	0.0123514934	0.521435681	0.009190834

**Table 2.** LncRNA UVC analysis results.

ID	Regression coefficient	HR	HR.95L	HR.96H	P
AC090181.2	-6.726490	0.002081	1.92E-11	0.219742	0.008102
LINC01235	0.257385	1.286387	0.928618	1.718638	0.117831
LINC01943	-2.207814	0.100398	0.020738	0.600936	0.009739

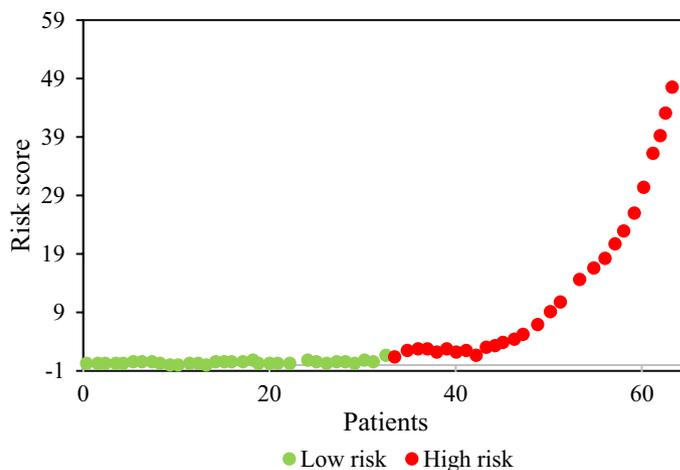
**Table 3.** LncRNA MVC results.

expression reflected a longer survival in patients, as displayed in Table 3. The TCBA PR model was constructed according to the levels of these 3 lncRNAs.

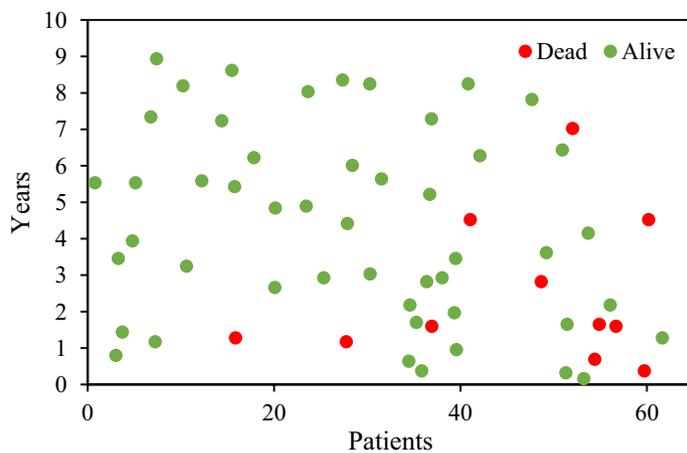
This prognostic model was employed to calculate the RS of 60 patients with TNBC who had met the conditions and complete clinical information. The expression of lncRNA selected by each patient was substituted by the established prognostic model expression equation to obtain the corresponding RS, and then the risk curve of patients was obtained. The abscissa of the curve was the ranking of patients, the ranking method was from left to right according to the risk score, and the ordinate was the risk value. Taking the median RS as the critical point, 300 patients were rolled into a HR group ( $n = 140$ ) and a LR group ( $n = 160$ ), as demonstrated in Fig. 2. The relationship between RS and prognosis obtained from the model was represented (Fig. 3) by a scatter plot of high follow-up endpoint and LR group survival status, with survival status represented in green and death status represented in red. As can be observed from the left to the right, there were more and more red parts, and the red parts were the most in the HR group.

### Verification results of PR model for TNBC

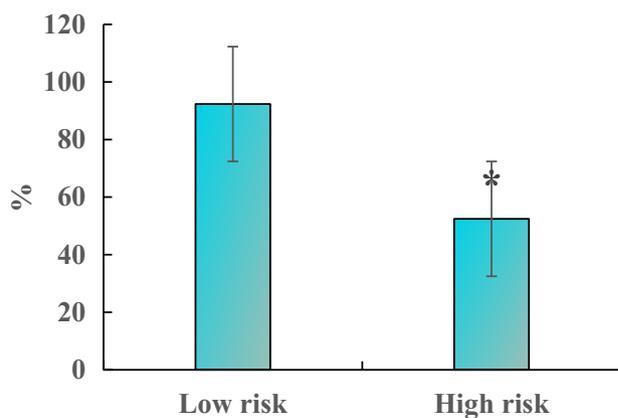
The KM survival analysis suggested that the median survival period (SP) of patients with  $SP \geq 30$  days was much shorter in the HR group in comparison with that in the LR group. The two survival curves were obviously separated, and Log-Rak test showed an extremely great difference ( $P < 0.001$ ). In addition, the survival rate of patients



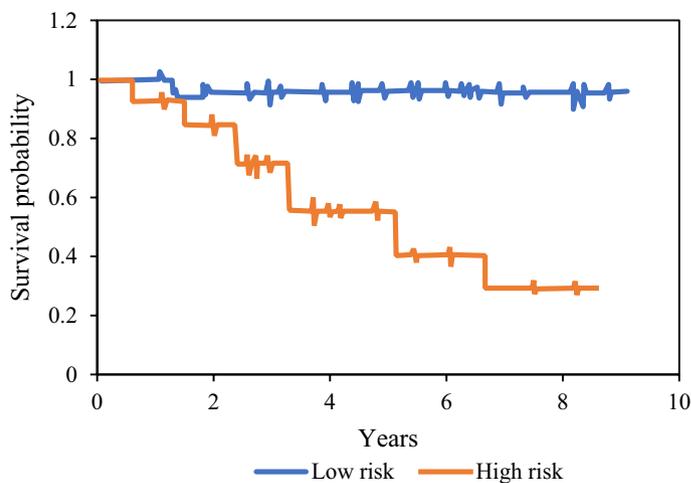
**Figure 2.** Scoring curve of risk model.



**Figure 3.** Scatter plot of the relationship between RS and prognosis.



**Figure 4.** The 5 year survival rates of two patient groups. (\* indicates significance compared to the low-risk group,  $P < 0.05$ ).



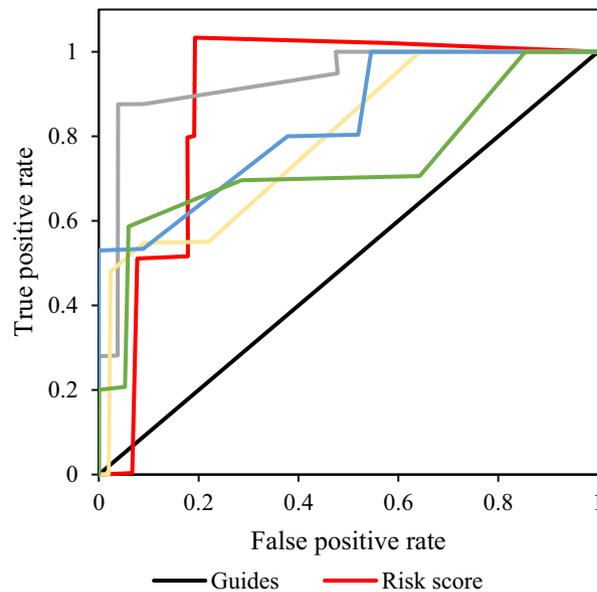
**Figure 5.** KM survival curve of the PR model.

in the LR group was higher and exhibited an extremely obvious difference with that in the HR group ( $P < 0.001$ ). It indicated a great difference in survival rate between patients in different groups, as illustrated in Figs. 4, 5.

It is believed that the higher the value of the AUC, the stronger the prediction ability, and the more accurate the prediction effect. To verify the predictive ability of PR model constructed by three immuno-related lncRNAs, ROC analysis was conducted on the model, which revealed that the predictive ability and accuracy of PR model were good, with an AUC of 0.928. Figure 6 illustrated the above results clearly.

### Independent prognostic value assessment of PR model

The constructed multi-gene PR model for TNBC patients was combined with clinicopathologic information and RS, which were processed by MVC analysis method. The results indicated that PR model constructed by immuno-related lncRNAs obviously correlated with the prognosis of TNBC was not associated with other clinical information ( $P > 0.05$ ). UVC implied that RS was a prognostic factor for TNBC ( $HR = 1.028745$ ,  $P < 0.05$ ), as displayed in Table 4. MVC analysis also told that RS was the prognostic IRF of TNBC patients ( $HR = 1.033286$ ,  $P < 0.05$ ) (Table 5). The UVC analysis is a statistical test method used to assess the association between categorical variables and outcomes, suitable for comparing multiple independent sample groups. In this study, UVC was employed to examine the association between the prognostic risk model based on immune-related



**Figure 6.** ROC curve of the PR model.

Index	HR	HR.95L	HR.96H	P
Age (years old)	0.957373	0.927844	1.194835	0.973824
Clinical staging	19.670391	4.892639	78.937462	0.000963
T stage	3.198741	1.028746	9.749845	0.043187
N stage	4.617384	2.643899	9.647383	0.000285
PR model scoring	1.028745	1.007832	1.068386	0.000017

**Table 4.** UVC analysis results of PR model.

Index	HR	HR.95L	HR.96H	P
Age (years old)	0.982673	0.937910	1.0510003	0.763993
Clinical staging	1.754926	0.108748	17.868439	0.608392
T stage	1.387649	0.510834	3.689301	0.501892
N stage	3.501007	0.831985	14.198763	0.097387
PR model scoring	1.033286	1.002751	1.072931	0.013678

**Table 5.** MVC analysis results of PR model.

lncRNAs and the OS of TNBC patients. The results demonstrated a significant correlation between the model and patient prognosis ( $P < 0.01$ ), indicating the efficacy of the selected lncRNAs in predicting patient survival outcomes. Additionally, UVC suggested that the RS is an independent prognostic factor for TNBC prognosis (HR = 1.028745,  $P < 0.05$ ), implying a direct correlation between high or low RS and patient prognosis, where higher RS typically corresponds to poorer survival expectations. In contrast, MVC is a more complex statistical method that evaluates the impact of certain variables on survival outcomes while considering other variables, such as clinical and pathological information. In this study, MVC analysis not only validated the association between the PR model constructed using immune-related lncRNAs and prognosis but also further investigated the importance of RS as an independent prognostic indicator, even after adjusting for other clinical variables ( $P < 0.05$ , HR = 1.033286). Here, “prognostic IRF” may refer to “prognostic impact risk factor,” emphasizing that RS is a robust predictive factor with a significant impact on patient prognosis.

### External validation of the model

The PR model of this study underwent external validation to assess its generalization capability. The results revealed an AUC of 0.853 in the external validation set, which slightly decreased compared to the internal validation set's AUC of 0.928. Nevertheless, it remains  $> 0.7$ , indicating the robust external generalization capability of the PR model constructed in this study.

## Discussion

### Research status of TNBC

As a variation of BC, TNBC exhibits negative progesterone receptor, ER, and Her-2. Due to its special biological behavior, strong heterogeneity, and insufficient therapeutic targets, TNBC presents extremely high drug resistance, faster progression rate and poor prognosis<sup>13</sup>. TNBC is common in premenopausal young women, and although it affects only one-fifth of all BC patients, it accounts for 80% of all BC deaths. Compared with other subtypes of BC, TNBC patients have a shorter SP, with a mortality rate of up to 40% within 5 years after diagnosis<sup>14</sup>. TNBC is highly invasive. About 46% of TNBC patients will develop distant metastases, which generally occur within 3 years of diagnosis. The average survival time after metastasis is less than 14 months, and metastasis generally involves intracranial and internal organs. The average recurrence time of BC patients with other subtypes ranges from 35 to 67 months, while the recurrence time of TNBC patients is significantly shortened to 19–40 months, and the mortality rate can reach 75% within 3 months after recurrence, indicating that the prognosis of TNBC patients is significantly worse than that of other subtypes<sup>15</sup>. Therefore, the search for biological molecular markers for early diagnosis, risk assessment and prognosis assessment of TNBC is very important for clinical treatment of TNBC. Several lncRNAs, such as USP30-AS1, AC090181.2, HCP5, and LINC01235, have been closely associated with the progression and prognosis of TNBC and have been demonstrated to independently predict the OS of TNBC patients. These lncRNAs exert their influence on tumor behavior by modulating pathways including immune response, cell proliferation, apoptosis, and signal transduction, thereby providing novel insights for early diagnosis and prognosis assessment of TNBC. Circular RNAs, as a special form of non-coding RNA, are increasingly recognized for their roles in disease progression. For instance, circSEPT9 has been identified as a potential biomarker and therapeutic target for TNBC, and through high-throughput RNA sequencing technology and subsequent validation, the abnormal expression of such molecules in TNBC cells and tissues may indicate their potential value in disease diagnosis and treatment. In TNBC, the incidence of BRCA1/2 gene mutations is relatively high, particularly BRCA1 mutations. Women carrying BRCA1 mutations have a significantly increased risk of developing TNBC, which not only has a significant impact on genetic counseling and risk assessment but also provides a basis for the adoption of targeted therapies such as PARP inhibitors. Characteristics of the immune microenvironment, such as PD-L1 expression levels and the quantity of tumor-infiltrating lymphocytes (TILs), are also considered important indicators of TNBC prognosis. Immune checkpoint inhibitors have shown therapeutic efficacy in some TNBC patients, further emphasizing the role of immune-related biomarkers in TNBC treatment decision-making. Gene expression profile analysis, including but not limited to PAM50 subtype classification and multi-gene scoring, can aid in identifying TNBC subgroups with distinct biological behaviors and prognosis characteristics. Basal-like TNBC is typically associated with poorer prognosis, and through specific gene expression profiling analysis, patients can be more accurately stratified, providing guidance for personalized treatment. This also encompasses microRNAs (miRNAs), DNA methylation, protein biomarkers, etc., which are studied to varying degrees as potential biomarkers for TNBC. They play roles in regulating gene expression, cell signaling, etc., and may be associated with tumor initiation, progression, and responsiveness to therapy. Although the aforementioned biomarkers have shown promise in theoretical and preliminary research, most are still in the experimental or early clinical validation stages. Transforming these biomarkers into clinical applications, such as developing diagnostic kits or incorporating them into routine clinical tests, requires addressing issues such as standardization, optimization of sensitivity and specificity, and large-scale validation. Additionally, the combined application of multiple biomarkers may be key to improving diagnostic accuracy and prognostic assessment. Future research should focus on multidimensional data integration and the development of comprehensive evaluation models for biomarkers.

The current treatment methods of TNBC include surgery, chemotherapy, and radiotherapy, among which systemic chemotherapy is the main treatment method (Li,<sup>16</sup> Bawaneh,<sup>17</sup> Manjunath,<sup>18</sup>. Chemotherapy regimens for TNBC mainly use anthracyclines, paclitaxel, platinum and cyclophosphamide, and there is no standard treatment regimen, and anthracyclines and/or paclitaxel are often ineffective after recurrence or metastasis (Dong,<sup>19</sup>. Because there are significant differences in survival prognosis among TNBC patients, patients have different sensitivity and efficacy to chemotherapy regimens. Patients with TNBC have poor treatment outcomes, low OS, and poor prognosis, so looking for new therapies to improve the prognosis of TNBC is extremely urgent.

With the development of next-generation gene sequencing technology and the continuous exploration of its molecular mechanism currently, immunotherapy and targeted therapy have gradually become the focus of researchers. In terms of targeted therapy, Trop-2 represents a pivotal emerging target in TNBC. Sacituzumab govitecan (Trodelvy) is an antibody–drug conjugate (ADC) targeting Trop-2, which delivers chemotherapy drugs directly to tumor cells by binding them with specific antibodies. It has demonstrated significant clinical benefits in the treatment of advanced TNBC, leading to improved patient survival rates. Furthermore, several studies are exploring the therapeutic potential of other targets such as PARP inhibitors for TNBC patients with BRCA mutations, as well as targeted drugs against the PI3K/Akt/mTOR signaling pathway, FGFR, and DDR pathways (Haque,<sup>20</sup>). In terms of immunotherapy, although TNBC exhibits high immunogenicity, not all patients respond to existing immune checkpoint inhibitors (such as PD-1/PD-L1 inhibitors). Clinical studies have shown that PD-L1-positive TNBC patients derive greater benefit from immunotherapy. For example, the use of atezolizumab in combination with chemotherapy has demonstrated prolonged progression-free survival (PFS) and OS in certain clinical trials (Hu,<sup>21</sup>). However, considering the heterogeneity of TNBC, a single biomarker such as PD-L1 is insufficient to comprehensively predict treatment response in all patients. Therefore, exploring combination therapies has become a focus of research, such as the combination of immune checkpoint inhibitors with anti-angiogenic agents, chemotherapy, or targeted therapies, aiming to enhance anti-tumor immune responses through multiple pathways.

The poor therapeutic effect and poor prognosis of TNBC have prompted people to urgently search for more effective biomarkers for prognosis prediction and immunotherapy intervention. From the perspective of immunotherapy, this work assisted clinical judgment of the prognosis of TNBC patients by screening immune-related biomarkers closely associated with the prognosis of TNBC patients and selected appropriate immunotherapy strategies according to this, so as to improve the prognosis of TNBC patients by reducing the recurrence rate and mortality and extending the survival period.

### Application status of bioinformatics analysis

Bioinformatics analysis is an analytical method that explores the laws of biological systems, including biological data retrieval (collection and screening), processing (editing, sorting, management, and display) and utilization (calculation and simulation). Thanks to the wide cross and comprehensive application of modern big data life science and bioinformation, bioinformatics analysis has been gradually applied in clinical diagnosis, accurate treatment, and precious prognosis assessment of various diseases, including different cancers. Pais et al.<sup>22</sup> performed clustering analysis on transcriptome data from 249 TNBC samples, identifying six distinct molecular subtypes, each with unique gene expression characteristics and potential treatment sensitivities (Pais,<sup>22</sup>). Using survival analysis and machine learning techniques, researchers constructed a multi-gene prognostic model (Chen,<sup>23</sup>). Pan et al.<sup>24</sup> integrated gene expression data and clinical information from the TCGA database, employing a LASSO Cox regression model to select a set of prognosis-related genes and developed a multi-gene risk scoring system, which effectively predicts the survival outcomes of TNBC patients (Pan,<sup>24</sup>).

Through the analysis of massive data, accurate evaluation of tumors can be realized to achieve breakthrough progress in the field of tumor disease monitoring, prevention, and clinical treatment (Xiao,<sup>25</sup>). The TNBC samples used in this study were all from the TCGA database, and a total of 60 tumor specimens were collected. Based on bioinformatics analysis for data processing, statistical analysis software is RStudio. The specific R packages used included: “Limma” package for the analysis and screening of the DEGs between different groups, and “survival” package for a survival analysis by combining the survival state with the time of impact. The most important feature is the combination of time factor. In addition, the “Survminer” package is the most used package in bioinformatics to implement survival analysis curve drawing. It also uses “ggplot” to draw beautiful and well-formatted survival curves. At the same time, it can also give P value, risk value and other parameters. “ROCR” package, ROC curve is a characteristic indicator reflecting sensitivity and continuity variables.

Compared with clinical case collection method, this analysis method can not only provide certain basis for research work, but also reduce the cost of scientific research and improve the efficiency of scientific research. Fang Jun et al. (Fang, 2021) found that elevated B7-H3 gene may be resulted from poor OS, especially in patients with lumen A and lumen B BC. Lin Jiaying et al. (Lin,<sup>26</sup> selected CCL5 gene as a prognostic biomarker through Kaplan–Meier analysis and Oncomine meta-analysis, and finally found that decreased expression of CCL5 reduced cell proliferation and invasion in clear cell renal cell cancer cell lines. Because the ability of single gene prediction based on bioinformatics analysis has certain limitations, the construction of multi-gene prognostic model to evaluate the relationship between each RNA and OS improves the reliability of prognostic prediction for tumor patients.

In this work, DEGs in TNBC samples were screened and IRG sets were obtained through online immune gene database. Immuno-related lncRNAs related to prognosis can be obtained after the intersection of the two. Further, a multi-gene risk prediction model can be constructed and the predictive ability of the model can be verified. Screening out meaningful immuno-related lncRNAs tightly linked to TNBC prognosis can further study the signaling pathway and molecular mechanism, and clarify the effect in tumors, prognosis, and other aspects. Therefore, it exhibited a certain guiding significance for immunotherapy strategy and prognosis evaluation of TNBC patients.

### Relationship between IRGs and tumors

Studies have shown that the interaction between tumor and the body’s immune system can affect the occurrence and progression of tumor, invasion, and metastasis, and the survival prognosis of tumor patients (Grabowski,<sup>27</sup>). The heterogeneity of tumor cells is caused by the immune microenvironment of tumor, and the proliferation, invasion ability, metastasis and drug resistance of tumor are related to the immune microenvironment of tumor

(Qian,<sup>28</sup> Different immune cell types express changing IRGs in the tumor immune microenvironment, so it is reasonable to speculate that IRGs are linked to the development and prognosis of tumors. Tumor cells evade host immune surveillance and attack by upregulating certain IRGs. For instance, the upregulation of PD-L1 can bind to PD-1 on the surface of T cells, leading to T cell exhaustion, which is a crucial mechanism of immune evasion. Blocking the function of such IRGs, such as using PD-1/PD-L1 inhibitors, can reinvigorate immune cells to combat tumors. IRGs influence the recruitment, activation, and polarization of immune cells in the tumor microenvironment. For example, members of the chemokine family such as CXCL9 and CXCL10 can promote the migration of T cells to the tumor area, enhancing anti-tumor immune responses. Conversely, some IRGs such as IL-10 and TGF- $\beta$  can induce the proliferation of regulatory T cells (Tregs), suppressing immune responses and favoring tumor growth. Inflammation plays a critical role in the initiation and progression of tumors, with numerous immune-related genes (IRGs) involved in regulating the inflammatory response. Under conditions of chronic inflammation, sustained activation of genes such as COX-2 and NF- $\kappa$ B can promote DNA damage, cell proliferation, and accelerate tumor progression. Additionally, inflammatory mediators can also influence the activity of immune cells, altering the immune balance within the tumor microenvironment. Qiu Pengjun et al. (Qiu,<sup>29</sup> established an exosome RS model and reported that the exosome related risk model was related to the prognosis and proportion of immune cell infiltration in TNBC patients. Xu Zhijie et al. (Xu,<sup>30</sup> based on nine lncRNA markers (CTD-2033A16.3, CTD-2116N20.1, CTD-2510F5.4, DDX11-AS1, LINC00942, LINC01224, LINC01231, LINC01508, and ZFPM2-AS1), constructed an iron death-related prognostic model for hepatocellular carcinoma, providing personalized prognostic tools for patients' prognosis and immune response.

It can be observed from literature review that correlation between IRGs and tumor prognosis is analyzed primarily based on bioinformatics methods. After screening out meaningful IRGs, prognostic prediction tools are used to verify their significance in tumor prognosis (Yang,<sup>31</sup> At present, it is necessary to understand the IRGs at the molecular mechanism level, which has certain research significance for the prognosis and prediction biomarkers of tumor therapy in the future. This work aimed to further refine the relationship between IRGs of TNBC in BC and prognosis of TNBC patients. Compared with other tumors, BC mostly belongs to non-immunogenic tumors (Torres,<sup>32</sup> and has a relatively low mutation rate. However, the high mutation rate of TNBC in BC subtype and the special tumor immune microenvironment provide a strong theoretical basis for the use of immunotherapy in TNBC. At the same time, it also proved that TNBC may have biomarkers with the potential to predict prognosis, which can be used to assist in the judgment of the patient's survival prognosis, offering a reference for the prognosis evaluation and improvement of TNBC.

Based on bioinformatics methods, the TNBC transcriptome dataset was downloaded from an online data database (TCGA). The IRGs associated with prognosis that have expression differences are screened out through data processing, and the relationship between IRGs and prognosis of TNBC is further analyzed and verified. It was proved that the immune genes are potential prognostic markers of TNBC, providing an individualized prognosis for the patient's survival prognosis. The results herein indicated that the screened IRGs supply a reference indicator for evaluating the survival prognosis of patients with TNBC, and may also provide a new therapeutic target for exploring the immunotherapy of TNBC.

### Relationship between lncRNA and tumor and its prognosis

Overexpression, deletion, and mutation of lncRNAs are key in tumor development, metastasis, and recurrence. Studies have shown that lncRNA importantly regulates the immune cell differentiation and development, autophagy, inflammation, cytokines, and other immune responses. It can be a potential biomarker for tumor immunotherapy and is expected to become a new target for tumor therapy, and is closely related to the survival prognosis of a variety of cancers (Park,<sup>33</sup> Therefore, it is of certain research significance to clarify the role of lncRNA in the occurrence, progression, and prognosis of tumors in terms of immunity. Certain lncRNAs can function as transcriptional regulatory factors, directly or indirectly affecting the expression of cell cycle proteins and apoptosis-related genes. For example, HOX transcript antisense RNA (HOTAIR) is overexpressed in various cancers, including TNBC, where it interacts with Polycomb repressive complexes to suppress the expression of tumor suppressor genes and promote tumor growth. lncRNAs can enhance tumor invasiveness and metastatic capacity by regulating cell adhesion molecules, matrix metalloproteinases (MMPs), and the epithelial-mesenchymal transition (EMT) process. MALAT1 (Metastasis-Associated Lung Adenocarcinoma Transcript 1) is highly expressed in various cancers, including TNBC, where it promotes the EMT process and increases the migration and invasion capabilities of tumor cells. This work explored from the perspective of immune regulation in TNBC patients. Currently, there are few studies on lncRNAs related to TNBC immunization, which cannot provide sufficient theoretical and experimental basis for clinical application. It is necessary to screen TNBC immune-related lncRNAs, and find those with prognostic evaluation significance as biomarkers by analyzing their relationship with TNBC prognosis.

In this work, three lncRNAs were screened out to be associated with TNBC immunity, namely LINC01235, AC090181.2, and LINC01943, and the results also showed that the PR model constructed by the three lncRNAs. LINC01235 is a newly discovered lncRNA. It is involved in the occurrence of tumors and is expressed in 12 kinds of human tissues, such as stomach, mammary gland, and bladder, on the 9p23 chromosome of human body. Zhang Cheng et al. (Zhang<sup>34</sup> confirmed that LINC01235 was observably overexpressed in GC cells and tissues by qRT-PCR. The study of Li Zhe et al. (Li<sup>35</sup> showed that Linc01235-ESR1 and Linc01235-Adra2a can be an important co-expression pair during progression of BC, and LINC01235 is a critical independent prognostic factor. Further literature search and reference in the database showed that there were no relevant studies and reports on AC090181.2 and LINC01943 in the field of cancer. Therefore, further study is necessary to explore the molecular mechanism and regulated signal transduction pathway of them to clarify their mechanism in

prognosis of TNBC patients, thus giving a reference for evaluating the survival prognosis of patients with TNBC and prognosis intervention.

## Conclusion

In this work, TNBC immune-related lncRNAs were screened out, and a PR model was established for exploring the correlation between immune-related lncRNAs and prognosis. The results told that using it as an independent prognostic biomarker of TNBC can effectively evaluate the survival prognosis of patients with TNBC, giving a reference indicator for clinical evaluation of the therapeutic efficacy of patients with TNBC. Among the 3 lncRNAs in PR model, LNC01235 has been confirmed to be associated with the prognosis of patients with gastric cancer and BC, but has not been reported in TNBC. AC090181.2 and LINC01943 have not been found in the field of cancer, and their functions are still unclear. They can be used as the newly discovered prognostic biomarkers of TNBC for further study. All data were from the TCGA database, and the sample size was also small. The influence of regional, population, and other factors on lncRNA should be further explored by expanding the selection range of patients.

## Data availability

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

Received: 13 December 2023; Accepted: 21 June 2024

Published online: 15 August 2024

## References

- Zhang, L., Chen, W., Liu, S. & Chen, C. Targeting breast cancer stem cells. *Int. J. Biol. Sci.* **19**(2), 552–570. <https://doi.org/10.7150/ijbs.76187> (2023).
- Baranovicova, E. *et al.* Circulating metabolites in the early stage of breast cancer were not related to cancer stage or subtypes but associated with ki67 level. Promising statistical discrimination from controls. *Mol. Cell. Probes.* **66**, 101862. <https://doi.org/10.1016/j.mcp.2022.101862> (2022).
- Howard, F. M. & Olopade, O. I. Epidemiology of triple-negative breast cancer: A review. *Cancer J.* **27**(1), 8–16. <https://doi.org/10.1097/PP0.0000000000000500> (2021).
- Yang, F. *et al.* Ferroptosis heterogeneity in triple-negative breast cancer reveals an innovative immunotherapy combination strategy. *Cell. Metab.* **35**(1), 84–100.e8. <https://doi.org/10.1016/j.cmet.2022.09.021> (2023).
- Lehmann, B. D. *et al.* Multi-omics analysis identifies therapeutic vulnerabilities in triple-negative breast cancer subtypes. *Nat. Commun.* **12**(1), 6276. <https://doi.org/10.1038/s41467-021-26502-6> (2021).
- Wang, X. Q. *et al.* Spatial predictors of immunotherapy response in triple-negative breast cancer. *Nature.* **621**(7980), 868–876. <https://doi.org/10.1038/s41586-023-06498-3> (2023).
- Bridges, M. C., Daulagala, A. C. & Kourtidis, A. LNCation: lncRNA localization and function. *J. Cell. Biol.* **220**(2), e202009045. <https://doi.org/10.1083/jcb.202009045> (2021).
- Zhu, S. *et al.* Recent advances in targeted strategies for triple-negative breast cancer. *J. Hematol. Oncol.* **16**(1), 100. <https://doi.org/10.1186/s13045-023-01497-3> (2023).
- Wigton, E. J. & Ansel, K. M. Noncoding RNAs in B cell responses. *RNA Biol.* **18**(5), 633–639. <https://doi.org/10.1080/15476286.2021.1885876> (2021).
- Li, L., Zhang, F., Liu, Z. & Fan, Z. Immunotherapy for triple-negative breast cancer: Combination strategies to improve outcome. *Cancers (Basel).* **15**(1), 321. <https://doi.org/10.3390/cancers15010321> (2023).
- Zhang, J. *et al.* lncRNA LINC00649 promotes the growth and metastasis of triple-negative breast cancer by maintaining the stability of HIF-1 $\alpha$  through the NF90/NF45 complex. *Cell. Cycle.* **21**(10), 1034–1047. <https://doi.org/10.1080/15384101.2022.2040283> (2022) (Epub 2022 Feb 21).
- Xiu, Y., Cao, S., Jiang, R. & Zhou, Y. lncRNA LINC01315 promotes malignancy of triple-negative breast cancer and predicts poor outcomes by modulating microRNA-876-5p/GRK5. *Bioengineered* **13**(4), 10001–10009. <https://doi.org/10.1080/21655979.2022.2062536> (2022).
- Vagia, E., Mahalingam, D. & Cristofanilli, M. The landscape of targeted therapies in TNBC. *Cancers (Basel).* **12**(4), 916. <https://doi.org/10.3390/cancers12040916> (2020).
- Qin, G. *et al.* NPM1 upregulates the transcription of PD-L1 and suppresses T cell activity in triple-negative breast cancer. *Nat. Commun.* **11**(1), 1669. <https://doi.org/10.1038/s41467-020-15364-z> (2020).
- Kwapisz, D. Pembrolizumab and atezolizumab in triple-negative breast cancer. *Cancer Immunol. Immunother.* **70**(3), 607–617. <https://doi.org/10.1007/s00262-020-02736-z> (2021).
- Li, Y. *et al.* Recent advances in therapeutic strategies for triple-negative breast cancer. *J. Hematol. Oncol.* **15**(1), 121. <https://doi.org/10.1186/s13045-022-01341-0> (2022).
- Bawaneh, A. *et al.* Intestinal microbiota influence doxorubicin responsiveness in triple-negative breast cancer. *Cancers (Basel)* **14**(19), 4849. <https://doi.org/10.3390/cancers14194849> (2022).
- Manjunath, M. & Choudhary, B. Triple-negative breast cancer: A run-through of features, classification and current therapies. *Oncol. Lett.* **22**(1), 512. <https://doi.org/10.3892/ol.2021.12773> (2021).
- Dong, S. *et al.* Ceritinib is a novel triple negative breast cancer therapeutic agent. *Mol. Cancer.* **21**(1), 138. <https://doi.org/10.1186/s12943-022-01601-0> (2022).
- Haque, S., Cook, K., Sahay, G. & Sun, C. RNA-Based therapeutics: Current developments in targeted molecular therapy of triple-negative breast cancer. *Pharmaceutics* **13**(10), 1694. <https://doi.org/10.3390/pharmaceutics13101694> (2021).
- Hu, S., Qu, X., Jiao, Y., Hu, J. & Wang, B. Immune classification and immune landscape analysis of triple-negative breast cancer. *Front. Genet.* **2**(12), 710534. <https://doi.org/10.3389/fgene.2021.710534> (2021).
- Pais, R. J., Iles, R. K. & Zmuidinaite, R. MALDI-ToF mass spectra phenomic analysis for human disease diagnosis enabled by cutting-edge data processing pipelines and bioinformatics tools. *Curr. Med. Chem.* **28**(32), 6532–6547. <https://doi.org/10.2174/0929867327666201027154257> (2021).
- Chen, W. *et al.* Study on the potential active components and molecular mechanism of Xiao Huoluo Pills in the treatment of cartilage degeneration of knee osteoarthritis based on bioinformatics analysis and molecular docking technology. *J. Orthop. Surg. Res.* **16**(1), 460. <https://doi.org/10.1186/s13018-021-02552-w> (2021).

24. Pan, J., Zhang, X., Fang, X. & Xin, Z. Construction on of a ferroptosis-related lncRNA-based model to improve the prognostic evaluation of gastric cancer patients based on bioinformatics. *Front. Genet.* **23**(12), 739470. <https://doi.org/10.3389/fgene.2021.739470> (2021).
25. Xiao, C., Wang, F., Jia, T., Pan, L. & Wang, Z. Big data analysis and application of liver cancer gene sequence based on second-generation sequencing technology. *Comput. Math. Methods Med.* **16**(2022), 4004130. <https://doi.org/10.1155/2022/4004130> (2022).
26. Lin, J. *et al.* Identification of biomarkers related to CD8+ T cell infiltration with gene co-expression network in clear cell renal cell carcinoma. *Aging (Albany NY)*. **12**(4), 3694–3712. <https://doi.org/10.18632/aging.102841> (2020).
27. Grabowski, M. M. *et al.* Immune suppression in gliomas. *J. Neurooncol.* **151**(1), 3–12. <https://doi.org/10.1007/s11060-020-03483-y> (2021).
28. Qian, Y. *et al.* Single-cell RNA-seq dissecting heterogeneity of tumor cells and comprehensive dynamics in tumor microenvironment during lymph nodes metastasis in gastric cancer. *Int. J. Cancer.* **151**(8), 1367–1381. <https://doi.org/10.1002/ijc.34172> (2022).
29. Qiu, P., Guo, Q., Yao, Q., Chen, J. & Lin, J. Characterization of exosome-related gene risk model to evaluate the tumor immune microenvironment and predict prognosis in triple-negative breast cancer. *Front. Immunol.* **1**(12), 736030. <https://doi.org/10.3389/fimmu.2021.736030> (2021).
30. Xu, Z. *et al.* Construction of a ferroptosis-related nine-lncRNA signature for predicting prognosis and immune response in hepatocellular carcinoma. *Front. Immunol.* **17**(12), 719175. <https://doi.org/10.3389/fimmu.2021.719175> (2021).
31. Yang, C., Huang, S., Cao, F. & Zheng, Y. A lipid metabolism-related genes prognosis biomarker associated with the tumor immune microenvironment in colorectal carcinoma. *BMC Cancer.* **21**(1), 1182. <https://doi.org/10.1186/s12885-021-08902-5> (2021).
32. Torres, E. T. R. & Emens, L. A. Emerging combination immunotherapy strategies for breast cancer: Dual immune checkpoint modulation, antibody-drug conjugates and bispecific antibodies. *Breast Cancer Res. Treat.* **191**(2), 291–302. <https://doi.org/10.1007/s10549-021-06423-0> (2022).
33. Park, E. G., Pyo, S. J., Cui, Y., Yoon, S. H. & Nam, J. W. Tumor immune microenvironment lncRNAs. *Brief Bioinform.* <https://doi.org/10.1093/bib/bbab504> (2022).
34. Zhang, C. *et al.* The novel role and function of LINC01235 in metastasis of gastric cancer cells by inducing epithelial-mesenchymal transition. *Genomics* **113**(3), 1504–1513. <https://doi.org/10.1016/j.ygeno.2021.03.027> (2021).
35. Li, Z. *et al.* Identification of key lncRNA-mRNA pairs and functional lncRNAs in breast cancer by integrative analysis of TCGA data. *Front Genet.* **20**(12), 709514. <https://doi.org/10.3389/fgene.2021.709514> (2021).

### Author contributions

SY and QW contributed to conception and design of the study. SY organized the database. QW performed the statistical analysis. SY wrote the first draft of the manuscript. QW wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

### Funding

This article was funded by the Medical Science and Technology, Project of Zhejiang Province (2022KY1299).

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to Q.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024, corrected publication 2024