



# OPEN SERPINA1 gene regulates the tumorigenesis and progression of breast cancer through PI3K/AKT signaling pathway and tumor immune microenvironment

Chuqi Lei<sup>1,2,5</sup>, Wei Lu<sup>1,5</sup>, Yuan Li<sup>3</sup>, Huaiyu yang<sup>1</sup>, Ke Zhang<sup>1</sup>, Nianchang Wang<sup>1</sup>, Lixue Xuan<sup>1</sup>✉ & Changyuan Guo<sup>4</sup>✉

This study aims to reveal the influence of SERPINA1 gene on the development, prognosis evaluation and immune environment changes of breast cancer. Cell lines with differential SERPINA1 expression were constructed. These models were used to explore its impact on the biological behavior of breast cancer cells. The expression of SERPINA1 in breast tumor tissues of patients undergoing surgery in our hospital was detected by immunohistochemistry to evaluate the regulatory effect of SERPINA1 gene on tumor microenvironment. The results found that the overexpression of SERPINA1 gene could significantly inhibit the proliferation and migration of breast cancer cells and promote apoptosis. Transcriptome sequencing analysis revealed that SERPINA1 may regulate the biological behavior of cells by affecting biological functional pathways such as adaptive immune response, natural killer cell-mediated cytotoxicity, and phosphatidylinositol signaling system<sup>1</sup>. Western blot analysis showed that SERPINA1 overexpression was accompanied by an increase in PTEN expression and a decrease in Akt and mTOR phosphorylation levels, suggesting a molecular mechanism by which SERPINA1 gene and PTEN cooperate to negatively regulate the passage of PI3K. In addition, the high expression of SERPINA1 is related to the recruitment of regulatory T cells (Treg) in breast cancer tissues, which may change the tumor microenvironment.

**Keywords** Breast cancer, SERPINA1 gene, PI3K/AKT, Tumor immune microenvironment

Breast cancer is a rising global health problem, affecting women of all ages in different countries<sup>2</sup>. The female mortality rate of breast cancer is 15.4%, which is the disease with the largest number of cancer-related deaths among women<sup>3</sup>. Breast cancer shows diversity at the biological level, and the prognosis is affected by a series of factors, such as different molecular subtypes, age of disease diagnosis, histopathological characteristics and so on<sup>4</sup>. In recent years, the increasing efforts of early screening for breast cancer and the continuous development of comprehensive treatment methods have significantly improved the prognosis of breast cancer patients<sup>5,6</sup>, but due to the substantial heterogeneity of the disease and the intrinsic or acquired drug resistance, the current clinical treatment still faces major challenges.

Serpin peptidase inhibitor clade A member 1 (SERPINA1) gene is a key member of coagulation related pathways. The  $\alpha$ 1-antitrypsin inhibitor (A1AT) encoded by this gene is a highly conserved protein in organisms<sup>7</sup>. It specifically inhibits a variety of proteolytic enzymes, thus playing a protective role in the circulatory system and tissues and avoiding the destruction of tissues by proteolytic enzymes<sup>8</sup>. Previous studies have shown that

<sup>1</sup>Department of Breast Surgical Oncology, National Cancer Center, National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences, Peking Union Medical College, Beijing 100021, China.

<sup>2</sup>Department of Breast Surgery, Beijing Chao-Yang Hospital, Capital Medical University, Beijing 100020, China.

<sup>3</sup>Department of Breast Surgical Oncology, National Cancer Center, National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen 518116, China. <sup>4</sup>Department of Pathology, National Clinical Research Center for Cancer/Cancer Hospital, National Cancer Center, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China. <sup>5</sup>Chuqi Lei and Wei Lu contributed equally to this work. ✉email: xuanlixuepumc@126.com; guocypumc@163.com

SERPINA1 gene expression patterns in a variety of solid tumors have significant changes and play a role in the key stages of tumor development<sup>9</sup>. The study of Evrim Ercetin pointed out that in patients with non-small cell lung cancer (NSCLC), the expression level of SERPINA1 in tumor tissue was low, while the expression level was high in non-tumor lung tissue<sup>10</sup>. The study of Lei Li et al. found that the expression level of SERPINA1 in colorectal cancer patients was significantly higher than that in healthy controls, and was closely related to the TNM stage of the tumor<sup>11</sup>. In thyroid cancer, SERPINA1 may serve as a biomarker of poor prognosis and promote tumor progression<sup>12</sup>. For breast cancer, Previous studies used bioinformatics methods to identify the prognostic significance of six coagulation related genes, including SERPINA1, in breast cancer<sup>9</sup>. SERPINA1 may play a tumor suppressive role in breast cancer, and its low expression may promote tumor development and metastasis<sup>13,14</sup>. These findings indicate that the tumorigenic effect of SERPINA1 is not unchanging, and its biological function and clinical significance in different tumor types need further study and elucidation.

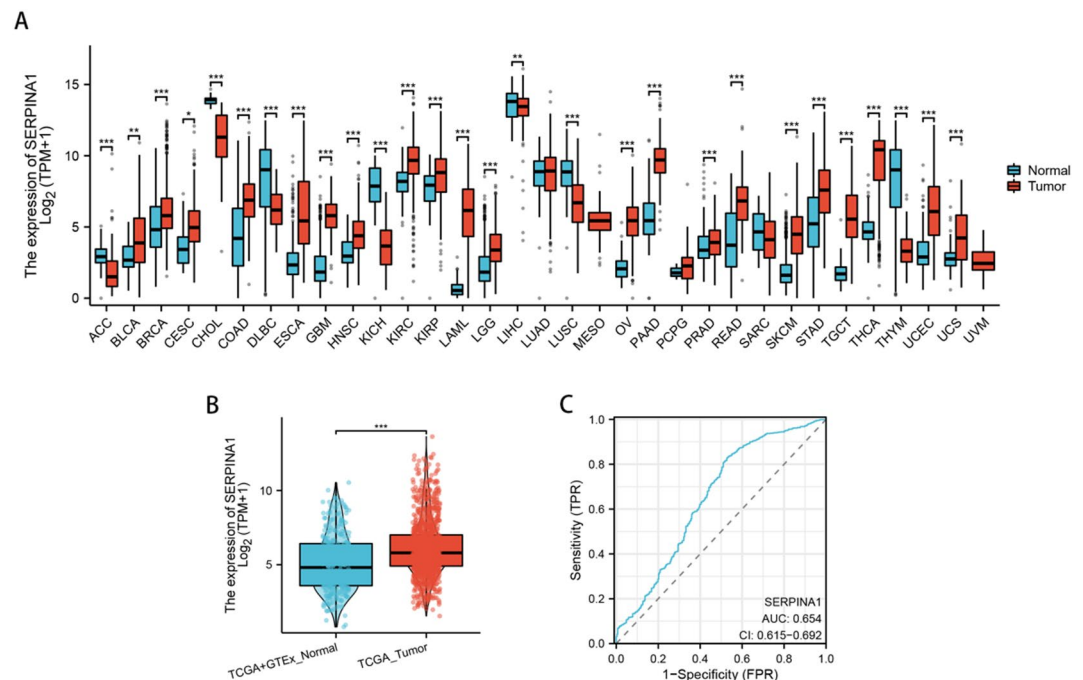
SERPINA1 protein (A1AT) can inhibit the release of pro-inflammatory cytokines in monocytes through NF $\kappa$ B signaling pathway<sup>15</sup>, and has the ability to activate leukocytes and promote inflammatory response<sup>16</sup>. In addition, SERPINA1 protein (A1AT) has also been observed to contribute to the generation of regulatory T cells (Treg), reduce the infiltration of lymphocytes, and inhibit the differentiation and maturation of dendritic cells, thus playing a role in the tumor microenvironment<sup>17–19</sup>. Tumor cells destroy surrounding tissues by releasing enzymes such as elastase, plasmin and cathepsin, and then achieve local diffusion. SERPINA1 protein (A1AT) affects the invasiveness and diffusion ability of tumors by neutralizing the activities of these enzymes. These research results not only show the great potential of SERPINA1 in regulating the tumor microenvironment (TME), but also emphasize its important value as a biomarker in guiding the development of immunotherapy strategies. Therefore, the in-depth study of SERPINA1 may provide new perspectives and therapeutic targets for tumor therapy, especially in the field of immunotherapy.

In this study, we first analyzed the role of SERPINA1 gene in the prognosis and immune regulation of breast cancer by using bioinformatics methods, and explored the differential expression of SERPINA1 and its prognostic value. Subsequently, the SERPINA1 gene was overexpressed and knocked down in breast cancer cell lines, and the function, signaling pathway and its role in the occurrence and development of breast cancer were explored by epigenetic experimental techniques such as proliferation, migration and apoptosis. In addition, this study used immunohistochemical technology to stain SERPINA1 gene and Treg lymphocyte count on the pathological sections of postoperative patients to explore the immune regulation role of this gene in breast cancer.

## Result

### Differential expression of SERPINA1 in tumor tissues and normal tissues and its diagnostic value

According to the results of Wilcoxon rank sum test, SERPINA1 expression was different in 33 cancer species (Fig. 1A). In most cancer species, SERPINA1 expression is elevated in tumor tissues, while in a few cancer species, such as ACC, DLBC, KICH and LUSC, SERPINA1 expression is elevated in normal tissues, and decreased in



**Fig. 1.** A. Differential expression of SERPINA1 in 33 different cancer species; B. SERPINA1 expression in breast cancer is higher than that in normal tissues; C. SERPINA1 is used to differentiate the diagnostic ROC curve of breast cancer tumor tissues and normal tissues. (\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ).

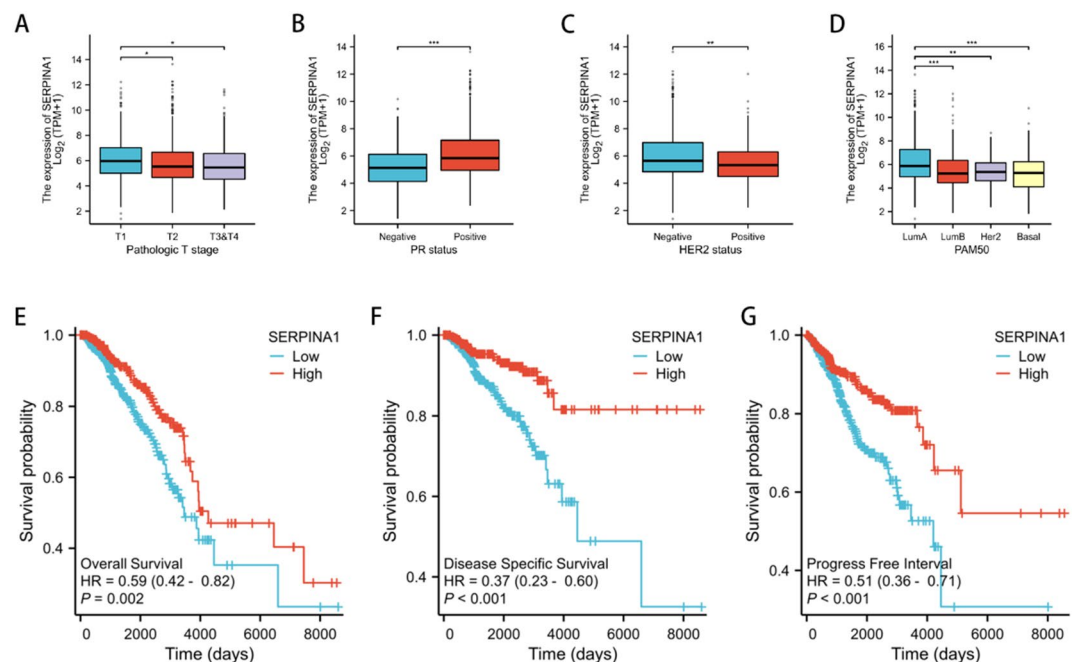
tumor tissues. Like most cancer species, SERPINA1 expression in breast cancer was increased in tumor tissues (Fig. 1B) ( $p < 0.001$ ). The diagnostic value of SERPINA1 for differentiating breast cancer tumor tissues from normal tissues was detected by receiver operating characteristic curve (ROC), and the AUC value obtained was 0.654 (95%CI:0.615–0.692,  $p > 0.05$ ), indicating the potential value of SERPINA1 as a diagnostic marker of breast cancer (Fig. 1C).

### SERPINA1 expression is correlated with clinicopathological characteristics and prognosis of breast cancer patients

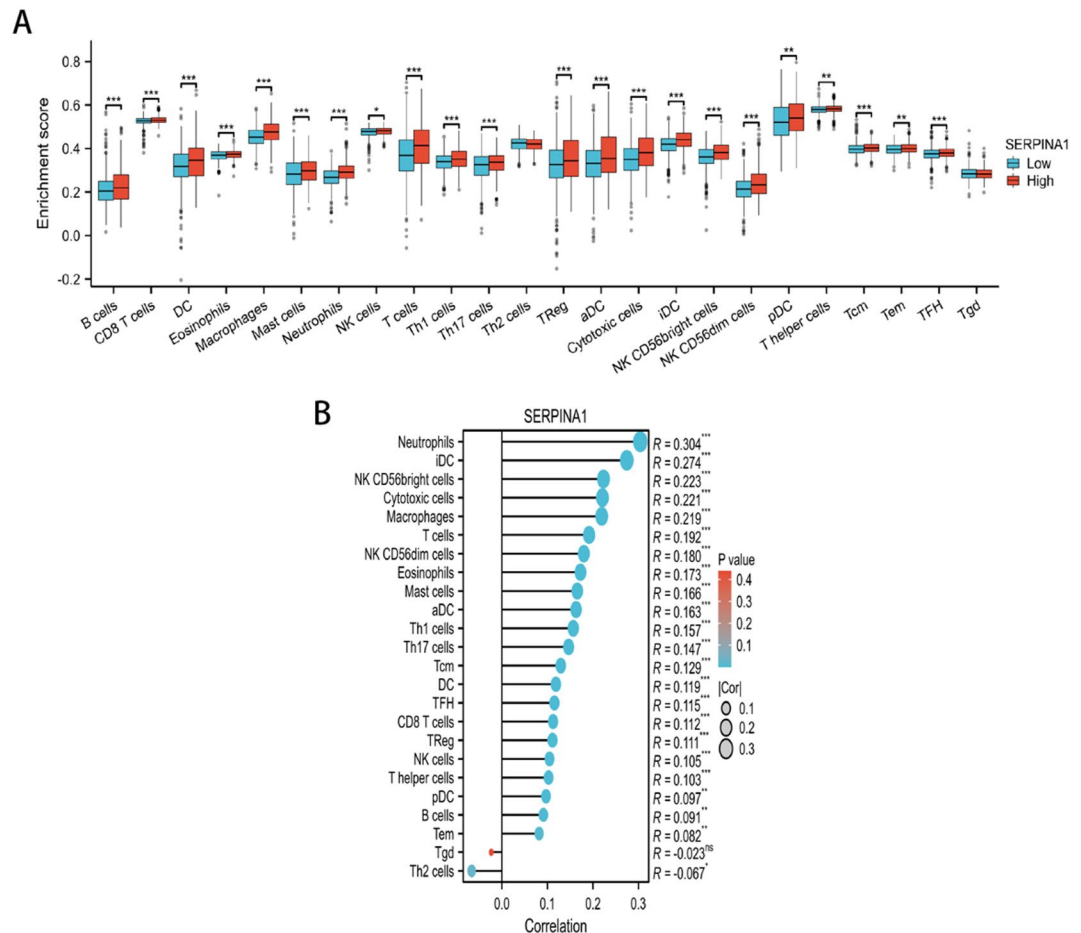
There were 543 patients in the low expression group and 544 patients in the high expression group. The analysis between groups found that there were no significant differences in age, N stage, M stage, clinical stage, pathological type, ER status, menopausal status and other characteristics between the low expression group and the high expression group, but there were significant differences in T stage ( $p = 0.001$ ), PR status ( $p < 0.001$ ), HER2 status ( $p = 0.035$ ), and molecular typing ( $p < 0.001$ ). The number of patients with high expression of SERPINA1 in T1 was significantly more than that of patients with low expression, while the number of patients with low expression in T2, T3 and T4 was more. Patients with high expression were more likely to be PR positive and HER2 negative, while patients with low expression were more likely to be PR negative and HER2 positive. In molecular typing, the expression of SERPINA1 in patients with Lum A type was significantly higher than that in patients with Lum B type, HER2 positive type and triple negative type (Fig. 2A–D). Further exploring the prognosis between the high and low groups, the results showed that there were significant differences in overall survival (OS) ( $p = 0.008$ ), disease-specific survival (DSS) ( $p < 0.001$ ) and progress free interval (PFI) ( $p < 0.001$ ) between the two groups. According to the Kaplan Meier curve, the OS, DSS and PFI of the patients in the low expression group were significantly lower than those in the high expression group, suggesting that the patients in the low expression group had a worse prognosis (Fig. 2E–G), which was consistent with the prognostic direction of the clinical characteristics of patients in the high and low groups (Supplementary Table 1).

### There were differences in immune infiltration between patients with high and low SERPINA1 expression groups

The results showed that except Th2 cells and TGD cells, the infiltration of immune cells in the other 22 groups between the high and low expression groups were statistically different, and the enrichment index of immune cell infiltration in the high expression group was higher than that in the low expression group, suggesting that patients in the high expression group had more abundant immune cell infiltration (Fig. 3A). Spearman correlation coefficient analysis showed that the immune cells positively correlated with SERPINA1 expression were neutrophils, interdigitating dendritic cells (IDC), cd56bright natural killer cells (NK cd56bright cells), cytotoxic cells, macrophages, etc. in addition, the infiltration of Th2 cells was negatively correlated with SERPINA1 expression level (Fig. 3B).



**Fig. 2.** A. There were differences in SERPINA1 expression in patients with different T stages; B. There were differences in SERPINA1 expression in patients with different PR status; C. There were differences in SERPINA1 expression in patients with different ER status; D. There were differences in SERPINA1 expression in patients with different molecular subtypes; E. Kaplan Meier curve with OS as endpoint; F. Kaplan Meier curve with DSS as endpoint; G. Kaplan Meier curves with PFI as endpoint. (\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ).



**Fig. 3.** A. Infiltration of 24 immune cells in high and low expression groups; B. Spearman correlation analysis results of SERPINA1 expression and immune cells. (\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ ).

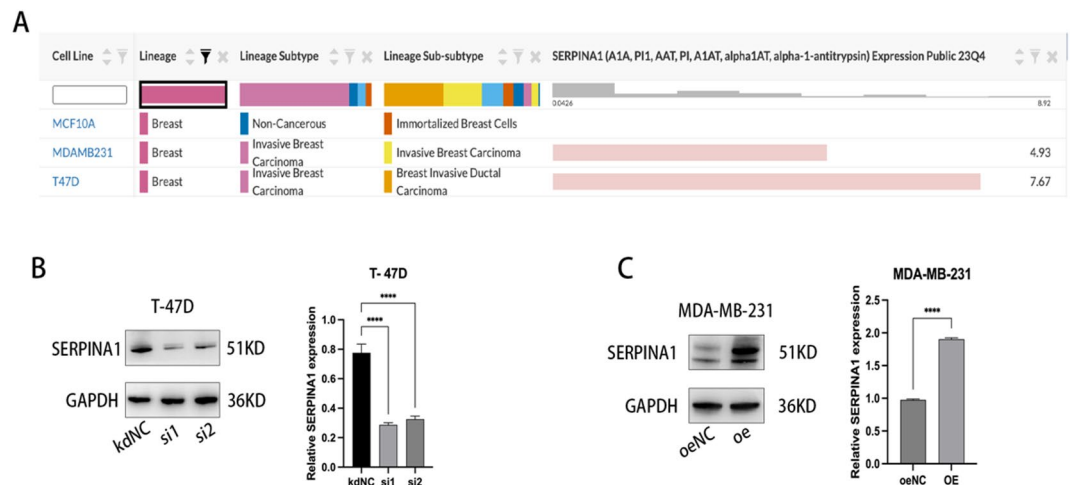
### Construction of breast cancer cell lines with overexpression of SERPINA1 and low expression of SERPINA1

According to the expression distribution of SERPINA1 gene in different breast cancer cell lines in CCLE database<sup>20</sup>, it can be seen that there is no expression of SERPINA1 in MCF-10 A, a normal breast ductal epithelial cell, but there is low expression and high expression in MDA-MB-231 and T-47D, respectively (Fig. 4A). Therefore, in this study, MDA-MB-231 breast invasive cancer cell line was used for overexpression of SERPINA1 gene, and t 47D breast invasive cancer cell line was used for knockdown of SERPINA1. The two sequences with the highest transfection efficiency [si1:5' gccagcttaccattacccaaa3'], [si2:5' ccagccaacacagcacaata3'] were used to knock down SERPINA1 gene in T-47D cell line. Protein samples were collected at the 48th hour of transfection experiment for Western blot analysis (Fig. 4B). Quantitative analysis of bands showed that both SI1 and SI2 sequences could significantly reduce the expression of SERPINA1 ( $p < 0.001$ ), with knockdown efficiencies of 78% and 69%, respectively. The SERPINA1 overexpression lentivirus was used to transfect MDA-MB-231 cell line, and the protein samples were collected for analysis. It can be seen that the expression of SERPINA1 in the expression group was significantly higher than that in the **vector-only** control group (Fig. 4C).

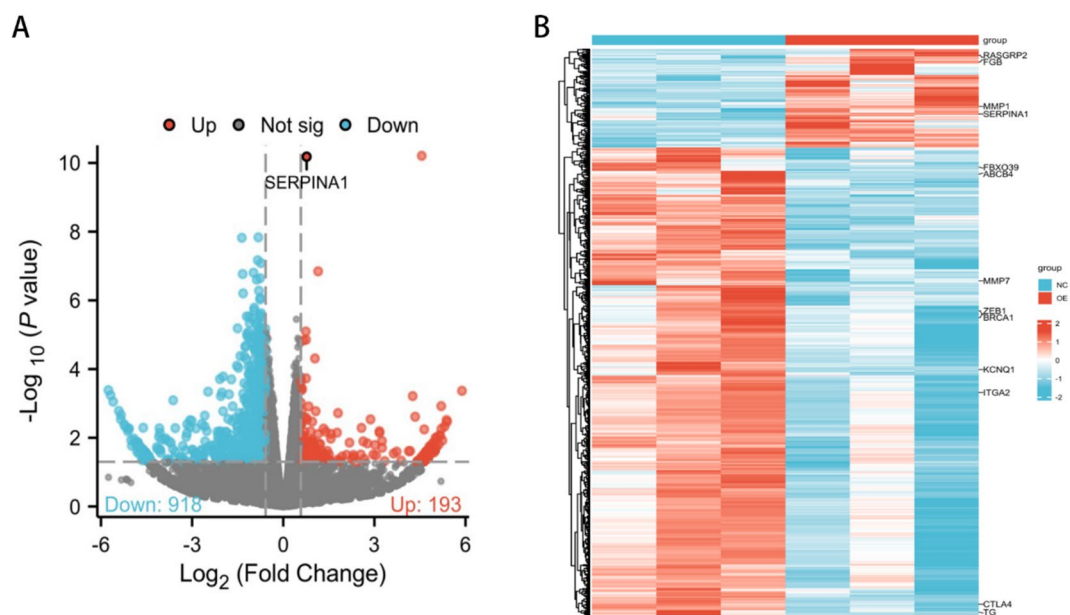
### High throughput sequencing results of overexpression group cells

Differential gene analysis of the sequencing results showed that the  $\log_{2}FC = 0.77$ ,  $p < 0.05$ , indicating that the expression of SERPINA1 in the overexpression group increased by about 2.15 times compared with the control group. Using  $|\log_{2}FC| > 0.58$  and  $p < 0.05$  as the standard to screen differentially expressed genes, a total of 1111 differential genes were obtained. The volcano map and heat map of differential genes are shown in Fig. 5. Go and KEGG functional enrichment analysis were performed on the differential genes (see Figs. 6 and 7).

According to the statistics and differential analysis of alternative splicing events in the transcriptome. The results are shown in Fig. 8. The highest type of differential alternative splicing events was se, with 1958 cases of differential events detected, followed by 503 Cases of MXE, followed by 273, 195 and 63 cases of a3s3, a5ss and RI, respectively. Go and KEGG functional enrichment analysis was performed for the differential splicing events. Fig. 9 and Fig. 10 show the p-value ranking (Supplementary Table 2).



**Fig. 4.** A. Expression distribution of SERPINA1 gene in different breast cancer cell lines in CCLE database; **B.** The knockdown of SERPINA1 gene was detected by Western blot analysis; **C.** The overexpression of SERPINA1 gene was detected by Western blot analysis. (\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ , original blots/gels are presented in Supplementary Figure).



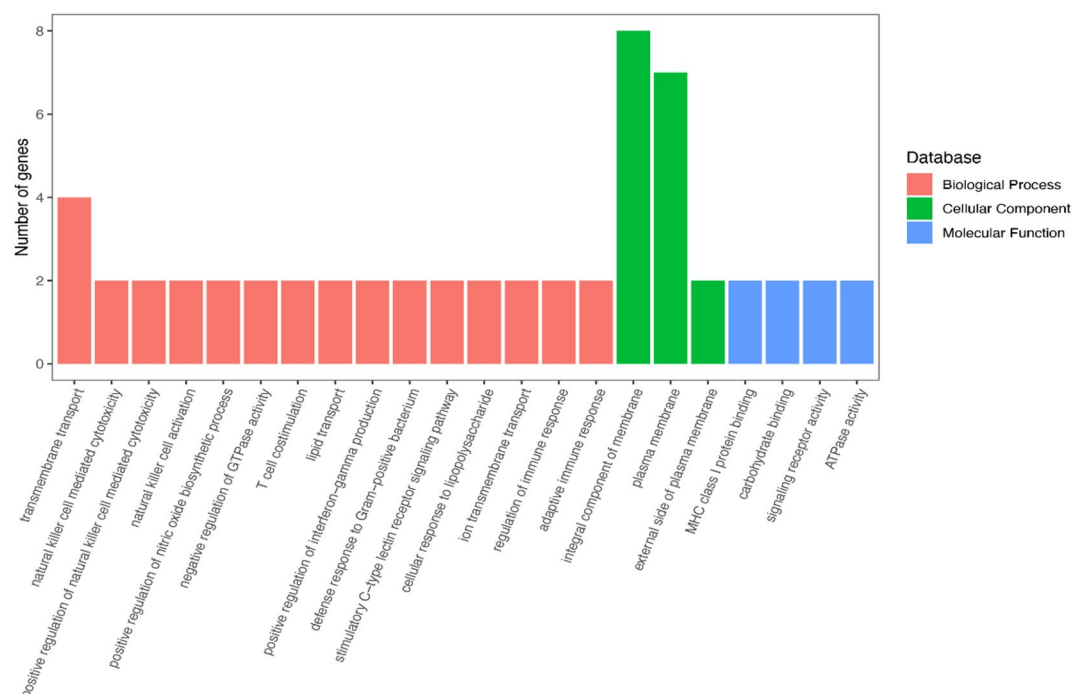
**Fig. 5.** Volcano plot and heat map of differentially expressed genes by transcriptome high-throughput sequencing.

### SERPINA1 overexpression inhibits breast cancer cell proliferation

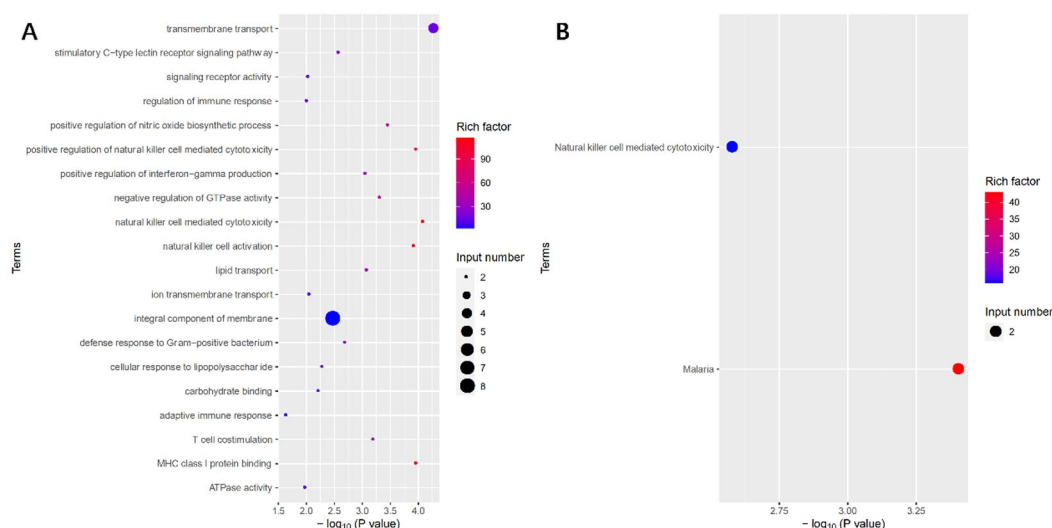
Real-time cell growth monitoring was used in this study RTCA xCELL system was used to monitor and record the cell proliferation process. The total observation and recording time of this experiment is 108 h, and the measurement is carried out once every 5 min for 1 min each time. The measured value is cell index, and the cell proliferation curve is drawn using the cell index value. The abscissa of the curve is time, and the ordinate is cell index. The curve drawn is shown in Fig. 11A & B. it can be seen in Fig. 10A that the proliferation of cells in the expression group is slower than that in the control group ( $p < 0.001$ ). As shown in Fig. 11B, the proliferation rate of knockdown groups in both sequences is higher than that in the control group ( $p < 0.05$ ).

As shown in Fig. 11C & D, the results of plate clonogenic assay showed that the number of clones formed in the knockdown group was significantly higher than that in the control group ( $p < 0.001$ ), while the number of clones in the overexpression group was significantly lower than that in the control group ( $p < 0.001$ ).





**Fig. 6.** Go enrichment analysis results of differentially expressed genes by transcriptome sequencing.



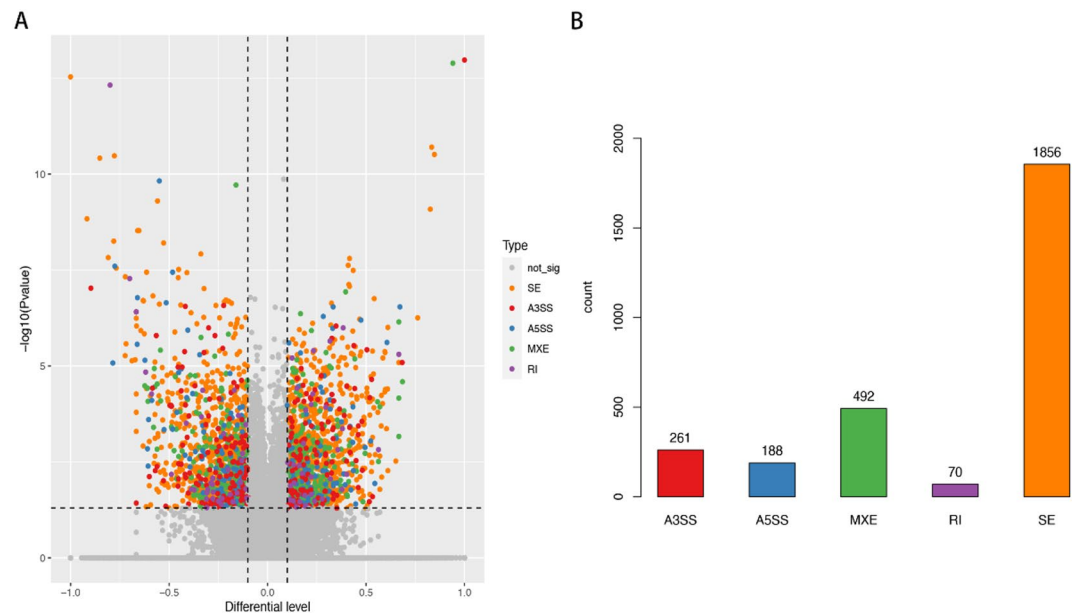
**Fig. 7.** Go and KEGG pathway enrichment results of differentially expressed genes by transcriptome sequencing. **A.** Go pathway enrichment results; **B.** KEGG pathway enrichment results<sup>21–23</sup>.

## Overexpression of SERPINA1 inhibits breast cancer migration

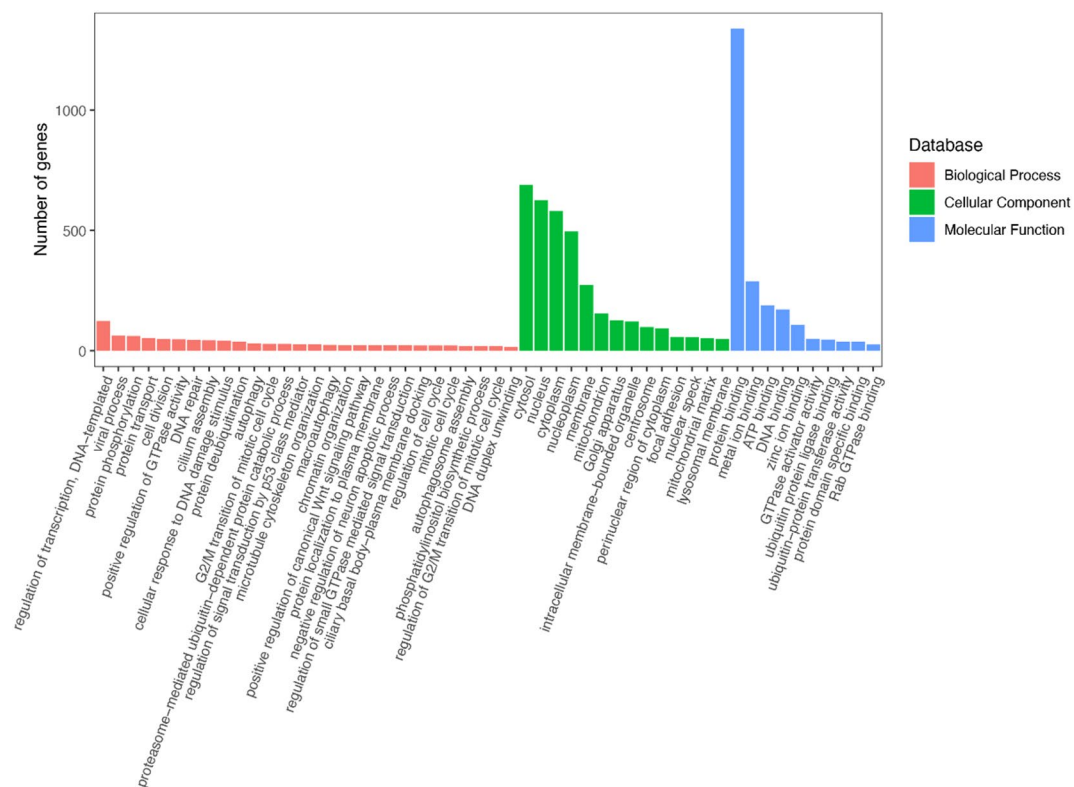
The results showed that the distance of cell migration in the overexpression group was significantly lower than that in the control group ( $p < 0.01$ ); In the cell lines in the knockdown group, it can be seen that the migration speed and ability of the cell lines after knockdown treatment are significantly improved compared with those in the control group, and the migration ability of the cell lines after SI1 sequence knockdown is more obvious (Fig. 12).

## SERPINA1 overexpression promotes apoptosis in breast cancer cell lines

It can be seen from the experiment (Fig. 13) that the expression level of Caspase-3 in the cell line treated with apoptosis induction is up-regulated, so as to increase the efficiency of the apoptotic signaling pathway. The results of this experiment showed that the expression of total caspase-3 in the knockdown group treated with apoptosis induction was higher than that in the control group, the expression of cleaved-caspase-3 (17kd) was

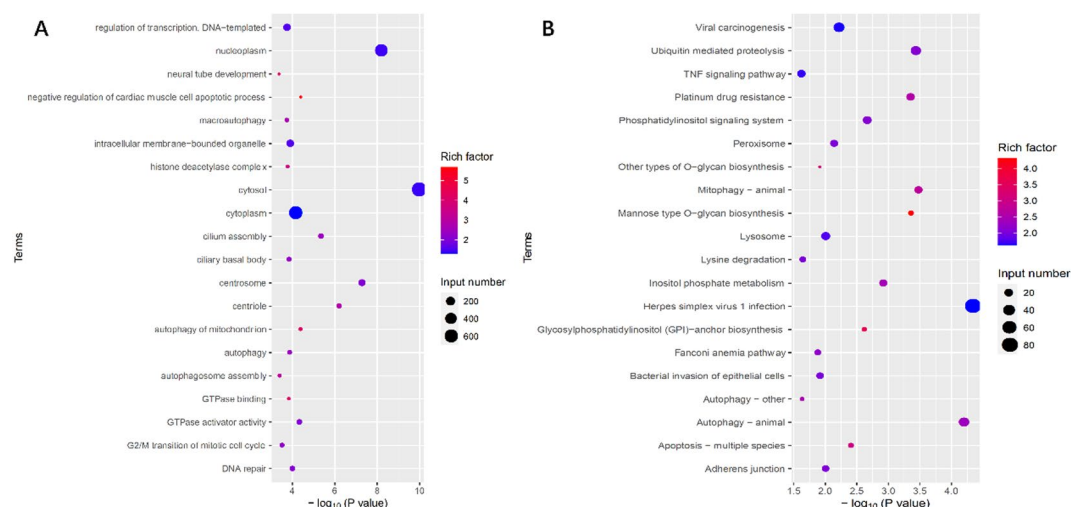


**Fig. 8.** Differential analysis of alternative splicing events in transcriptome sequencing of overexpressed and control cell lines **A**. Volcano plot of differential alternative splicing events; **B**. Number of different types of differential alternative splicing events.

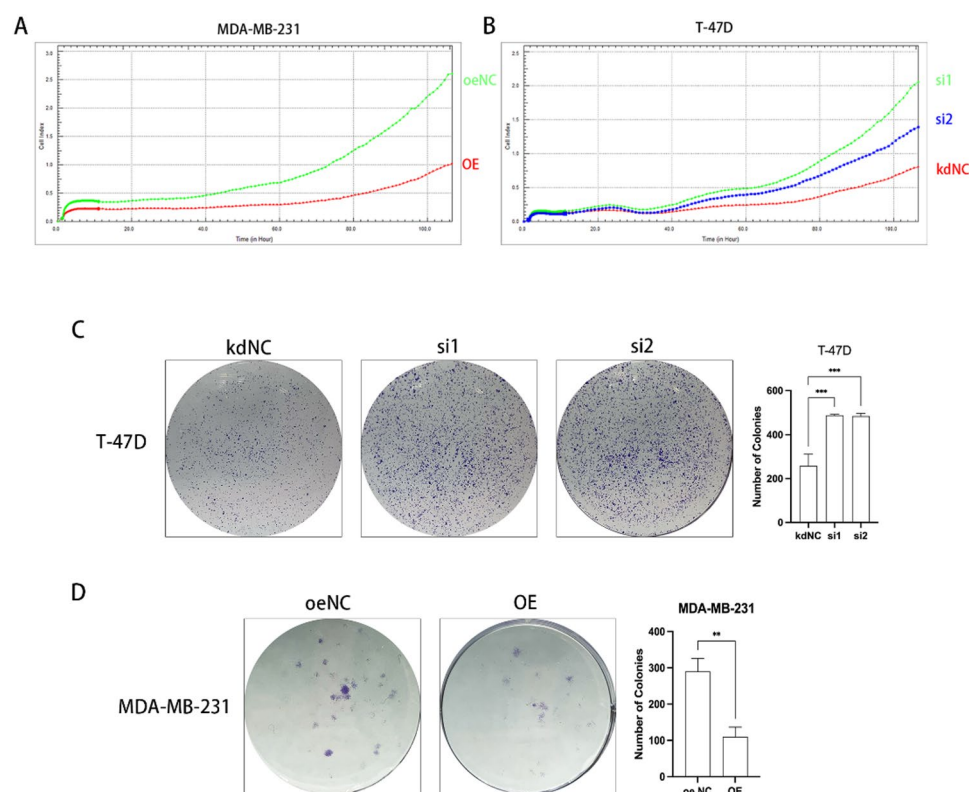


**Fig. 9.** Go enrichment analysis results of differential alternative splicing events.

significantly lower than that in the control group, the expression of total PARP was higher than that in the control group, and the expression level of cleaved PARP was significantly lower than that in the control group; In the overexpression group, the expression of total caspase-3 and cleaved PARP decreased, while the expression of cleaved caspase-3 and total PARP increased. Original blot images are provided in Supplementary File.

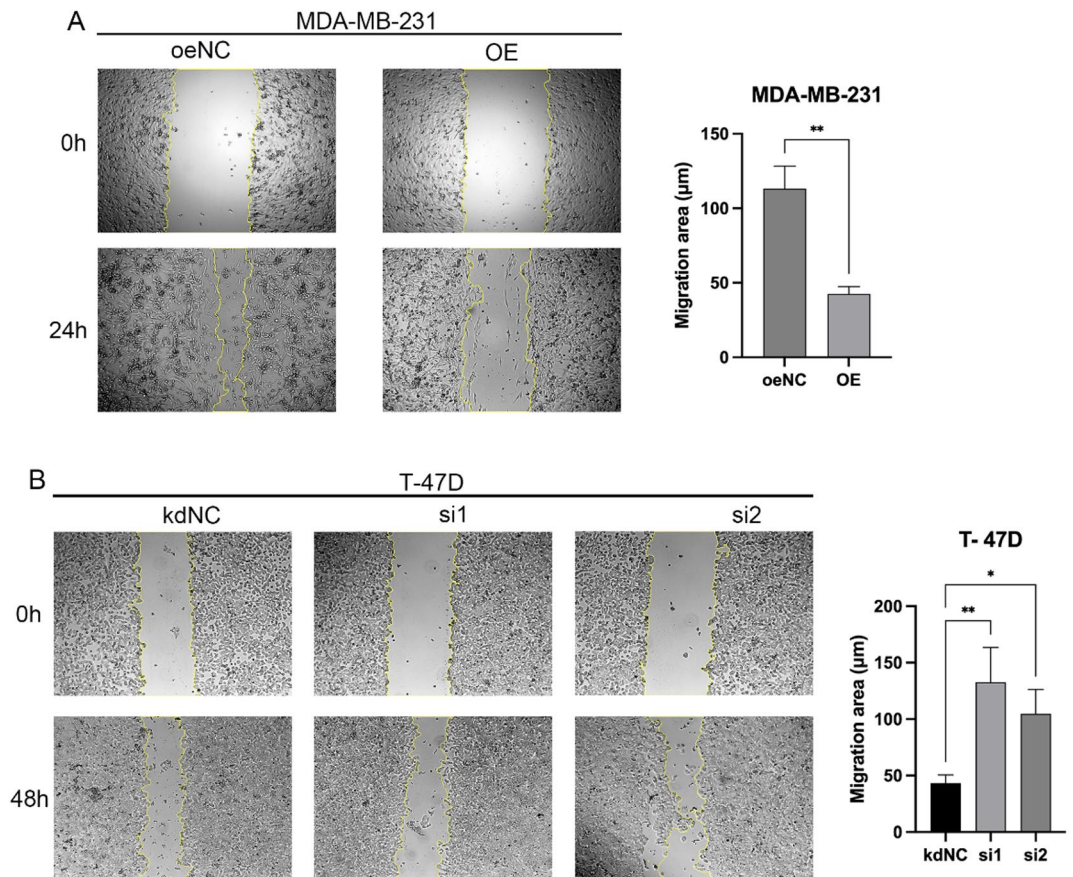


**Fig. 10.** Go and KEGG pathway enrichment analysis results of differential alternative splicing events. **A.** Go pathway enrichment analysis results; **B.** KEGG pathway enrichment analysis results.



**Fig. 11.** Cell proliferation assay to explore the effect of serpina1 gene on the proliferation of breast cancer cells. **A.** The results of cell proliferation experiment based on the real-time monitoring cell system showed that the proliferation rate of the overexpression group was lower than that of the control group (the green dotted line represents the cells of the control group, and the red dotted line represents the cells of the overexpression group); **B.** The results of cell proliferation experiment based on real-time monitoring cell system showed that the proliferation rate of the knockdown group was higher than that of the control group (green dotted line represents SI1 sequence knockdown group cells, blue dotted line represents SI2 sequence knockdown group cells); **C.** The results of plate cloning experiment showed that the proliferation ability of knockdown group was stronger than that of control group; **D.** The results of plate cloning experiment showed that the proliferation ability of the overexpression group was weaker than that of the control group. (\*  $p < 0.05$ ; \*\*  $p < 0.01$ ; \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$ ).





**Fig. 12.** Effect of *SERPINA1* gene on the migration ability of breast cancer cell lines by cell scratch assay. **A.** The cell migration ability of overexpression group was decreased; **B.** The cell migration ability of knockdown group was improved (\*  $p < 0.05$ ; \*\*  $p < 0.01$ ).

### SERPINA1 gene negatively regulates PI3K/AKT signaling pathway through PTEN gene

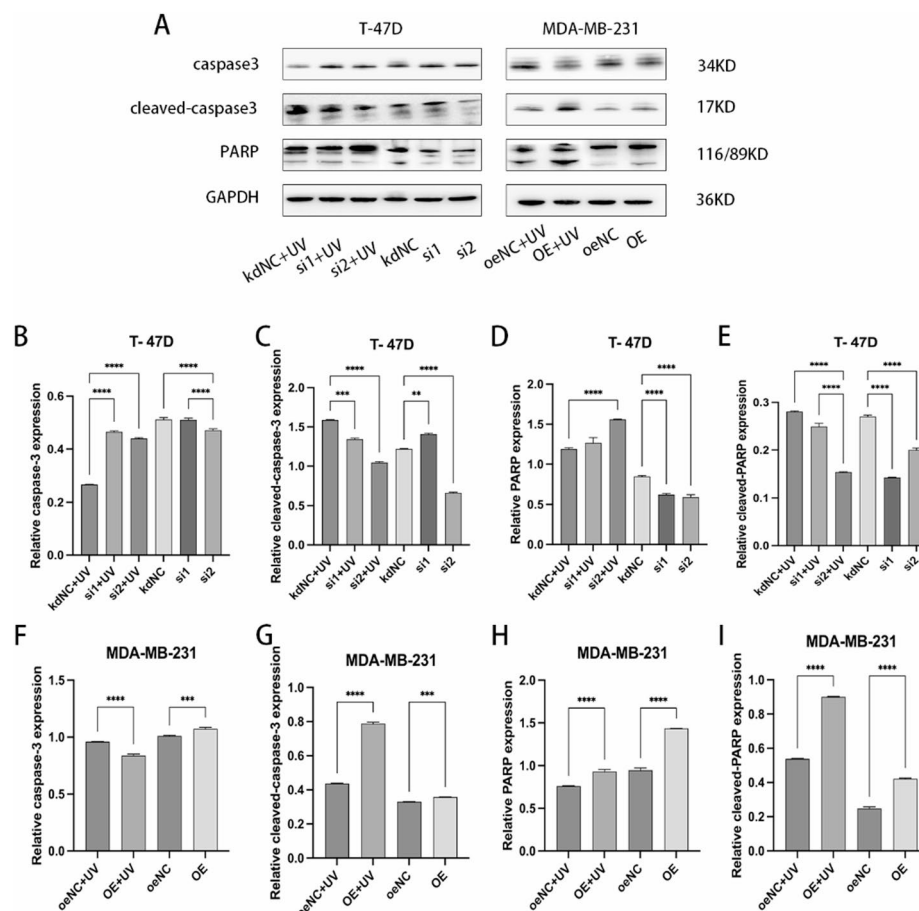
Results as shown in Fig. 14, there was no significant difference in the expression of PI3K between the knockdown group and the control group. Although there was a difference in the phosphorylation state of PI3K between the experimental group and the control group, the difference was not significant. The performance of PI3K and p-pi3k in the overexpression group was similar to that in the knockdown group. As for the important downstream effectors of PI3K, Akt and mTOR molecules, their phosphorylation status in the knockdown group was significantly enhanced compared with that in the control group, while the phosphorylation expression of this molecule in the overexpression group was decreased compared with that in the control group. Original blot images are provided in Supplementary File.

### Basic information of enrolled patients

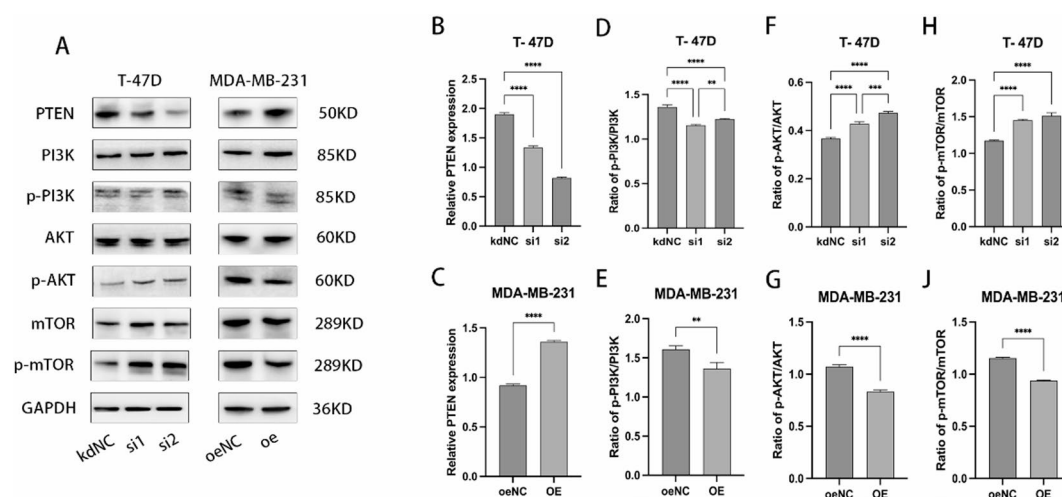
A total of 95 patients who visited the Department of breast surgery, Chinese Academy of Medical Sciences, National Cancer Center from January 2018 to December 2021 were included in this study. Among the enrolled patients, 49 were aged  $< 50$  years and 46 were aged  $\geq 50$  years; There were 26 patients (27.4%) with luminal A type, 39 patients (41.1%) with luminal B type, 20 patients (21.1%) with HER2 positive type, and 10 patients (11.6%) with triple negative type. 57 patients (60%) with clinical T stage T1, 38 patients (40%) with T2 and T3, 52 patients (54.7%) with postoperative pathology suggestive of lymph node metastasis (i.e. clinical N stage N0), and 43 patients (45.3%) with postoperative pathology suggestive of lymph node metastasis (including N1, N2, and N3). There were 42 patients (44.2%) with clinical stage I and 53 patients (55.8%) with clinical stages II and III. patients with clinical stage IV were not included in this study. The histological stage of enrolled patients was grade 1 in 11 (11.6%), grade II in 55 (57.9%), grade III in 29 (30.5%), and vascular tumor thrombus or nerve invasion positive in 45 (47.4%). Among the enrolled patients, 11 patients developed local recurrence or metastasis during the follow-up period, with a recurrence rate of 11.6%, and 1 death occurred during the follow-up period (Supplementary Table 3).

### Expression of SERPINA1 in breast cancer histopathological sections

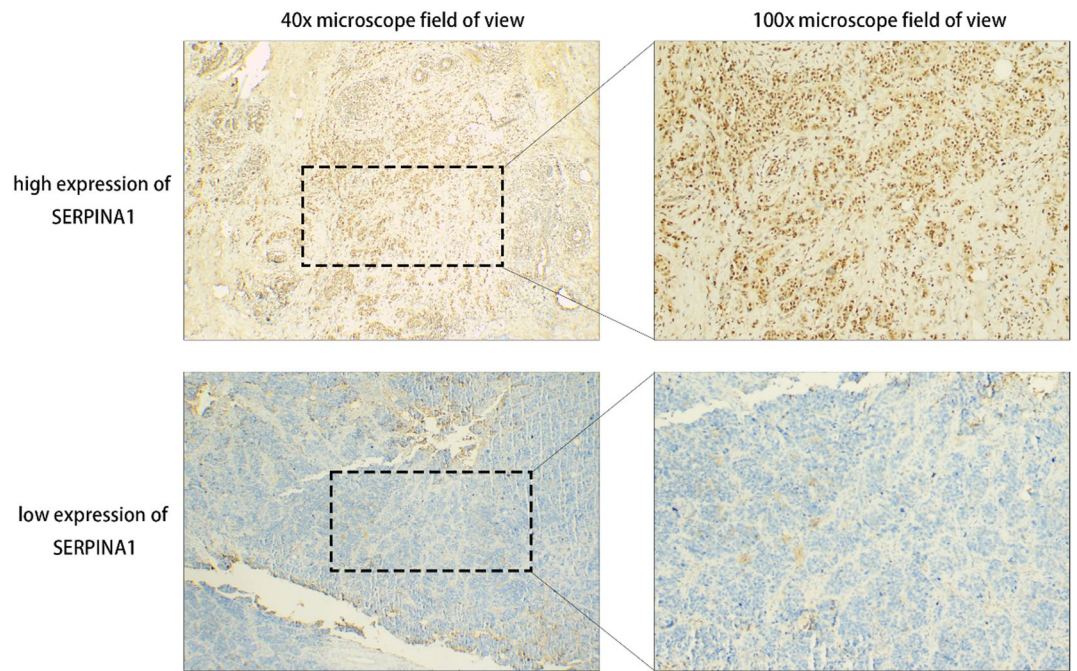
The immunohistochemical results showed that SERPINA1 was negative in most normal cells and significantly positive in the extracellular matrix. In breast cancer cells, it was low to moderate positive and mainly stained in the cytoplasm and cell membrane (Fig. 15). The results of immunohistochemical staining of pathological



**Fig. 13.** The effect of SERPINA1 on apoptosis of breast cancer cells was explored by detecting the expression of apoptosis related molecular markers caspase-3, cleavedcaspase-3, PARP, cleaved PARP.



**Fig. 14.** The key molecules and phosphorylation status of PTEN and PI3K/Akt/mTOR signaling pathways were detected by Western blot analysis.



**Fig. 15.** Examples of different SERPINA1 expression in breast cancer histopathological sections.

sections were interpreted by semi quantitative method. The results showed that among the enrolled patients, 51 patients (53.7%) had low expression of SERPINA1 and 44 patients (46.3%) had high expression.

#### Foxp3 expression and Treg lymphocyte count

Foxp3 was localized in the nuclei of Treg lymphocytes in the proliferative phase, and the blue black nuclei were positive. The FOXP3 positive Treg lymphocytes were counted, and the mean value of the counts in five random different fields under 400 times of the microscope was the final result. The results showed that the count interval of Treg lymphocytes was 0–94.2.2, with a median of 8.8. Patients with count results  $\geq 8.8$  were included in the high infiltration group of Treg lymphocytes, and patients with count results  $< 8.8$  were included in the low infiltration group. A total of 47 patients (49.5%) in the high infiltration group and 48 patients (50.5%) in the low infiltration group were enrolled (Supplementary Table 4).

#### Correlation of SERPINA1 expression with Treg lymphocyte infiltration and clinicopathological characteristics

Comparing the clinicopathological characteristics between patients with high and low SERPINA1 levels, it can be seen that there are significant differences in Treg lymphocyte infiltration status ( $p = 0.031$ ), HER2 status ( $p = 0.014$ ) and molecular typing ( $p = 0.006$ ). 27 patients in the SERPINA1 high expression group showed high infiltration of Treg lymphocytes, which was significantly more than that in the low infiltration group, while the number of patients in the low expression group showed low infiltration was significantly more than that in the high infiltration group. The positive rate of HER2 in SERPINA1 high expression group was higher (13.7%), and the number of molecular HER2 positive patients was more (15.8%), while the number of patients in low expression group was higher than that in high expression group in triple negative (9.5% vs. 1.1%), lumina A (16.8% vs. 10.5%) and luminal B (22.1% vs. 18.9%).

In this study, patients with high and low SERPINA1 expression had no significant difference in recurrence/metastasis, death and other prognosis related characteristics. Further comparing the differences between the clinicopathological characteristics of patients with high and low infiltration of Treg lymphocytes, the results showed that there were statistical differences in histological grade ( $p < 0.001$ ), ER status ( $p < 0.001$ ), PR status ( $p = 0.027$ ), HER2 status ( $p = 0.001$ ), Ki67 expression ( $p < 0.001$ ), clinical T stage ( $p = 0.009$ ) and molecular typing ( $p = 0.002$ ). Patients with high infiltration of Treg lymphocytes had higher histological grade, negative rate of ER and PR, positive rate of HER2 and expression rate of Ki67, but patients in high infiltration group tended to have lower clinical T stage. In molecular typing, patients in high infiltration group were more HER2 positive and triple negative, while patients in low infiltration group were more hormone receptor positive (Supplementary Table 5).

#### Discussion

This study utilized the TCGA database to analyze the role of SERPINA1 gene in breast cancer prognosis and immune regulation through bioinformatics methods, exploring its differential expression patterns and prognostic value. The previous research on TCGA biomarkers conducted by Liu's Laboratory provided preliminary insights for the design of this study<sup>24–27</sup>. This study found that the expression of SERPINA1 was significantly



different between tumor tissues and normal tissues, which has a certain diagnostic value. Moreover, SERPINA1 expression is also correlated with clinical and molecular typing. Further grouping analysis showed that patients in the SERPINA1 high expression group had lower T stage and higher hormone receptor expression level. In molecular typing, patients with hormone receptor positive type were more, while HER2 positive type and triple negative type were less. It can be seen that the tumorigenicity of SERPINA1 in breast cancer may be caused by the loss of its expression or abnormal function. This conclusion was also supported by the results of SERPINA1 immunohistochemical expression from clinicopathological sections. For the prognostic prediction ability, the high expression of SERPINA1 gene is significantly positively correlated with OS, DSS and PDI of breast cancer patients, which shows that SERPINA1 has a strong independent prognostic prediction ability in breast cancer, and its increased expression is associated with a better prognosis of breast cancer patients. SERPINA1 gene has shown its influence on tumor microenvironment and immune regulation in many cancer species<sup>28,29</sup>. Li et al. found 24 differentially expressed coagulation genes in breast cancer, SERPINA1 is also one of them, and believes that it has a certain role in predicting the prognosis of breast invasive ductal carcinoma<sup>30</sup>. Different from the results of this study, the differential genes obtained in this study come from the invasive ductal carcinoma and ductal carcinoma in situ.

In order to further confirm the specific mechanism of SERPINA1 gene in breast cancer, this study studied it from the perspective of cell biology. The study constructed breast cancer cell lines with high and low SERPINA1 expression as the research object. First of all, this study used two cell lines with different molecular types, and found that the role of SERPINA1 gene is not only in triple negative breast cancer, but also in T-47D hormone receptor positive breast cancer. Secondly, this study showed through two proliferation experiments that overexpression of SERPINA1 can significantly inhibit cell proliferation, while knockdown of SERPINA1 significantly enhances cell proliferation ability; The results of scratch test show that when the expression of SERPINA1 increases, the migration ability of cells decreases, and when the expression of SERPINA1 decreases, the migration ability of cells can be significantly improved (especially in the inert breast cancer cell lines such as T-47D<sup>31</sup>). In addition, the results of apoptosis experiment showed that cell lines with high expression of SERPINA1 accelerated apoptosis in the environment of apoptosis induction, while cells in the SERPINA1 knockdown group could effectively resist the stimulation from apoptosis induction, which suggested that high expression of SERPINA1 could promote apoptosis of breast cancer cells, while low expression could inhibit apoptosis. Previous studies by Zhao et al. also tried to explain the biological role of SERPINA1 in breast cancer<sup>14</sup>. They used MDA-MB 231 cell line to knock down SERPINA1, and found that silencing SERPINA1 gene can inhibit the proliferation and migration of triple negative breast cancer cells. To sum up, the cell phenotype experiments showed that the high expression of SERPINA1 inhibited the proliferation and migration of breast cancer and promoted the specific biological function of apoptosis.

Through transcriptome sequencing, this study found that the genetic differences generated by the SERPINA1 gene were mainly enriched in biological functional pathways such as adaptive immune response, natural killer cell-mediated cytotoxicity, and T cell co stimulation, which provide important clues for SERPINA1 in immune microenvironment regulation. In terms of cellular phenotype and biological function, changes are more reflected in differential alternative splicing events. Functional enrichment analysis was conducted on the differential alternative splicing events between SERPINA1 gene overexpression and the control group, indicating changes in the phosphatidylinositol signaling system. The Phosphoinositide signaling system is one of the important signaling pathways in cells, involved in regulating physiological processes such as cell growth, proliferation, apoptosis, metabolism, and intracellular signaling<sup>32,33</sup>. The phosphatidylinositol signaling system is closely related to cellular physiological activities and is crucial for maintaining normal cellular function. Phosphatidylinositol 3,4,5-trisphosphate (PIP3) is phosphorylated from phosphatidylinositol diphosphate (PI (4,5) P2). PI3K (phosphoinositol 3-kinase) is a type of kinase that can phosphorylate phosphatidylinositol diphosphate (PI (4,5) P2) and convert it into PIP3, thereby playing a signaling role in cells. PIP3 plays a crucial role in cell signaling, especially in the PI3K/AKT signaling pathway, where it can participate in regulating important physiological processes such as cell growth, proliferation, survival, and metabolism by activating downstream molecules such as protein kinase B (PKB, also known as AKT)<sup>34</sup>. To verify the role and influence of SERPINA1 gene in the molecular regulatory mechanism mentioned above, this study detected the expression of key molecules on the PI3K/Akt/mTOR signaling pathway in two cell lines. The results showed that the phosphorylation status of AKT and mTOR was significantly inhibited in SERPINA1 overexpressing cell lines, while the phosphorylation expression of these two molecules was significantly increased in knockdown cell lines. This situation suggests that high expression of SERPINA1 gene has a negative regulatory effect on the PI3K/Akt/mTOR pathway, and that SERPINA1 gene mainly regulates this pathway on AKT molecules.

Previous studies have shown that the function of AKT can be inhibited by a classic phosphatase called PTEN (phosphatase and tensin kinase)<sup>35</sup>. The classic function of PTEN is to dephosphorylate PIP3 and convert it to PIP2, thereby inhibiting the activation of the PI3K/AKT signaling pathway. The deficiency or dysfunction of PTEN can lead to excessive activation of the PI3K/AKT signaling pathway, thereby promoting cancer cell proliferation, metastasis, and drug resistance<sup>36</sup>. This is similar to the ability demonstrated by SERPINA1 in this study, suggesting that SERPINA1 and PTEN may have a synergistic effect. To verify this hypothesis, this study further examined the expression of PTEN gene under different expression states of SERPINA1. The results showed that in the SERPINA1 gene overexpression group, the expression level of PTEN significantly increased, while in the knockdown group, the expression of PTEN also decreased. It can be seen that the SERPINA1 gene and PTEN do have a synergistic effect, jointly participating in the regulation of the PI3K/Akt pathway and playing an important negative regulatory role in cell proliferation, migration, and other aspects. Further experiments are needed to demonstrate how these two genes interact with each other. Existing research has demonstrated that new technologies, such as AI, exhibit rapid development and potential in the biomedical

field<sup>37</sup>. Subsequent studies may consider further exploring the mechanism and clinical applications of SERPINA1 by integrating relevant new technologies based on this research.

The influence of SERPINA1 in breast cancer is not only limited to intracellular signal regulation, but also the influence on tumor immune microenvironment. As mentioned earlier, changes in the SERPINA1 gene can affect tumor immune regulation through high-throughput transcriptome sequencing. This conclusion has been mentioned in multiple studies, such as the pan cancer study by Kuai et al.<sup>38</sup>, which confirmed a strong link between SERPINA1 and numerous tumor infiltrating lymphocytes (TILs). But the specific mechanism is not yet clear. Christian Mueller and others have done a clinical study on recombinant adeno-associated virus (rAAV)<sup>39</sup>. This result implied that SERPINA1 may have the function of recruiting regulatory T cells to a certain extent. Therefore, our study detected the correlation between the expression of SERPINA1 and Foxp3 positive regulatory T cells (Treg) in breast tumor tissues by immunohistochemistry. The results showed that the number of Foxp3 positive Treg cells in the high expression group was significantly higher than that in the low expression group, which verified that SERPINA1 high expression can recruit Treg cells, thus producing changes in the tumor microenvironment. It is thought-provoking that the functions of Treg cells in the tumor microenvironment are complex and diverse. They inhibit the anti-tumor immune response, promote the growth of tumors and evade the surveillance of the immune system through a variety of mechanisms<sup>40</sup>. In this study, patients in the high Treg cell infiltration group also did have poor prognosis related factors (such as higher histological grade, hormone receptor negativity, poor prognosis analysis typing, etc.), which seems to conflict with the function of SERPINA1 gene in tumor cells. It can be inferred that SERPINA1 gene plays a dual role in the occurrence and development of breast cancer cells. The differences in pathological features observed in the clinic may stem from the dominant gene function, while the immune regulation function is relatively weak. This functional imbalance ultimately affects the clinical manifestations in the real world.

In addition, the synergistic effect of SERPINA1 and PTEN may jointly negatively regulate the PI3K/AKT pathway and affect cell proliferation and migration. In terms of tumor immune microenvironment, the high expression of SERPINA1 is related to the recruitment of regulatory T cells (Treg), which may change the tumor microenvironment and affect tumor growth and immune surveillance. Although Treg cells have immunosuppressive effects in the tumor microenvironment, the high infiltration of Treg cells in this study is associated with poor prognosis related factors, which indicates that there may be duality in the function of SERPINA1 in breast cancer. The latest study suggests that tumor-associated macrophages (TAMs) are plastic and can modify their functions in response to different environmental stimuli to inhibit or promote tumor progression<sup>41</sup>, similar to the dual role observed in this study. In addition, recent studies have shown that the expression levels of hub genes (CRTAM, CLEC2D and KLRB1) in depletion-type CD8 + T cells (CD8Tex) in breast cancer are closely related to the clinical characteristics of patients<sup>42</sup>. The discovery of these immune-related biomarkers provides a new perspective for immunotherapy in breast cancer. This study reveals the potential of SERPINA1 in breast cancer immunotherapy, though its specific mechanisms require further investigation. Immunotherapy has emerged as a pivotal breakthrough in malignant tumor treatment. Recent research indicates that multiple blood biomarkers can effectively predict immune response, promising to enhance both precision and patient prognosis<sup>43</sup>. Splendiani et al. demonstrated that multi-gene editing and combination therapies show remarkable potential to overcome immune suppression in tumor microenvironments and boost anti-tumor immunity<sup>44</sup>. Meanwhile, Yu et al. highlighted that checkpoint inhibitors-based immunotherapy, as an innovative approach, have been widely adopted in clinical oncology<sup>45</sup>. In combination with the above studies, SERPINA1 may be used as a biomarker to predict the response of breast cancer immunotherapy. In addition, it can be combined with gene editing technology in the future as a new therapeutic target for breast cancer. In conclusion, SERPINA1 gene plays a complex role in the occurrence, development and prognosis of breast cancer. Its role in intracellular signal regulation and tumor immune microenvironment deserves further study, in order to provide new strategies and targets for the treatment of breast cancer.

Our research still has certain limitations, and there are unavoidable biases when using extensive datasets such as TCGA and sequencing techniques for research<sup>46</sup>. In the future, we need to further improve our experiments in order to draw more scientific conclusions. We can consider using patient -derived xenograft models to further validate our findings<sup>47,48</sup>.

This study revealed the high expression of SERPINA1 gene is associated with better prognosis of breast cancer patients, and may affect the development of breast cancer by negatively regulating PI3K/AKT signaling pathway and tumor immune microenvironment in cooperation with PTEN. This study provides a new molecular target for the precise treatment of breast cancer and a scientific basis for the development of future treatment strategies.

## Materials and methods

### Experimental materials

#### *Human breast cancer cell lines and culture conditions*

The cell lines used in the experiment were MDA-MB-231 human breast cancer cell line and T-47D human breast duct cancer cell line. Both cell lines were purchased from Wuhan Pricella Biotechnology Co., Ltd., and the cell lines were identified as correct cell lines by STR. MDA-MB-231 human breast cancer cells were cultured in DMEM high glucose medium containing 10% fetal bovine serum, and T-47D human breast cancer cells were cultured in RPMI-1640 medium containing 0.1% human insulin injection and 10% fetal bovine serum. All cells were cultured in a 37 ° C incubator with 5%CO<sub>2</sub> saturated with humidity.

### Overexpression lentivirus

The design and chemical synthesis of SERPINA1 gene overexpression lentivirus were completed by Shanghai Genechem Co., Ltd. Both the overexpression lentivirus and its negative control lentivirus used GV492 vector, and the sequence of vector elements was Ubi-MCS-3FLAG-CBh-gcGFP-IRES-puromycin.



## Design and synthesis of small interfering RNA (siRNA)

The double stranded siRNA targeting *serpina1* gene was designed and completed by jintuosi (Wuhan) Biotechnology Co., Ltd. Three different *serpina1* sequences were used for synthesis, si1: 5'-GCCAGCTTACA TTTACCCAAA-3', si2: 5'-CCAGTCCAACAGCACCAATA-3', si3: 5'-GCCAAGAAACAGATCAACGAT-3'. Two sequences with high knockdown efficiency were screened for subsequent experiments. Main experimental reagents and antibodies (Supplementary tables 6 and table 7), Main experimental instruments (Supplementary table 8).

## Experimental methods

### Obtain the expression of SERPINA1 from public databases

Download RNAseq data from TCGA database and GTEx data, Download TCGA RNAseq data from TCGA website and GTEx RNAseq data from XENA database. Data acquisition time is August 2022 and the data will be uniformly processed into TPM format through the trail process. The corresponding TCGA data of 33 types of cancer and the corresponding normal tissue data in GTEx were extracted, including 9807 tumor samples from TCGA database, 727 paracancerous samples, and 7568 normal samples from GTEx data. The expression of SERPINA1 in all samples was extracted, and the Wilcoxon rank sum test was used to analyze the expression of SERPINA1 between tumor tissues and normal and adjacent tissues. ROC analysis of breast cancer data was performed using the *proc* package in R software (Version 4.2.1) to explore the clinical diagnostic significance of SERPINA1 in breast cancer<sup>42,49</sup>. Data were visualized using the *ggplot2* package in R software.

### Correlation analysis between SERPINA1 expression and clinicopathological characteristics of breast cancer patients

RNAseq data and corresponding clinical information of breast cancer patients were obtained through the public database TCGA. Patients were divided into different groups according to their clinicopathological characteristics (such as tumor size, grade, lymph node metastasis status, etc.). The SERPINA1 gene expression level of breast cancer patients in each group was calculated. According to the median expression of SERPINA1 as the cutoff value, the patients were divided into low expression group and high expression group. T-test or analysis of variance were used to compare the differences between the clinicopathological characteristics of patients in high and low expression groups. Subsequently, Pearson correlation coefficient test was used to analyze the correlation between SERPINA1 gene expression level and clinical pathology of breast cancer patients.

### Correlation analysis between SERPINA1 expression and prognosis of breast cancer patients

To further explore the effect of SERPINA1 on the prognosis of breast cancer patients, Cox regression analysis was used to compare the prognosis between the high and low SERPINA1 groups. First, the survival package was used to test the proportional hazards assumption between the two groups. When the independent variable met the equal proportional hazards assumption ( $p > 0.05$ ), then the fitting survival regression analysis was carried out. The *survminer* package in R software was used for survival analysis, and the *ggplot2* package was used for visualization of the results. The overall survival (OS), disease-specific survival (DSS), and progression free interval (PFI) were set as prognostic parameters, and Kaplan Meier curves were drawn.

### SERPINA1 expression and immune cell infiltration analysis

To explore the effect of SERPINA1 expression on immune cell infiltration in breast cancer patients, the *ssgsea* algorithm in R software was used to calculate the immune cell infiltration in different groups according to the expression of 24 kinds of immune cell marker antibodies<sup>50</sup>, and then the Mann Whitney U test was used to compare the difference of immune infiltration results between higher and lower expression groups. Spearman correlation coefficient was used to analyze the correlation between SERPINA1 expression level and immune infiltration data, and the results were visualized using *ggplot2* package.

### Cell line selection and culture conditions

The cell lines used in the experiment were human breast cancer cell line MDA-MB-231 and human breast duct cancer cell line T-47D. The experimental cell lines were stored in liquid nitrogen in 1.5 ml cryotubes for long-term storage. Put the frozen cell strain into a 37 °C thermostatic water bath and shake it for 1 min. After the cell suspension is completely dissolved, centrifuge it at 800 rpm for 5 min. Remove the supernatant, add fresh medium to resuspend the cells and transfer it to a sterile cell culture dish. Supplement the culture medium according to the size of the dish. Shake the cell suspension by cross method and place it in a 37 °C thermostatic incubator with 5% CO<sub>2</sub> for culture. Observe the cell growth and fusion with a microscope, and change the culture medium in time to ensure that the cell growth environment is free of pollution. The cells were subcultured when the cell fusion rate reached 80%–90% by microscopic observation. The cells in good growth condition were gently washed with PBS for three times, digested with 0.25% trypsin, and the cytoplasm retracted under the microscope. After the cell space increased, sufficient complete medium was added in time to stop digestion. The adherent cells were gently blown with a pipette to form a cell suspension. According to the number of cells, the cell suspension was inoculated into 2–3 sterile culture dishes, and the passage cycle was about 2–3 days. Follow up experiments were carried out after observing the cells in logarithmic growth period under the microscope.

### Lentiviral transfection experiments and screening of stable Transgenic strains

Cells in logarithmic growth phase were seeded into 6-well plates with  $1 \times 10^5$  cells per well, and three duplicate wells were set for parallel infection. After the cells were cultured overnight, the corresponding virus amount was added according to the MOI value of different cells, and the virus amount was calculated according to the following formula: virus amount = (MOI × number of cells)/virus titer. Continue to culture for 72–96 h, observe

the cell growth and activity, and use puromycin to screen stably transfected cell lines on the premise of good cell growth. Use the lowest concentration of puromycin that can kill all cells determined by the pre experiment to treat the cells continuously for a week. Observe the cell death every day. When the cell fusion rate begins to increase gradually and the death rate continues to decrease, reduce the concentration of puromycin to 1/2 of the initial concentration and continue to culture to ensure the stability of transfection.

### Small interfering RNA transfection experiments

Cells with good growth and in logarithmic growth phase were selected one day in advance and seeded into 6-well plates. The number of cells in each well was controlled to be  $1 \times 10^5$ , and 3 multiple wells were set for parallel transfection. After inoculation, the cells were transfected when the confluence rate was 70–90%. Use serum-free medium to dilute lipofectamin3000 reagent and siRNA respectively in each well. Gently mix the diluted lipo3000 reagent with siRNA, and let the mixture stand for 15 min to fully combine siRNA with lipo3000 reagent to form a complex. The cells to be transfected were washed twice with PBS, and an appropriate amount of culture medium was added. The mixture of activated siRNA and lipo3000 was added into the culture medium, and the total amount of culture medium per well was 2 ml. The medium was changed to complete medium 4–6 h after transfection. The transfected cells were cultured in a 5%CO<sub>2</sub> cell incubator at 37 °C for 24–96 h, and the cells were collected for protein detection at 48 h of culture. Other cell function tests were performed within 96 h of transfection.

### Western blot analysis was used to detect the protein expression

#### *Extraction of total protein and determination of protein concentration*

Discard the supernatant of the cell sample to be extracted, rinse twice with precooled PBS solution, carefully suck up the residual liquid, add an appropriate amount of freshly prepared cell lysate (configuration ratio: pmsf: np-40 = 1:200–250) according to the number of cells, quickly scrape off the cells with a cell scraper, and collect them into the marked precooled EP tube, lyse for 40 min on ice, and shake several times in the middle. The lysed cell suspension was placed in a low-temperature high-speed centrifuge at 4 °C and centrifuged at 12,000 rpm for 15 min. Suck the supernatant into the newly labeled EP tube to extract the total cellular protein. Bradford method was used to determine the concentration of freshly extracted total cell protein: 2 µl of protein sample was added to 400 µl of G250 staining solution for shaking incubation, and the concentration of each sample was calculated by comparing with the standard curve. According to the sample concentration, add 5×protein loading buffer for concentration balancing, and ensure that the loading amount of each sample protein is 30 µg. The protein samples were boiled in a 100 °C water bath for 10 min for protein denaturation, and the extracted protein samples were stored in a –80 °C refrigerator.

#### *SDS-PAGE gel electrophoresis and dry transfer membrane*

Use 10% denatured protein electrophoresis precast gel for electrophoresis and the glue we used is resolving gel. Pour enough electrophoresis buffer into the electrophoresis tank, load the protein samples to be tested in sequence, run the electrophoresis at 130v for 40–60 min, and observe that the bromophenol blue is moved down to the outside of the gel to stop the electrophoresis. After electrophoresis, the protein was transferred to the PVDF membrane using a semi dry membrane rotator under the condition of 15 V for 90 min.

#### *After antibody incubation and gel imaging*

Use 5% skimmed milk to block for 1 h at room temperature on a low-speed shaker. After blocking, rinse three times with sufficient PBST solution for 5 min each time. After adding the primary antibody diluted in proportion, incubate it overnight in a low-speed shaker in a 4 °C refrigerator (the primary antibody used in the experiment and their dilution ratios are shown in Supplementary Table 7). The next day, the primary antibody was recovered, and the membrane was washed three times with a sufficient amount of PBST solution. Goat anti mouse or Goat anti rabbit horseradish peroxidase labeled secondary antibody (dilution ratio 1:5000) was added according to the species of the primary antibody, incubated for 1 h at room temperature on a low-speed shaker, washed the membrane again, developed color with ECL chemiluminescence solution, exposed according to the band brightness in the LAS-4000 imaging system, and saved the band diagram. Membranes were physically trimmed around target protein regions prior to antibody incubation where appropriate. Original, uncropped blot images are provided in Supplementary File. Image J software was used to quantitatively analyze the gray value of the strip graph, and GraphPad Prism 9.0 software was used to process and visualize the results. The above experimental steps were repeated three times independently.

### Cell proliferation experiment based on realtime cellular analysis (RTCA)

The cell proliferation curve was drawn using icelligence RTCA system to observe the proliferation and growth process of cells. Prepare sample cells and cell suspensions of the control group at a concentration of  $5 \times 10^4$  cells/ml. After adding 50 µl of complete medium to the e-plate L8 detection plate, place the RTCA station to detect the baseline and confirm that the electrode contact of the detection hole is normal. Take out the detection plate, add 100 µl of evenly mixed cell suspension, so that the number of cells in each well is 5000 cells/100 µl, and set two duplicate wells for each kind of cells. The test plate was re placed in the RTCA station and allowed to stand for 30 min before starting the experiment. The system records the cell proliferation process according to the program settings, and draws the cell proliferation curve with the abscissa axis as the time and the ordinate axis as the cell index (CI value) (Supplementary Table 9).

### Clonogenic experiments

Plate clonogenic assay was used to explore the effect of SERPINA1 expression changes on the proliferation and survival ability of breast cancer cells. Single cell suspensions were prepared by conventional digestion, centrifugation and resuspension of sample cells in logarithmic growth phase and their control group. The cells were seeded in a six well plate at a density of 1000 cells per well, supplemented with sufficient complete medium, mixed with cells to prevent cells from clumping, and then transferred to the incubator for continuous culture for 7–14 days after standing for 30 min. During the period, pay attention to the cell growth and the color of the medium. If necessary, replace the medium in time to ensure the normal growth of cells. The culture can be terminated when clone formation (most clones contain more than 50 cells) is observed by naked eyes. After discarding the culture medium, wash the cells twice with PBS, add 4% paraformaldehyde tissue fixative to each well, fix for 30 min, discard the fixative, wash the cells twice with PBS again, add 0.1% crystal violet dye to each well for 30 min, recover the dye, wash the cells slowly with water, and dry the air. The number of clonogenic cells per well was observed and recorded.

### Cell scratch test

The effect of SERPINA1 on the motility and migration of breast cancer cells was observed by cell scratch assay. Mark the bottom of the six well plate, inoculate the sample cells and the control group cells into the marked six well plate, the number of cells in each well is about  $8 \times 10^5$ , and set up three duplicate wells for parallel experiments. Replenish the medium in the six well plate, and place the cells in a constant temperature incubator for culture. When the cells grow well and the confluence reaches 80%, replace the medium with serum-free medium for overnight culture. The next day, the 20  $\mu$ l gun head was used to scratch the surface of the vertical orifice plate and the horizontal line at the bottom of the plate. The floating cells were gently washed away using PBS to visualize the gap. Three fixed points were selected for each well to observe and take photos, and the scratch healing conditions of 0 h, 24 h, 48 h and 72 h were recorded under the microscope. ImageJ software was used to measure the wound healing area to calculate the cell migration rate, and graphpad prism 9.0 was used for the result statistics and visualization.

### Apoptosis experiment

Apoptosis is a highly conserved programmed cell death pathway, which is regulated by complex molecular cascades. Endogenous apoptosis is triggered by different intracellular stress signals, leading to caspase dependent or caspase independent cell death cascades. Caspase-3, as a key executive molecule, plays an important role in the process of apoptosis signaling. In the early stage of apoptosis, caspase-3 is activated to form activated caspase-3 formed by two large subunits (17kd) and two small subunits (12kd), which starts to cleave related substrates leading to apoptosis. Western blot analysis of cleavage of Caspase-3 and its substrate poly (ADP ribose) polymerase [poly (ADP ribose) polymerase, PARP]] can detect cell apoptosis. The cells were seeded in 60 mm  $\times$  15 mm dishes, and the cultured cells grew to the logarithmic phase. One group of overexpression and control group cells were selected, and the knockdown and control group cells were used for apoptosis induction. Remove the culture medium of the cell samples in the apoptosis induction group, use PBS to clean twice, add 2 ml of culture medium respectively, remove the cover of the culture dish, place the cells 50 cm away from the 15 W UV lamp for irradiation for 10 min, and then put the cells back into the incubator for further culture. After 6 h, collect the protein samples to detect the protein expression of Caspase3 and PARP.

### Transcriptome sequencing

Transcriptome sequencing was performed on cell lines transfected with SERPINA1 overexpression lentivirus and their control groups. Under the microscope, the cell samples with good growth and confluence greater than 80% were sampled, including three overexpression groups and three control groups. According to the ratio of adding 1 ml Trizol per  $2 \times 10^6$  cells, add an appropriate amount of Trizol to fully digest the cells to extract total RNA. Transfer the cell suspension to RNase free 1.5 ml centrifuge tube, use 1 ml gun head to repeatedly blow and mix until Trizol is clear and bright, and put the sample in  $-80^\circ\text{C}$  refrigerator for storage after marking. The purity, concentration and integrity of the extracted total RNA were checked. After strict quality control, the Illumina hiseq 4000 platform was used to build the database to obtain the original sequencing data file in FASTQ format. Fastp software was used to filter the original sequencing data to obtain clean data. Clean data was compared with the reference genome (GRCh38\_release110) to obtain comprehensive transcript information. To analyze the gene expression level of transcripts, first, the gene expression was counted, and the rpkm (reads per kilobase per million reads) value was used as a measure of gene expression. Subsequently, the differential expression analysis was performed between the high expression group and the control group. The deseq2 process was used to screen the differentially expressed genes with the absolute value of  $\log_2$  fold change  $> 0.58$  and  $p < 0.05$  as the standard. Finally, go (gene ontology) functional enrichment analysis and KEGG (Kyoto Encyclopedia of genes and genomes) pathway enrichment analysis were performed on the differentially expressed genes. All analyses and visualization of results were performed in R software, using the cluster profiler package for enrichment analysis, the org.hs.eg.db package for ID conversion, and the ggplot2 package for visualization of results.

### Immunohistochemical staining

The paraffin embedded postoperative pathological specimens were sectioned and fixed on glass slides, the samples were dehydrated in different concentrations of ethanol, the transparent agent was used to make the tissue transparent, the heat induced method was used for antigen thermal repair, and the nonspecific binding material was used to block the areas on the glass slides that were not bound by antibodies, so as to reduce the nonspecific background signal. The primary antibodies of SERPINA1 (dilution ratio 1:400) and Foxp3 (dilution ratio 1:500) were diluted to the appropriate concentration and dropped onto the glass slide to bind to the target

protein. The primary antibody was incubated at 4 °C overnight. The next day, the sections were washed with PBS for several times to remove the unbound antibody and nonspecific conjugate. The secondary antibody labeled with horseradish peroxidase corresponding to the primary antibody was added to the glass slide and incubated for 20 min at room temperature. The freshly prepared DAB chromogenic solution was used to stain the labeled secondary antibody, hematoxylin was used for nuclear staining, and deionized water was used to wash the sections to stop staining. The slices were placed in ammonia water for bluing treatment, and the slices were naturally dried at room temperature before sealing. After all sections were processed according to the above steps, two experienced pathologists independently scored the expression of SERPINA1 in the samples under the microscope and counted the FOXP3 positive lymphocytes.

### Immunohistochemical interpretation method

#### *SERPINA1 expression in breast cancer*

In breast cancer tissues, SERPINA1 was mainly stained in the cytoplasm and membrane, showing low to moderate positivity. The results of immunohistochemical interpretation were interpreted by semi quantitative method. Five fields of view were randomly selected under the 200 × field of the microscope, and the staining intensity of cells and the range of positive cells under the microscope were analyzed and scored. The scoring criteria are as follows: (1) according to the cell staining intensity, colorless is 0, light yellow is 1, brown yellow is 2, and 3 is brown; (2) According to the percentage of positive cells, ≤ 10% was 0, 11% – 25% was 1 point, 26% – 50% was 2 points, 51% – 75% was 3 points, and ≥ 76% was 4 points; The final score is obtained by multiplying the two scores. Patients with a score of 0–4 were assigned to the SERPINA1 low expression group, and patients with a score of 6–12 were assigned to the SERPINA1 high expression group.

#### *Foxp3 positive lymphocyte count*

Foxp3 is localized in the nucleus of Treg cells in the proliferative phase, and the blue black staining in the nucleus is positive<sup>51</sup>. The number of Foxp3 positive Treg lymphocytes was calculated by randomly selecting five representative areas that were not repeated under the 400 × field of the microscope, and the final number was the average number of positive lymphocytes in the five fields. The median of the mean value of the final count was taken as the cutoff value. Patients below the median were defined as low infiltration of Treg cells, and patients above the median were defined as high infiltration of Treg cells.

### Clinical data collection

#### **Clinical data**

Through the electronic medical record system and telephone follow-up, the information of patients who visited the Department of breast surgery, Chinese Academy of Medical Sciences, National Cancer Center from January 2018 to December 2021 was retrospectively collected. A total of 95 patients were included. All patients were diagnosed with breast cancer and received surgical treatment. The surgical scheme included modified radical mastectomy, simple mastectomy + sentinel lymph node biopsy, breast conserving + sentinel lymph node biopsy and other standard breast cancer treatment methods. All patients received regular internal medicine or radiotherapy treatment in our hospital, and were reviewed regularly after surgery. The clinical data of the patients were collected, including gender, age of diagnosis, pathological characteristics (histological analysis, tumor size, lymph node involvement, occurrence of vascular tumor thrombus, immunohistochemical indicators), clinical stage, etc. The TNM stage and clinicopathological stage of patients were determined according to the 8th edition of breast cancer clinical staging guidelines issued by the American Joint Committee on cancer (AJCC) and the Union for international cancer control (UICC).

### Discharge standard

#### *Inclusion criteria*

1) breast cancer patients who visited our hospital in 2018 and received surgical treatment; 2) The age was 18–80 years old; 3) Female; 4) Did not receive preoperative neoadjuvant chemotherapy, endocrine, targeted therapy; 5) The patients were reviewed regularly in our hospital after operation; 6) Patients with non-specific invasive breast cancer; 7) The patient's clinical personal information, case data and follow-up data are complete.

#### *Exclusion criteria*

1) male; 2) Breast cancer patients who did not receive surgical treatment; 3) Patients who received neoadjuvant chemotherapy, endocrine or targeted therapy before surgery; 4) Patients with other malignant tumors during perioperative period; 5) Patients who had undergone organ transplantation; 6) Patients with long-term oral anticoagulant drugs; 7) Patients with other serious underlying diseases; 8) Patients who underwent breast reconstruction surgery; 9) Patients with special pathological type of invasive breast cancer.

### Statistical analysis

The statistical analysis involved in the experiment in this chapter is mainly completed by graphpad prism 9.0 and R software (Version 4.2.1). The counting data are described in the form of mean ± standard deviation ( $\bar{x} \pm s$ ). Differences between two related samples were judged using paired t-test or Wilcoxon rank sum test. Pearson or Spearman tests were used to compare the correlations between continuous variables satisfying normal distribution. Mann Whitney U test was used to compare the differences between the medians of independent samples that did not meet the normal distribution. Chi square test or Fisher's exact test was used to compare the association or independence between categorical variables. If the P value of the Q test is equal to or less than 0.05, significant heterogeneity exists in the study.  $I^2$  greater than 50% is also considered heterogeneous<sup>52</sup>.



# Data availability

All datasets used and analysed during the current study available from the corresponding author on reasonable request. The sequencing data are available in the SRA repository. The accession numbers are PRJNA1251032 and associated SRA metadata are available at <https://dataview.ncbi.nlm.nih.gov/object/PRJNA1251032?reviewer=v50ulitql03eg0k7j8j4du3vh>.

Received: 29 March 2025; Accepted: 30 September 2025

Published online: 06 November 2025

# References

1. Sun, L. et al. Activating a collaborative innate-adaptive immune response to control metastasis. *Cancer cell* 39, 1361–1374.e1369, (2021). <https://doi.org/10.1016/j.ccell.2021.08.005>
2. Bray, F. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J. Clin.* 74, 229–263. <https://doi.org/10.3322/caac.21834> (2024).
3. Siegel, R. L., Giaquinto, A. N., Jemal, A. & Cancer statistics CA: a cancer journal for clinicians 74, 12–49, (2024). <https://doi.org/10.3322/caac.21820> (2024).
4. Brown, R. B., Bigelow, P., Dubin, J. A. & Mielke, J. G. High dietary phosphorus is associated with increased breast cancer risk in a U.S. Cohort of Middle-Aged women. *Nutrients* 15 <https://doi.org/10.3390/nu15173735> (2023).
5. Brown, R. B., Bigelow, P., Dubin, J. A. & Neiterman, E. Breast cancer, alcohol, and phosphate toxicity. *J. Appl. Toxicology: JAT.* 44, 17–27. <https://doi.org/10.1002/jat.4504> (2024).
6. Will, M., Liang, J., Metcalfe, C. & Chandarlapaty, S. Therapeutic resistance to anti-oestrogen therapy in breast cancer. *Nat. Rev. Cancer.* 23, 673–685. <https://doi.org/10.1038/s41568-023-00604-3> (2023).
7. Schroeder, W. T., Miller, M. F., Woo, S. L. & Saunders, G. F. Chromosomal localization of the human alpha 1-antitrypsin gene (PI) to 14q31-32. *Am. J. Hum. Genet.* 37, 868–872 (1985).
8. Di Francesco, A. et al. Novel RNA-binding activity of NQO1 promotes SERPINA1 mRNA translation. *Free Radic. Biol. Med.* 99, 225–233. <https://doi.org/10.1016/j.freeradbiomed.2016.08.005> (2016).
9. Lei, C. et al. Unraveling breast cancer prognosis: a novel model based on coagulation-related genes. *Front. Mol. Biosci.* 11, 1394585. <https://doi.org/10.3389/fmolb.2024.1394585> (2024).
10. Ercetin, E. et al. Clinical Significance of SERPINA1 Gene and Its Encoded Alpha1-antitrypsin Protein in NSCLC. *Cancers* 11, (2019). <https://doi.org/10.3390/cancers11091306>
11. Li, L. et al. Plasma Exosomal protein PLG and SERPINA1 in colorectal cancer diagnosis and coagulation abnormalities. *J. Cancer Res. Clin. Oncol.* 149, 8507–8519. <https://doi.org/10.1007/s00432-023-04776-1> (2023).
12. Li, Q. et al. Identification of Serpin peptidase inhibitor clade A member 1 (SERPINA1) might be a poor prognosis biomarker promoting the progression of papillary thyroid cancer. *Life Sci.* 329, 121938. <https://doi.org/10.1016/j.lfs.2023.121938> (2023).
13. Chan, H. J. et al. SERPINA1 is a direct Estrogen receptor target gene and a predictor of survival in breast cancer patients. *Oncotarget* 6, 25815–25827. <https://doi.org/10.18632/oncotarget.4441> (2015).
14. Zhao, Z. et al. Silence of  $\alpha$ 1-Antitrypsin inhibits migration and proliferation of triple negative breast cancer cells. *Med. Sci. Monitor: Int. Med. J. Experimental Clin. Res.* 24, 6851–6860. <https://doi.org/10.12659/msm.910665> (2018).
15. Bai, X. et al.  $\alpha$ 1-Antitrypsin Binds to the Glucocorticoid Receptor with Anti-Inflammatory and Antimycobacterial Significance in Macrophages. *Journal of immunology (Baltimore, Md.)* 209, 1746–1759, (2019). <https://doi.org/10.4049/jimmunol.2200227> (2022).
16. Ehlers, M. R. Immune-modulating effects of alpha-1 antitrypsin. *Biol. Chem.* 395, 1187–1193. <https://doi.org/10.1515/hsz-2014-0161> (2014).
17. Elshikha, A. S. et al. Alpha 1 antitrypsin inhibits dendritic cell activation and attenuates nephritis in a mouse model of lupus. *PloS One.* 11, e0156583. <https://doi.org/10.1371/journal.pone.0156583> (2016).
18. Marcondes, A. M. et al.  $\alpha$ 1-Antitrypsin (AAT)-modified donor cells suppress GVHD but enhance the GVL effect: a role for mitochondrial bioenergetics. *Blood* 124, 2881–2891. <https://doi.org/10.1182/blood-2014-04-570440> (2014).
19. Tawara, I. et al. Alpha-1-antitrypsin monotherapy reduces graft-versus-host disease after experimental allogeneic bone marrow transplantation. *Proc. Natl. Acad. Sci. U.S.A.* 109, 564–569. <https://doi.org/10.1073/pnas.1117665109> (2012).
20. Ghandi, M. et al. Next-generation characterization of the cancer cell line encyclopedia. *Nature* 569, 503–508. <https://doi.org/10.1038/s41586-019-1186-3> (2019).
21. Kanehisa, M. Toward Understanding the origin and evolution of cellular organisms. *Protein Science: Publication Protein Soc.* 28, 1947–1951. <https://doi.org/10.1002/pro.3715> (2019).
22. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe, M. KEGG: biological systems database as a model of the real world. *Nucleic Acids Res.* 53, D672–d677. <https://doi.org/10.1093/nar/gkac909> (2025).
23. Kanehisa, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. <https://doi.org/10.1093/nar/28.1.27> (2000).
24. Li, Y. & Liu, H. Clinical powers of aminoacyl tRNA synthetase complex interacting multifunctional protein 1 (AIMP1) for head-neck squamous cell carcinoma. *Cancer Biomark. A.* 34, 359–374. <https://doi.org/10.3233/cbm-210340> (2022).
25. Liu, H. & Li, Y. Potential roles of Cornichon family AMPA receptor auxiliary protein 4 (CNIH4) in head and neck squamous cell carcinoma. *Cancer Biomark. A.* 35, 439–450. <https://doi.org/10.3233/cbm-220143> (2022).
26. Liu, H. & Weng, J. A comprehensive bioinformatic analysis of cyclin-dependent kinase 2 (CDK2) in glioma. *Gene* 822, 146325. <https://doi.org/10.1016/j.gene.2022.146325> (2022).
27. Liu, H., Weng, J., Huang, C. L. & Jackson, A. P. Is the voltage-gated sodium channel  $\beta$ 3 subunit (SCN3B) a biomarker for glioma? *Funct. Integr. Genom.* 24, 162. <https://doi.org/10.1007/s10142-024-01443-7> (2024).
28. Sun, Y. F. et al. Predictive potentials of glycosylation-related genes in glioma prognosis and their correlation with immune infiltration. *Sci. Rep.* 14, 4478. <https://doi.org/10.1038/s41598-024-51973-0> (2024).
29. Ma, Y. et al. CEBPB-mediated upregulation of SERPINA1 promotes colorectal cancer progression by enhancing STAT3 signaling. *Cell. Death Discovery.* 10 <https://doi.org/10.1038/s41420-024-01990-9> (2024).
30. Li, J., Du, J., Wang, Y., Jia, H. A. & Coagulation-Related Gene-Based prognostic model for invasive ductal carcinoma. *Front. Genet.* 12, 722992. <https://doi.org/10.3389/fgene.2021.722992> (2021).
31. Albert, V. et al. Protein kinase C targeting of luminal (T-47D), luminal/HER2-positive (BT474), and triple negative (HCC1806) breast cancer cells in-vitro with AEB071 (Sotrastaurin) is efficient but mediated by subtype specific molecular effects. *Arch. Gynecol. Obstet.* 306, 1197–1210. <https://doi.org/10.1007/s00404-022-06434-2> (2022).
32. Kakar, R., Ghosh, C. & Sun, Y. Phosphoinositide signaling in immune cell migration. *Biomolecules* 13 <https://doi.org/10.3390/biom13121705> (2023).
33. Owusu Obeng, E. et al. Phosphoinositide-Dependent signaling in cancer: A focus on phospholipase C isozymes. *Int. J. Mol. Sci.* 21 <https://doi.org/10.3390/ijms21072581> (2020).



34. He, X. et al. The PI3K/AKT signalling pathway in inflammation, cell death and glial Scar formation after traumatic spinal cord injury: mechanisms and therapeutic opportunities. *Cell Prolif.* **55**, e13275. <https://doi.org/10.1111/cpr.13275> (2022).
35. Hlozkova, K. et al. PTEN/PI3K/Akt pathway alters sensitivity of T-cell acute lymphoblastic leukemia to L-asparaginase. *Sci. Rep.* **12**, 4043. <https://doi.org/10.1038/s41598-022-08049-8> (2022).
36. Hashemi, M. et al. Progress in targeting PTEN/PI3K/Akt axis in glioblastoma therapy: revisiting molecular interactions. *Biomed. pharmacotherapy = Biomedicine Pharmacotherapie.* **158**, 114204. <https://doi.org/10.1016/j.biopha.2022.114204> (2023).
37. Chen, Z. & Er Saw, P. Integration in Biomedical Science : Emerging Trends in the Post-Pandemic Era. (2024).
38. Kuai, X. et al. Member 1 is prognostic and involved in immunological regulation in human cancers. *Int. J. Mol. Sci.* **24** <https://doi.org/10.3390/ijms241411566> (2023).
39. Mueller, C. et al. Human Treg responses allow sustained Recombinant adeno-associated virus-mediated transgene expression. *J. Clin. Invest.* **123**, 5310–5318. <https://doi.org/10.1172/jci70314> (2013).
40. Ma, M. et al. Lymphatic endothelial cell-mediated accumulation of CD177(+)Treg cells suppresses antitumor immunity in human esophageal squamous cell carcinoma. *Oncoimmunology* **13**, 2327692. <https://doi.org/10.1080/2162402x.2024.2327692> (2024).
41. Chen, S., Li, S. & Wang, H. Remodeling tumor-associated macrophages in the tumor microenvironment. **10**, 281–285, (2024). <https://doi.org/10.1097/ot9.000000000000063>
42. Liu, H., Dong, A., Rasteh, A. M., Wang, P. & Weng, J. Identification of the novel exhausted T cell CD8+ markers in breast cancer. *Sci. Rep.* **14**, 19142. <https://doi.org/10.1038/s41598-024-70184-1> (2024).
43. Ramapriyan, R. et al. Altered cancer metabolism and implications for next-generation CAR T-cell therapies. *Pharmacol. Ther.* **259**, 108667. <https://doi.org/10.1016/j.pharmthera.2024.108667> (2024).
44. Splendiani, E. et al. Immunotherapy in melanoma: can we predict response to treatment with Circulating biomarkers? *Pharmacol. Ther.* **256**, 108613. <https://doi.org/10.1016/j.pharmthera.2024.108613> (2024).
45. Yu, X. et al. Evolving perspectives regarding the role of the PD-1/PD-L1 pathway in gastric cancer immunotherapy. *Biochim. Et Biophys. Acta Mol. Basis Disease.* **1870**, 166881. <https://doi.org/10.1016/j.bbadis.2023.166881> (2024).
46. Liu, H., Li, Y., Karsidag, M., Tu, T. & Wang, P. Technical and biological biases in bulk transcriptomic data mining for cancer research. *J. Cancer.* **16**, 34–43. <https://doi.org/10.7150/jca.100922> (2025).
47. Kang, Z. R. et al. Deficiency of BCAT2-mediated branched-chain amino acid catabolism promotes colorectal cancer development. *Biochim. Et Biophys. Acta Mol. Basis Disease.* **1870**, 166941. <https://doi.org/10.1016/j.bbadis.2023.166941> (2024).
48. Lee, S. J. et al. hsa-miR-CHA2, a novel microRNA, exhibits anticancer effects by suppressing Cyclin E1 in human non-small cell lung cancer cells. *Biochim. Et Biophys. Acta Mol. Basis Disease.* **1870**, 167250. <https://doi.org/10.1016/j.bbadis.2024.167250> (2024).
49. Liu, H. & Tang, T. A bioinformatic study of IGFs in glioma regarding their diagnostic, prognostic, and therapeutic prediction value. *Am. J. Translational Res.* **15**, 2140–2155 (2023).
50. Bindea, G. et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* **39**, 782–795. <https://doi.org/10.1016/j.immuni.2013.10.003> (2013).
51. Georgiev, P., Charbonnier, L. M., Chatila, T. A. & Regulatory, T. Cells: the many faces of Foxp3. *J. Clin. Immunol.* **39**, 623–640. <https://doi.org/10.1007/s10875-019-00684-7> (2019).
52. Higgins, J. P. & Thompson, S. G. Quantifying heterogeneity in a meta-analysis. *Stat. Med.* **21**, 1539–1558. <https://doi.org/10.1002/sim.1186> (2002).

## Author contributions

CL and WL are responsible for data analysis of the database. WL, CL and YL completed cell experiments. LX, CG and NW provided patients. HY and KZ collected the samples. LX and CG guided experiments. CL and WL analyzed and interpreted the data. WL finished manuscript editing. LX and CG reviewed the manuscript. All authors read and approved the final manuscript.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Natural Science Foundation of China (No.82372078) and Beijing Xisike Clinical Oncology Research Foundation (No.Y-NESTLE2022MS-0036).

## Declarations

## Competing interests

The authors declare no competing interests.

## Ethics committee approval and patient consent

All clinical samples were collected from the Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College. Written informed consent was obtained from all participants enrolled in this study, and ethical approval was obtained from the following institutional review boards in accordance with the Declaration of Helsinki: National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College. Approval number: NCC2023C-479. Issued date: 2023-07-18.

## Consent for publication

All authors approved the final manuscript and the submission to this journal.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-22658-z>.

**Correspondence** and requests for materials should be addressed to L.X. or C.G.

**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) 2025