Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Prognosis and immunotherapy response in breast cancer can be predicted using ARHGAP26, IL33, and BST2

Mingyuan Zhao
Zhejiang Hospital
QiLong Zhang
Zhejiang Hospital
Luoning Zheng
Zhejiang Hospital
Jing Li
Zhejiang Hospital
Fang Peng
Zhejiang Hospital
Shuangyan Lin
linshy1986@163.com
Zhejiang Hospital

Research Article

Keywords: Breast cancer, ARHGAP26, IL33, BST2, Autophagy, Tumor Microenvironment

Posted Date: September 8th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10739395/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background: Breast cancer remains the most common malignancy among women worldwide, with clinical outcomes frequently compromised by drug resistance and immune evasion.

Autophagy plays a dual role in both tumor progression and immune microenvironment remodeling. This study aimed to develop an autophagy-related prognostic signature to predict patient survival and therapeutic response.

Methods: Transcriptomic and clinical data were obtained from the TCGA-BRCA and GSE1456 cohorts. Univariate and multivariate Cox regression analyses combined with consensus clustering were applied to identify autophagy-related prognostic genes and establish a risk signature. Patients were stratified into high- and low-risk subgroups, and between-group differences were evaluated using Kaplan–Meier analysis, time-dependent ROC curves, immune cell infiltration profiling, tumor mutational burden, immune checkpoint expression, and chemotherapy and immunotherapy sensitivity predictions. Single-cell RNA sequencing and immunohistochemical staining validated the expression patterns of the core genes.

Results: Three autophagy-related genes—ARHGAP26, IL33, and BST2—were ultimately identified to construct the prognostic signature. High-risk patients exhibited significantly worse overall survival, accompanied by decreased infiltration of antitumor immune cells, lower immune and stromal scores, reduced expression of multiple immune checkpoint genes, a weaker response to immune checkpoint blockade immunotherapy, and greater resistance to conventional chemotherapeutic and targeted agents. Single-cell analysis revealed cell-type-specific expression patterns of the three hub genes within tumor tissues. Immunohistochemistry showed that ARHGAP26 and IL33 were downregulated, whereas BST2 was upregulated in breast cancer.

Conclusion: This validated autophagy-related signature reliably predicts survival, chemosensitivity, and immunotherapy efficacy in breast cancer, offering novel biomarkers to guide individualized clinical treatment.

1. Introduction

Breast cancer represents the most prevalent malignancy in women globally, and its main treatments—surgery, chemotherapy, endocrine therapy, and targeted therapy—each have considerable limitations[1, 2]. Triple-negative breast cancer, for instance, lacks established hormone receptors and HER2 targets, making chemotherapy the primary treatment option. This approach often leads to the rapid development of drug resistance, and patients with advanced-stage disease have a five-year survival rate of only about 12%[3]. Hormone receptor-positive breast cancer initially responds to endocrine therapy but frequently acquires resistance and relapses after extended treatment[4, 5]. Although targeted therapy for HER2-positive disease has shown efficacy, some patients still experience treatment failure or disease progression[6]. These ongoing challenges have driven the search for alternative therapeutic approaches.

Autophagy is a highly conserved cellular process of degradation and recycling, exerting a complex, bidirectional, and stage-dependent role in tumors [7]. Consequently, it has emerged as a biological mechanism of profound interest. Moreover, many chemotherapeutic agents function by either blocking the access of tumor cells to essential nutrients or by damaging their genetic material[8]. Under these circumstances, tumor cells may develop drug resistance, as autophagy can be activated within them, thereby enabling their survival despite exposure to chemotherapeutic drugs[9]. From an immunological perspective, autophagy attenuates antitumor immune responses by degrading molecules critical for cancer cell destruction, such as granzyme B, thereby diminishing the efficacy of T cells and natural killer cells. Through hypoxia-triggered signaling pathways, autophagy also enhances immunosuppression by reducing the intensity of immune attacks[10, 11]. Furthermore, autophagy in tumor-associated macrophages alleviates oxidative stress, which contributes to the creation of a microenvironment that shields tumors from immune clearance[12, 13]. In established breast tumors, autophagy facilitates the clearance of damage induced by chemotherapy and radiotherapy, thereby preventing apoptosis and enabling cancer cell survival during treatment [14]. Therefore, autophagy represents an ideal target for cancer therapies due to its multifaceted roles in survival, metabolism, dissemination, and immune evasion[15]. Combination strategies involving the modulation of autophagy are attracting increasing attention from researchers.

Building on recent advances in high-throughput sequencing and bioinformatics, we used publicly available data to construct a prognostic model based on autophagy-related genes, evaluated its predictive value for immunotherapy response, and examined the association between the expression of the three model genes and patient survival in breast cancer.

2. Methods

2.1 Data collection

Gene expression RNA sequencing (RNA-seq) data from 1053 breast cancer samples and 113 normal breast samples were obtained from the University of California, Santa Cruz (UCSC) Xena database (https://xenabrowser.net/datapages/), and the GSE1456 and GSE GSE305078 dataset from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). Using the GeneCards (https://www.genecards.org/), National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/gene) and KEGG (https://www.kegg.jp/kegg/pathway) databases, we compiled a set of 487 autophagy-related genes.

2.2 DEGs analysis

Differential expression analysis was conducted using the DESeq2 package, with $|\log_2$ fold change$| > 1$ and $P < 0.05$ set as the thresholds for statistical significance. Differentially expressed autophagy-related genes (DEARGs) were identified by overlapping the differentially expressed genes (DEGs) with the autophagy-related genes (ARGs).

2.3 Pathway enrichment analysis

Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and Gene Set Enrichment Analysis (GSEA) were conducted using the "clusterProfiler" R package to identify biological processes (BP), molecular functions (MF), cellular components (CC) relevant to breast cancer subtypes, applying a significance level of $p < 0.05$.

2.4 Univariate and Multivariate COX analysis

Univariate and multivariate Cox regression analyses were conducted to evaluate the association between autophagy-related genes and the prognosis of breast cancer patients. Hazard ratios and their corresponding 95% confidence intervals were calculated for each gene, with statistical significance defined as $p < 0.05$. Risk score distributions, survival outcomes, and expression heatmaps were generated using the "ggrisk" R package.

2.5 Clustering analysis

Unsupervised clustering was conducted using the 'ConsensusClusterPlus' R package based on the DEARGs identified by univariate Cox regression, categorizing breast cancer patients into molecular subtypes.

2.6 Survival analysis

Survival differences between breast cancer subtypes were assessed using Kaplan–Meier curves. Patients were stratified according to optimal gene expression cut-offs (statistical threshold: $p < 0.05$). Predictive performance was evaluated using time-dependent receiver operating characteristic (ROC) curves via the "timeROC" R package, and predictive accuracy for survival probabilities at specified time points was quantified by the area under the curve (AUC).

2.7 Establishment of a prognostic scoring model

A multivariate Cox model was established to derive a prognostic scoring system. Each patient's risk score was calculated as the linear combination of gene expression values weighted by their respective regression coefficients: Risk Score = $\Sigma(Exp_i \times \beta_i)$. Based on the median risk score, patients in both the development and validation cohorts were divided into high- and low-risk subsets

2.8 Immune infiltration

Immune cell infiltration in breast cancer was quantified using the CIBERSORT algorithm, while the ESTIMATE algorithm was employed to compute ImmuneScore, StromalScore, ESTIMATEScore, and tumor purity. Violin plots were generated to visualize the association between tumor microenvironment characteristics and clinical outcomes stratified by the prognostic model.

2.9 Drug sensitivity and mutation burden analysis

Drug sensitivity prediction for BRCA was conducted using the "oncoPredict" R package, which estimated IC_{50} values for conventional breast cancer chemotherapies. Pharmacogenomic profiles were downloaded from the GDSC database (https://osf.io/c6tfx/), and somatic mutations were acquired via the GDC data portal for TCGA-BRCA (https://portal.gdc.cancer.gov/projects/TCGA-BRCA). Subsequent analyses with the "maftools" package focused on comparing tumor mutation burden (TMB) and detecting significantly mutated genes ($P < 0.05$) between risk-stratified patient groups.

2.10 Immunophenoscore and Checkpoint Gene Profiling

The immunophenoscore (IPS), a surrogate for tumor immunogenicity, predicts the likelihood of response to immune checkpoint blockade. IPS values were retrieved from The Cancer Immunome Atlas (http://tcia.at/home), and differences in checkpoint inhibitor sensitivity between the high- and low-risk groups were evaluated. Furthermore, the expression patterns of over 40 breast cancer-related immune checkpoint genes were compared across the two risk strata.

2.11 Clinical features and Immunohistochemistry (IHC)

Gene–clinical association analyses in breast cancer were performed using data sourced from bc-GenExMiner v5.2 (Breast Cancer Gene-Expression Miner v5.2). Immunohistochemical staining data were obtained from The Human Protein Atlas (HPA) public database.

2.12 Single-cell RNA transcriptomics analysis

Single-cell transcriptomic data for breast cancer were retrieved from the GEO database (accession: GSE305078). Raw expression profiles were processed with the Seurat R package through a standard workflow encompassing preprocessing, quality control, dimensionality reduction, and unsupervised cell clustering to define transcriptionally distinct populations. Cluster-specific differentially expressed genes (DEGs) were identified using the FindAllMarkers function, with statistical significance set at FDR-adjusted $p < 0.05$. Cell types were assigned to each cluster via manual annotation based on established signature gene sets from previous publications. Expression distributions of target genes across annotated cell populations were visualized using violin plots generated with the Seurat VlnPlot function.

2.13 Statistical analysis

All statistical analyses were conducted using R software (v4.4.2). Intergroup comparisons of continuous variables were performed via independent-samples t-tests, with $p < 0.05$ considered statistically significant.

3. Results

3.1 Identification of autophagy-related DEGs in breast cancer

To investigate autophagy-related genes implicated in breast cancer progression, we initially analyzed RNA-seq data from 1053 breast cancer patients and 113 normal controls in the TCGA database. This analysis identified 11,039 differentially expressed genes (DEGs), comprising 6392 upregulated and 4647 downregulated transcripts (Fig. 1A). Overlap of these DEGs with 487 curated autophagy-related genes (ARGs) yielded 82 differentially expressed autophagy-related genes (DEARGs) for downstream investigation (Fig. 1B). KEGG enrichment analysis revealed that these 82 genes were predominantly enriched in autophagy- and mitochondrial metabolism-related pathways, including autophagy of mitochondrion, mitophagy, macroautophagy, regulation of autophagy et.al (Fig. 1C). Univariate Cox regression analysis further identified eight autophagy-related genes significantly associated with breast cancer patient prognosis (Fig. 1D), highlighting their potential as prognostic biomarkers.

3.2 Autophagy-based molecular subtyping and subtype-specific DEGs

Using the eight prognosis-related autophagy genes, consensus clustering was applied to classify the 1053 TCGA breast cancer patients. Clustering stability was evaluated for k values ranging from 2 to 5 (Fig. 2A), with k = 2 yielding the most stable classification. Accordingly, all patients were partitioned into two distinct molecular subtypes (Cluster 1 and Cluster 2) for subsequent analyses (Fig. 2B). Differential expression analysis between the two subtypes identified 2289 DEGs (1042 upregulated and 1247 downregulated; Fig. 2C), revealing substantial transcriptional heterogeneity between these autophagy-based molecular subgroups.

3.3 Functional enrichment and prognostic validation of molecular subtypes

Functional enrichment analysis of the 2289 subtype-associated DEGs was conducted to explore the functional basis of this transcriptional heterogeneity. These genes were prominently enriched in KEGG immune-related signaling pathways. Neuroactive ligand–receptor interaction and cytokine–cytokine receptor interaction exhibited the highest enrichment ratios (GeneRatio > 0.10, Fig. 2D). GO enrichment further confirmed marked enrichment in adaptive immune processes, including lymphocyte-mediated immunity and B/T cell receptor complexes, as well as immune-related molecular functions such as cytokine activity and chemokine binding ($p < 0.001$) (Fig. 2E). Additionally, GSEA demonstrated significant positive enrichment of multiple immune-related pathways, including the chemokine signaling pathway, hematopoietic cell lineage, and T cell receptor signaling pathway (NES > 0, Fig. 2F), suggesting that these immune pathways are broadly activated and contribute to remodelling of the tumor immune microenvironment.

Prognostic analysis confirmed significant survival disparities between the two molecular subtypes: patients in Cluster 2 exhibited markedly better overall survival than those in Cluster 1 (Fig. 2G and H). Time-dependent ROC analysis further validated the predictive performance of this molecular subtyping,

with AUC values exceeding 0.6 for 1-, 3-, and 5-year survival prediction (Fig. 2I and J). Collectively, these findings demonstrate that autophagy-based molecular subtyping possesses reliable prognostic value.

3.4 Establishment and Validation of an Autophagy-Related Prognostic Signature for BRCA

A parsimonious prognostic model was constructed by intersecting the 2289 subtype-associated DEGs with the eight prognosis-related autophagy genes, which ultimately yielded three key autophagy-associated genes: ARHGAP26, IL33, and BST2. Using these three genes, we established a breast cancer prognostic risk model via multivariate Cox regression. The risk score for each patient was calculated as follows: Risk score = (− 0.29964 × ARHGAP26 expression) + (− 0.14506 × IL33 expression) + (− 0.08824 × BST2 expression). All TCGA breast cancer patients were stratified into high-risk (n = 530) and low-risk (n = 529) subgroups according to the median risk score. The scatter plot revealed that higher risk scores were accompanied by increased patient mortality (Fig. 3A), whereas the heatmap visualized the distinct expression patterns of the three signature genes between the two risk groups (Fig. 3B). Kaplan–Meier survival analysis demonstrated that low-risk patients had significantly superior overall survival compared with high-risk patients (Fig. 3C). Time-dependent ROC analysis confirmed robust predictive accuracy, with all 1-, 3-, and 5-year AUC values exceeding 0.6 (Fig. 3D). Notably, the external GSE1465 cohort further validated the stable prognostic performance of the risk model in both survival and ROC analyses (Fig. 3E and F).

Furthermore, univariate Cox regression analysis identified clinical subgroup, age, T stage, M stage, and risk score as significant prognostic factors ($p < 0.05$) (Fig. 3G). Subsequent multivariate Cox regression confirmed that risk score, age, T stage, and M stage were independent prognostic indicators for breast cancer patients (Fig. 3H). These findings verify the reliable prognostic value and clinical applicability of the three-gene autophagy-related signature.

3.5 Risk score-related differences in immune cell infiltration and TME scores

Given the marked immune-pathway enrichment and prognostic differences between the risk subgroups, we further explored the tumor immune microenvironment (TME) of high- and low-risk patients. The CIBERSORT algorithm was applied to quantify the infiltration proportions of 22 immune cell subtypes in breast cancer samples. The results showed that the low-risk group exhibited significantly higher infiltration of nine immune cell subsets, including B cells, CD4 + T cells, CD8 + T cells, and M1 macrophages (Fig. 4A). The distribution of all 22 immune cell subtypes across individual samples is displayed in Fig. 4B.

Consistently, ESTIMATE analysis revealed significant TME differences between the two subgroups. The low-risk group had markedly higher ImmuneScore, StromalScore, and ESTIMATEScore and lower tumor purity than the high-risk group (Fig. 4C–F). Because elevated immune and stromal scores and reduced tumor purity are closely correlated with favorable clinical outcomes[16], these differences were highly

consistent with the superior survival of low-risk patients. Moreover, stratified analysis across clinic stages and the TNM staging system yielded consistent results, with the high-risk group maintained significantly lower immune and stromal scores at all clinical stages (Fig. 4G–R). In summary, the low-risk subgroup is characterized by a highly activated immune microenvironment and an enhanced anti-tumor immune response.

3.6 Mutational spectrum analysis and risk-based chemotherapy sensitivity evaluation

To explore the genomic heterogeneity and clinical therapeutic implications of the two risk subgroups, we analyzed their somatic mutation landscapes and drug sensitivity. Waterfall plots of frequently mutated breast cancer genes showed that TP53 had the highest overall mutation frequency, with a significantly higher mutation rate in the high-risk group, whereas the low-risk group had markedly higher mutation frequencies of PIK3CA and CDH1 (Fig. 5A and B), indicating distinct genomic alteration patterns between the two subgroups.

Risk-stratified drug sensitivity analysis of commonly used breast cancer chemotherapeutic and targeted agents revealed that low-risk patients were more sensitive to a wide range of drugs, including Cisplatin, Docetaxel, Olaparib, Palbociclib, 5-Fluorouracil, Paclitaxel, Oxaliplatin, Tamoxifen, Fulvestrant, Cyclophosphamide, Ribociclib, Dactinomycin, and Vinorelbine (Fig. 5C). Together, these genomic and pharmacogenomic differences further elucidate the molecular heterogeneity of breast cancer stratified by the autophagy-related signature and provide evidence to support precision, individualized treatment strategies.

3.7 Prediction of immunotherapy efficacy and immune checkpoint expression

Considering the significant differences in immune microenvironment activity between the two subgroups, we evaluated their immunotherapy responsiveness using the immunophenoscore (IPS). The low-risk group had significantly higher IPS values than the high-risk group under single PD-1 blockade, single CTLA-4 blockade and combined immune checkpoint blockade (Fig. 5D), suggesting that low-risk patients are more sensitive to immune checkpoint inhibitor therapy.

Immune checkpoint gene expression was systematically compared between subgroups to elucidate the molecular basis of differential immunotherapy sensitivity. A total of 43 immune checkpoint genes, including CTLA-4, were significantly upregulated in the low-risk group (Fig. 5E), indicating that low-risk breast tumors possess stronger immunogenicity and a more activated immune microenvironment, and are therefore more likely to respond to immune checkpoint blockade.

3.8 Single-Cell RNA-Seq Analysis of Cell Subsets and Candidate Gene Expression Patterns

Single-cell RNA-seq data from the GSE305078 dataset were analyzed to define the cellular expression distribution of the three signature genes (ARHGAP26, IL33, and BST2) in breast cancer tissues. After stringent quality control and filtering, 65,199 cells were retained. Dimensionality reduction via t-SNE and cell annotation identified 14 distinct cell populations, among which T cells were the most abundant, indicating a T cell-dominant immune microenvironment in breast cancer tissues (Fig. 6A).

Expression profiling of the three signature genes across cell clusters revealed distinct patterns. ARHGAP26 was expressed sparsely and at low levels in a small number of cells, with almost no expression in T cells and mast cells and negligible levels in neutrophils, endothelial cells, macrophages, and vascular smooth muscle cells. IL33 was specifically and highly enriched in neutrophils, with minimal or undetectable expression in fibroblasts, B cells, and endothelial cells. In contrast, BST2 was widely expressed across multiple cell types, with prominent expression in basal epithelial cells, neutrophils, mast cells, NK cells, and endothelial cells, demonstrating widespread but variable expression across the tumor microenvironment (Fig. 6B).

3.9 Expression and Prognostic Characteristics of Three Candidate Genes

We next verified the differential expression and prognostic significance of the three signature genes in the TCGA breast cancer cohort. Compared with normal breast tissues, ARHGAP26 and IL33 were significantly downregulated in breast cancer, whereas BST2 was markedly upregulated (Fig. 6C-F). Survival analysis showed that high expression of ARHGAP26 and IL33 was significantly associated with prolonged overall survival, indicating protective prognostic roles in breast cancer (Fig. 6G and H). Although BST2 was dysregulated in tumor tissues, it showed no significant correlation with prognosis in the overall cohort (Fig. 6I). Notably, high BST2 expression predicted better survival in triple-negative breast cancer, suggesting a subtype-specific prognostic value (Fig. 6J).

3.10 Clinical Relevance and Immunohistochemical Expression of Three Candidate Genes

Based on the above cohort findings regarding gene expression and prognosis, we further explored the clinical relevance of the three candidate genes and performed immunohistochemistry to verify their protein expression patterns. We compared the mRNA expression levels of ARHGAP26, IL33, and BST2 among breast cancer subgroups stratified by TP53 status, ER status, PR status, combined ER/PR phenotypes, and HER2 status (Fig. 7A-R). All three genes exhibited significantly differential expression with respect to patient age (Fig. 7A-C). ARHGAP26 and IL33 expression correlated with P53 status (Fig. 7D and E), whereas BST2 expression showed no statistically significant association with TP53 mutation status (Fig. 7F). Expression of ARHGAP26, IL33, and BST2 was significantly associated with ER-, PR-, and ER-PR- double negativity (Fig. 7G-O). Among the three genes, only IL33 expression was correlated with HER2 status (Fig. 7P-R). These results indicate that ARHGAP26, IL33, and BST2 may participate in the molecular pathogenesis of the ER-PR- double-negative subtype.

3.11 Immunohistochemical staining

Immunohistochemical staining validated the protein expression patterns of the three genes in normal and malignant breast tissues. ARHGAP26 and IL33 were specifically expressed in the myoepithelial cells of normal breast tissues: ARHGAP26 localized to the cell membrane and cytoplasm (Fig. 8A), whereas IL33 was mainly localized to the nucleus (Fig. 8C). Both proteins were almost undetectable in breast cancer cells (Fig. 8B, 8D), indicating that the loss of ARHGAP26 and IL33 expression is a critical event in breast cancer malignant transformation. In contrast, BST2 was moderately expressed in the glandular epithelial cells of normal breast tissues (Fig. 8E) but significantly upregulated in breast cancer cells (Fig. 8F), confirming its tumor-associated expression pattern.

4. Discussion

In this study, we constructed an autophagy-related prognostic model comprising ARHGAP26, IL33 and BST2 through bioinformatics analysis. Higher risk scores were directly associated with poorer patient prognosis. Immune infiltration analysis using the CIBERSORT and ESTIMATE algorithms revealed significant differences in immune infiltration between the high-risk and low-risk groups, indicating that autophagy-related risk stratification is closely linked to the tumor immune microenvironment and suggesting that autophagy pathways may regulate tumor immune infiltration. To assess the clinical utility of the model for guiding individualized treatment, we compared drug sensitivity, immunotherapy response potential and immune checkpoint expression between the two groups. Low-risk patients were significantly more sensitive to cisplatin, paclitaxel, oxaliplatin, and other first-line chemotherapeutic and targeted agents than high-risk patients, indicating a higher likelihood of benefit from conventional therapy. Across the CTLA-4 and PD-1 subgroups, low-risk patients consistently had higher checkpoint blockade response scores, suggesting greater potential benefit from anti-PD-1/CTLA-4 immunotherapy. Immune gene expression profiling clarified this difference at the molecular level: core immune checkpoint genes, including CD274 (PD-L1), PDCD1, CTLA4, LAG3 and TIGIT, were generally upregulated in the low-risk group, reflecting greater immune infiltration and a higher abundance of checkpoint molecules in the low-risk microenvironment—features that likely underpin its sensitivity to immunotherapy. Taken together, low-risk patients exhibit both high sensitivity to chemotherapy and targeted agents and a strong response to immunotherapy, whereas high-risk patients respond poorly to conventional treatment, show low checkpoint expression and derive limited benefit from immunotherapy. These findings suggest that the autophagy-related risk model incorporating ARHGAP26, IL33 and BST2 may support clinical stratification to guide treatment decisions.

Subsequently, we analyzed the relationship between the expression of three genes (ARHGAP26, IL33 and BST2) and the survival prognosis of breast cancer patients. The results indicate that these genes are closely associated with cancer initiation and progression. ARHGAP26 is located on chromosome 5q31.3 and encodes a protein belonging to the Rho GTPase-activating protein family, involved in key cellular processes including cytoskeletal organization, cell motility, and signal transduction[17]. ARHGAP26 has been studied more extensively in tumors such as gastric cancer, and the CLDN18-ARHGAP26 fusion

gene has been identified in diffuse-type gastric cancer. This fusion results in the loss of Rho GTPase regulatory function of ARHGAP26, thereby activating the RHOA signaling pathway and inducing malignant transformation of cancer cells[18, 19]. In this study, RNA and protein expression levels of ARHGAP26 were lower in breast cancer tissues than in normal breast tissue. It is therefore plausible that ARHGAP26 may contribute to breast cancer development through analogous regulatory mechanisms. However, specific studies investigating mutations, fusions, or aberrant expression of ARHGAP26 in breast cancer remain limited, representing an area that warrants further investigation.

IL-33 is not a classical tumor suppressor gene but rather an immune modulator, and exhibits context-dependent antitumor activity in breast cancer. It can suppress tumor growth and metastasis by activating $CD8^+$ T and NK cell-mediated Th1 immune responses, recruiting eosinophils, and remodeling the tumor microenvironment toward immunosurveillance[20, 21]. IL-33 exerts dual roles in tumors: its expression in tumor cells promotes anti-tumor immune responses, whereas stromal expression may facilitate immunosuppression[22]. Xin Gao et al. reported that tumor-derived IL-33 suppresses tumor growth through $CD8^+$ T cells and NK cells and remodels the tumor microenvironment [23]. The present study found that high IL-33 expression correlates with better survival in breast cancer patients, suggesting an anti-tumor function of IL-33 in this context, which aligns with earlier reports.

BST2 (bone marrow stromal antigen 2, also known as Tetherin) is a transmembrane protein localized at the cell membrane[24]. Although BST2 is not a canonical autophagy gene, it can bidirectionally modulate immune activity within the breast cancer TIME through selective autophagy, regulation of immune molecule expression, and remodeling of immune cell infiltration patterns, ultimately linking these effects to immune evasion, therapeutic sensitivity, and prognosis[25]. Its influence varies by breast cancer subtype and is more pronounced in triple-negative breast cancer (TNBC). In this study, we analyzed BST2 expression levels in normal breast tissue and breast cancer (including TNBC), finding that BST2 expression was higher in both breast cancer and TNBC than in normal breast tissue. However, survival curve analysis showed that BST2 expression was not statistically significantly associated with survival rate among breast cancer patients overall. In contrast, TNBC patients with high BST2 expression had a higher survival rate. These findings suggest a closer link between BST2 expression and TNBC, supporting its relevance for breast cancer subtyping and the selection of subtype-based treatment regimens.

This study has several limitations. Our findings are based on retrospective public datasets and have not been validated in vivo or in vitro. In future work, we plan to perform in vitro assays using breast cancer cell lines and to collect pathological specimens from breast cancer patients for further validation. Despite these limitations, this study provides a useful direction for future investigation.

5. Conclusions

This study reveals the potential value of a prognostic model based on three autophagy-related genes—ARHGAP26, IL33, and BST2—in breast cancer. The model serves as a robust tool for assessing patient

survival and may aid in guiding the selection of individualized pharmacological treatment. This work advances the understanding of autophagy in breast cancer progression and therapy, and lays a foundation for future translational studies.

Declarations

Ethical approval and accordance

Not applicable

Data availability statement

The TCGA-BRCA dataset and corresponding clinical information used in this study were obtained from the University of California, Santa Cruz (UCSC) Xena database (https://xenabrowser.net/datapages/). Additionally, the GSE1456 and GSE305078 datasets analyzed in this work were retrieved from the Gene Expression Omnibus (GEO) database and are accessible via the following links:
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE1456
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE305078.

Conflict of interest

The authors declare no relevant conflict of interest.

Consent to Publish declaration

Not applicable.

Consent to Participate declaration

Not applicable.

Funding

This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author Contributions

Shuangyan Lin and Fang Peng conceived and designed the research framework of this study. Mingyuan Zhao conducted data collection and bioinformatics analysis, generated all figures, and drafted the original manuscript. Qilong Zhang assisted in public dataset analysis and participated in manuscript writing. Luoning Zheng and Jing Li contributed to manuscript layout and proofreading. All authors have reviewed and approved the final version of the manuscript.

Acknowledgments

Not applicable.

References

1. Waks AG, Winer EP. Breast Cancer Treat Jama, 2019. 321(3).
2. Bray F et al. *Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.* CA: A Cancer Journal for Clinicians, 2024. 74(3): pp. 229–63.
3. Hu Y et al. The treatment landscape of triple-negative breast cancer. Med Oncol, 2024. 41(10).
4. Manna PR et al. Hormonal and Genetic Regulatory Events in Breast Cancer and Its Therapeutics: Importance of the Steroidogenic Acute Regulatory Protein. Biomedicines, 2022. 10(6).
5. Cardoso F et al. Correction to: 3rd ESO–ESMO International Consensus Guidelines for Advanced Breast Cancer (ABC 3). Ann Oncol, 2017. 28(12).
6. Swain SM, Shastry M, Hamilton E. Targeting HER2-positive breast cancer: advances and future directions. Nat Rev Drug Discovery. 2022;22(2):101–26.
7. Vessoni AT, et al. Autophagy and genomic integrity. Cell Death Differ. 2013;20(11):1444–54.
8. Onorati AV, et al. Targeting autophagy in cancer. Cancer. 2018;124(16):3307–18.
9. Chu Y, et al. LUBAC and OTULIN regulate autophagy initiation and maturation by mediating the linear ubiquitination and the stabilization of ATG13. Autophagy. 2020;17(7):1684–99.
10. Janji B, Berchem G, Chouaib S. Targeting Autophagy in the Tumor Microenvironment: New Challenges and Opportunities for Regulating Tumor Immunity. Frontiers in Immunology; 2018. p. 9.
11. Noman MZ, et al. Blocking Hypoxia-Induced Autophagy in Tumors Restores Cytotoxic T-Cell Activity and Promotes Regression. Cancer Res. 2011;71(18):5976–86.
12. Xia H et al. Autophagic adaptation to oxidative stress alters peritoneal residential macrophage survival and ovarian cancer metastasis. JCI Insight, 2020. 5(18).
13. Xia H, Green DR, Zou W. Autophagy in tumour immunity and therapy. Nat Rev Cancer. 2021;21(5):281–97.
14. Han Y, et al. Role of autophagy in breast cancer and breast cancer stem cells (Review). International Journal of Oncology; 2018.
15. Liu H, et al. Integrating the hallmarks of cancer into autophagy: a perspective from underlying mechanisms to therapeutic strategies. Theranostics. 2026;16(12):6948–79.
16. Yoshihara K et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun, 2013. 4(1).
17. Sadok A, Marshall CJ. Rho GTPases Small GTPases, 2014. 5(4).
18. Yao F, et al. Recurrent Fusion Genes in Gastric Cancer: CLDN18-ARHGAP26 Induces Loss of Epithelial Integrity. Cell Rep. 2015;12(2):272–85.
19. Zhang F, et al. Recurrent RhoGAP gene fusion CLDN18-ARHGAP26 promotes RHOA activation and focal adhesion kinase and YAP-TEAD signalling in diffuse gastric cancer. Gut. 2024;73(8):1280–91.

20. Arrizabalaga L et al. Unveiling the multifaceted antitumor effects of interleukin 33. Front Immunol, 2024. 15.
21. Saglimbeni J, et al. Peritumoral macrophages recruit eosinophils to promote antitumor immune responses in breast cancer. Proc Natl Acad Sci U S A. 2025;122(38):e2504645122.
22. Hollande C, et al. Inhibition of the dipeptidyl peptidase DPP4 (CD26) reveals IL-33-dependent eosinophil-mediated control of tumor growth. Nat Immunol. 2019;20(3):257–64.
23. Gao X, et al. Tumoral Expression of IL-33 Inhibits Tumor Growth and Modifies the Tumor Microenvironment through CD8 + T and NK Cells. J Immunol. 2015;194(1):438–45.
24. Erikson E et al. *In vivo expression profile of the antiviral restriction factor and tumor-targeting antigen CD317/BST-2/HM1.24/tetherin in humans.* Proceedings of the National Academy of Sciences, 2011. 108(33): pp. 13688–13693.
25. Kim C, Choi S, Park J-W. Emerging Oncogenic and Immunoregulatory Roles of BST2 in Human Cancers. Biomedicines, 2026. 14(1).

Figures

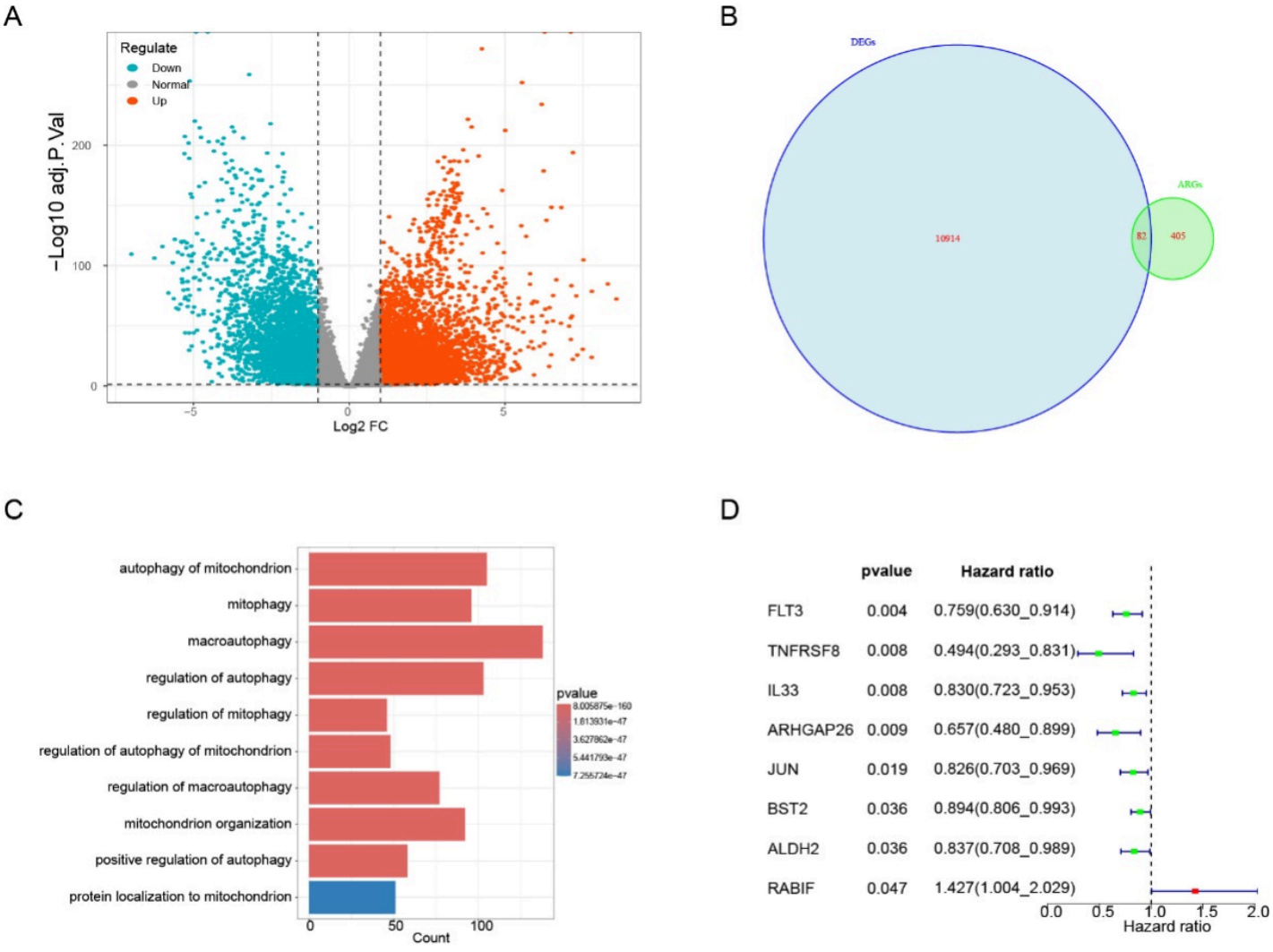


Figure 1

Autophagy-related gene candidate identification in TCGA-BRCA.

(A) Volcano plot of differential expression (tumor vs. normal). (B) Venn diagram of intersection genes between DEGs and ARGs. (C) KEGG enrichment analysis of DEGs. (D) Univariate Cox regression for prognosis-associated candidates.

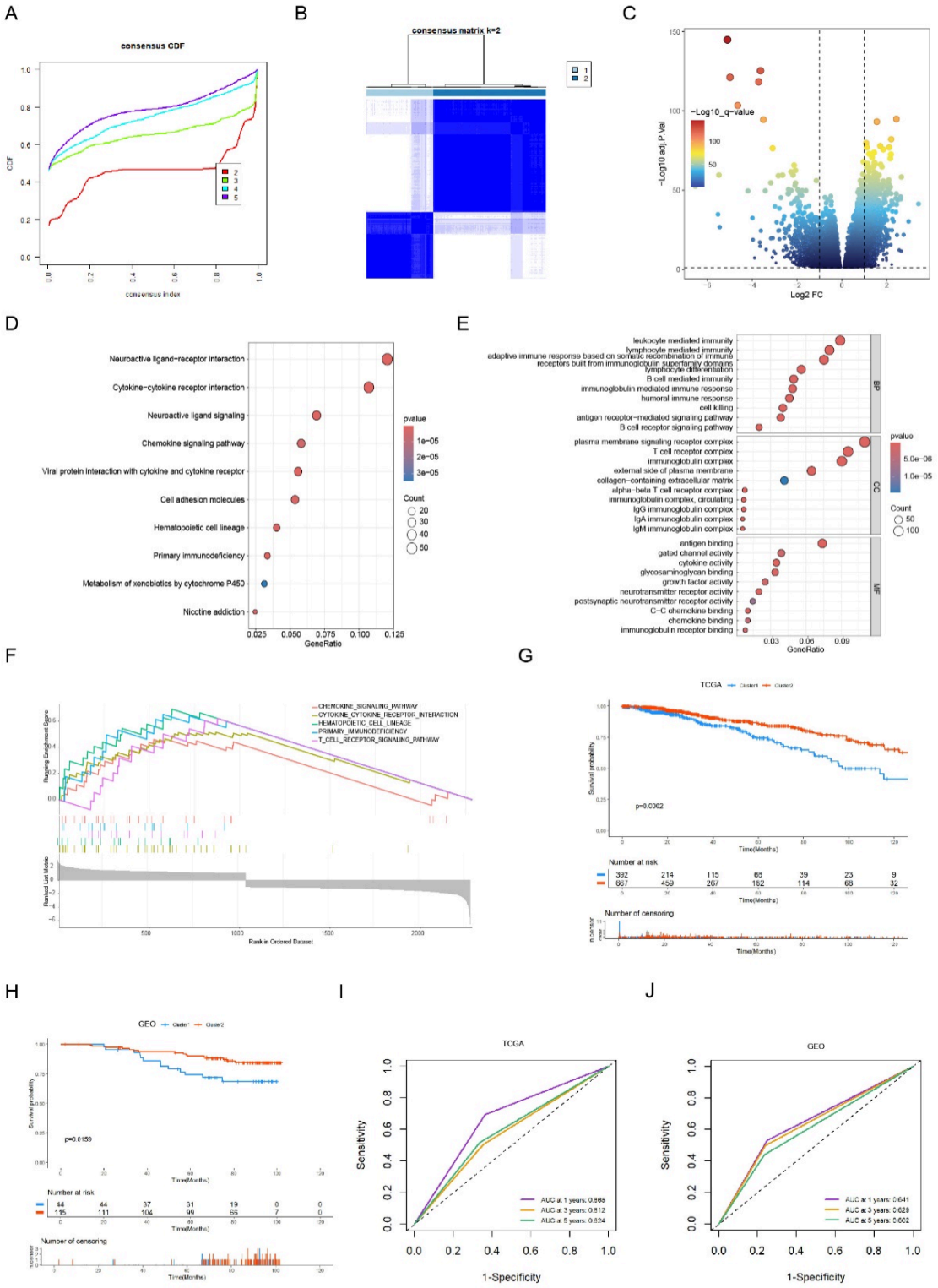


Figure 2

Consensus clustering and transcriptomic differences in TCGA-BRCA.(A) CDF curves and Δ area under the CDF curve for k = 2-5. (B) Patient assignment at the selected k=2. (C) Volcano plot of differential expression (Cluster 1 vs. Xluster2). (D) KEGG analysis. (E) GO analysis. (F) GSEA between Cluster 1 and Cluster 2. (G) Kaplan–Meier survival curves for the two clusters in the TCGA and (H) GSE1456 cohorts. (I-J) Time-dependent ROC curves at 1-, 3-, and 5- year AUC values for the TCGA (I) and GSE1456 (J) cohorts.

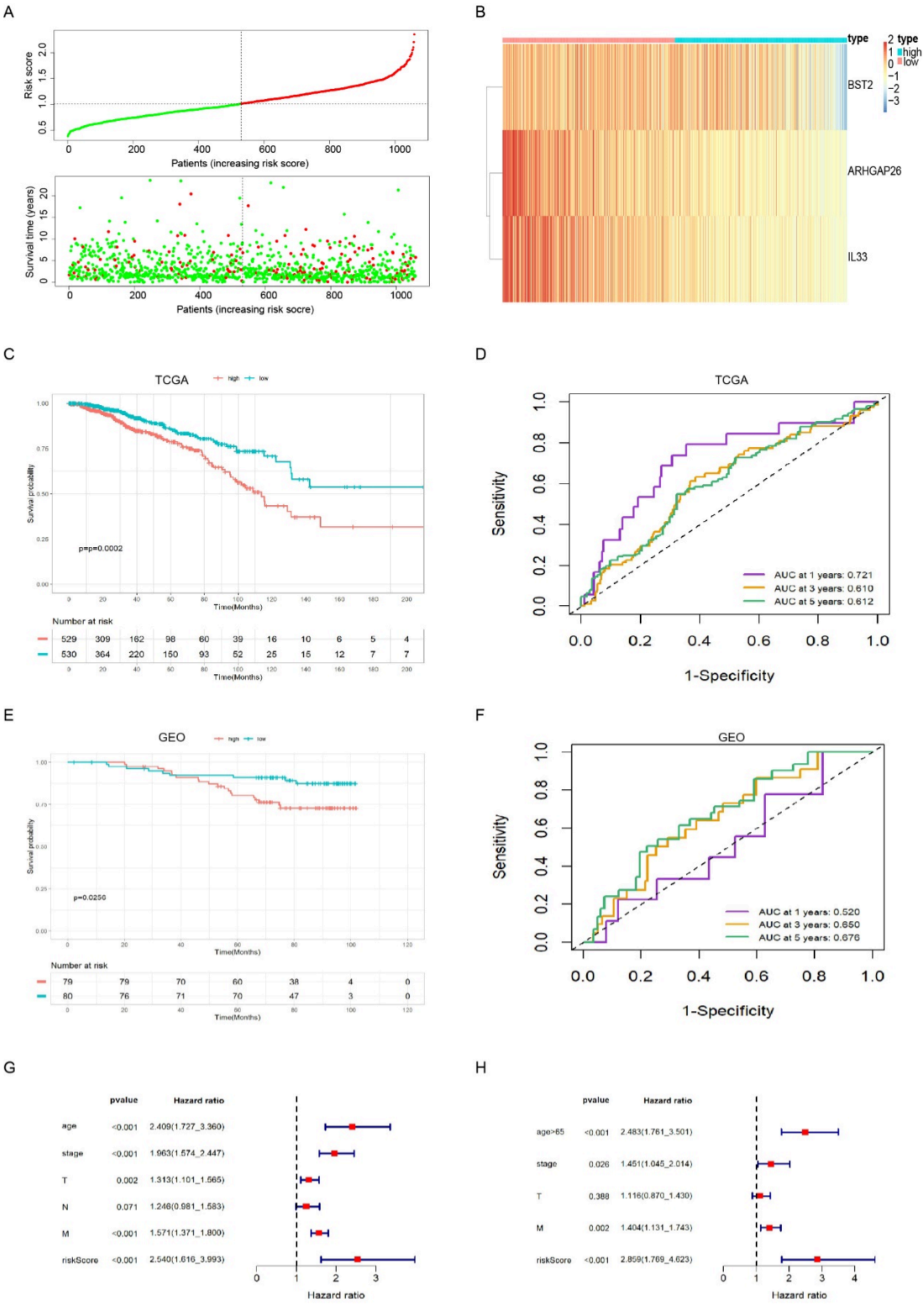


Figure 3

Development and external validation of a three-gene prognostic signature.

(A) Risk score distribution and patient survival status in the TCGA cohort. (B) Expression heatmap of BST2, ARHGAP26, and IL33. (C) Kaplan–Meier curves for the TCGA cohort (log-rank p = 0.0002) and (D) the GEO cohort (log-rank p = 0.0256). (E) Time-dependent ROC curves with 1-, 3-, and 5-year AUC values

for the TCGA cohort and (F) for the GEO cohort. (G) Univariate and (H) multivariate Cox regression forest plots for clinical factors and risk score.

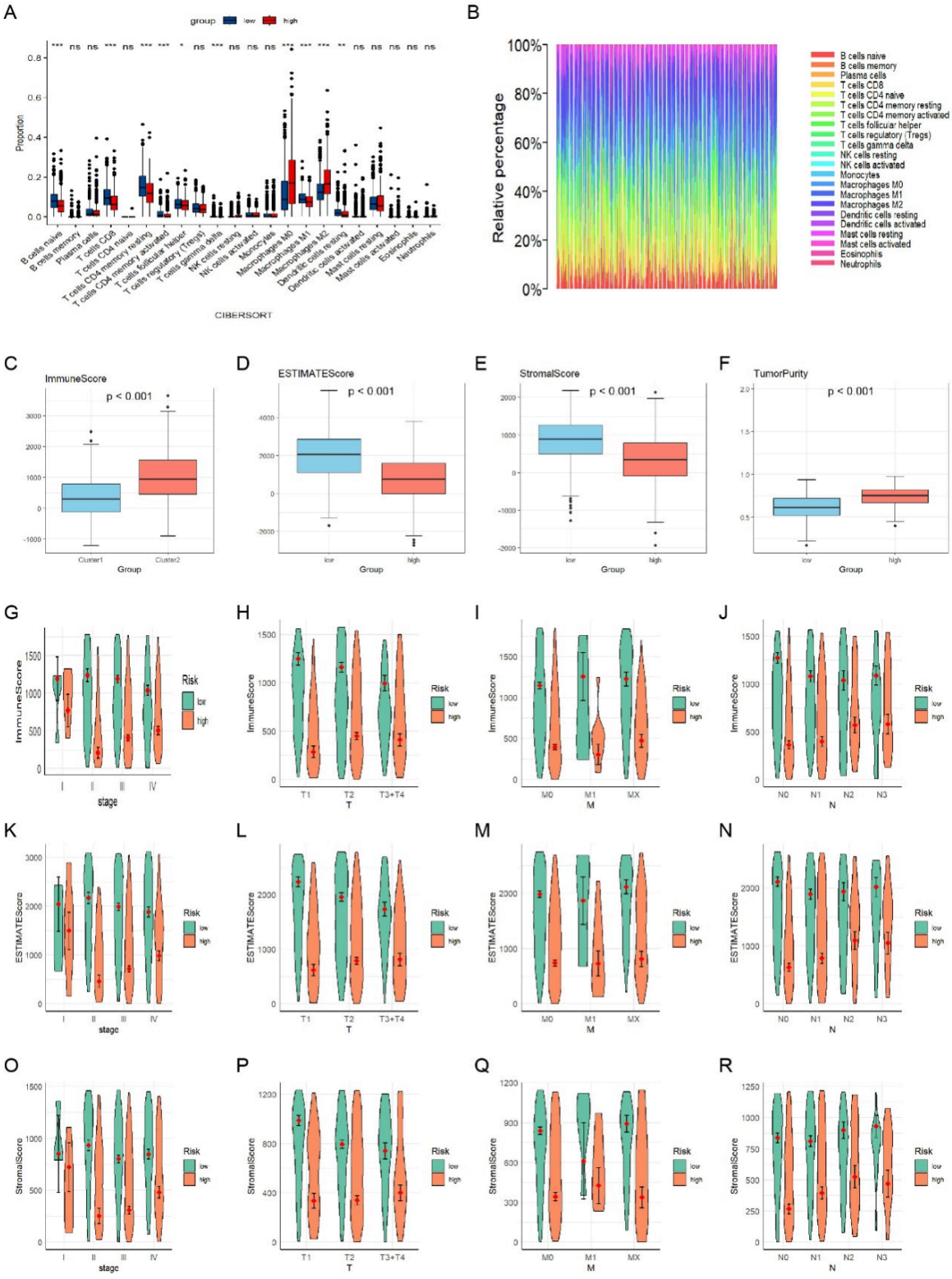


Figure 4

Association of the risk score with tumor microenvironment features. (A) Comparison of 22 immune cell types between high- and low-risk groups. (B) Stacked bar chart of 22 immune cell proportions across

samples. (C-F) ImmuneScore, ESTIMATEScore, StromalScore, and tumor purity in high- vs. low-risk groups. (G-J) ImmuneScore stratified by pathologic stage (G), T stage (H), M stage (I), and N stage (J). (K–N) ESTIMATE Score stratified by pathologic stage (K), T stage (L), M stage (M), and N stage (N). (O–R) StromalScore stratified by pathologic stage (O), T stage (P), M stage (Q), and N stage (R).

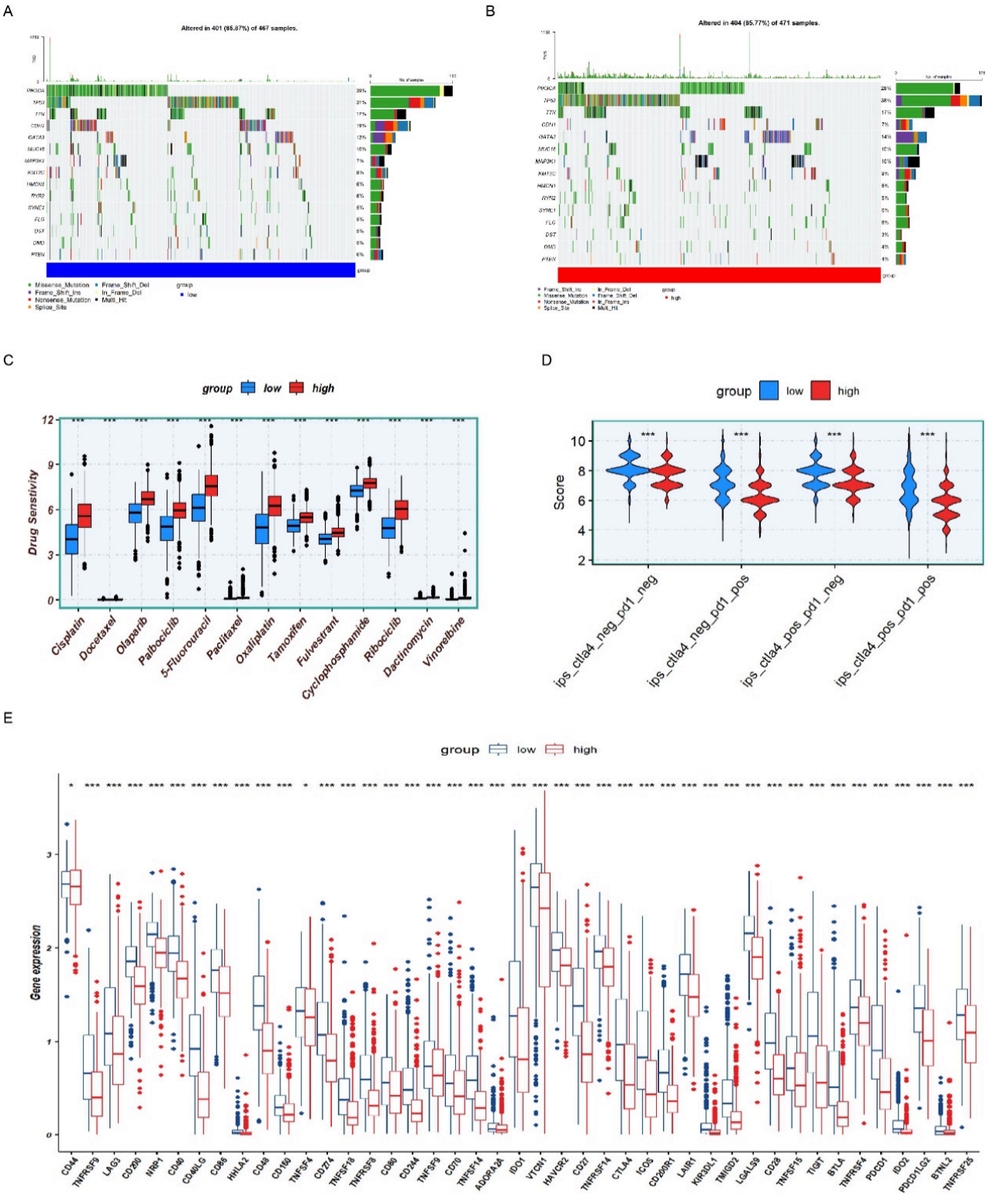


Figure 5

Genomic mutation landscape and chemosensitivity differences between risk subgroups. (A-B) Waterfall plots of the top 15 mutated genes in the high-risk (A) and low-risk (B) groups. (C) Association of the risk score with drug sensitivity. (D) Immunophenoscore (IPS) comparison between the two risk groups. (E) Expression levels of common immune checkpoint genes in the high- versus low-risk groups.

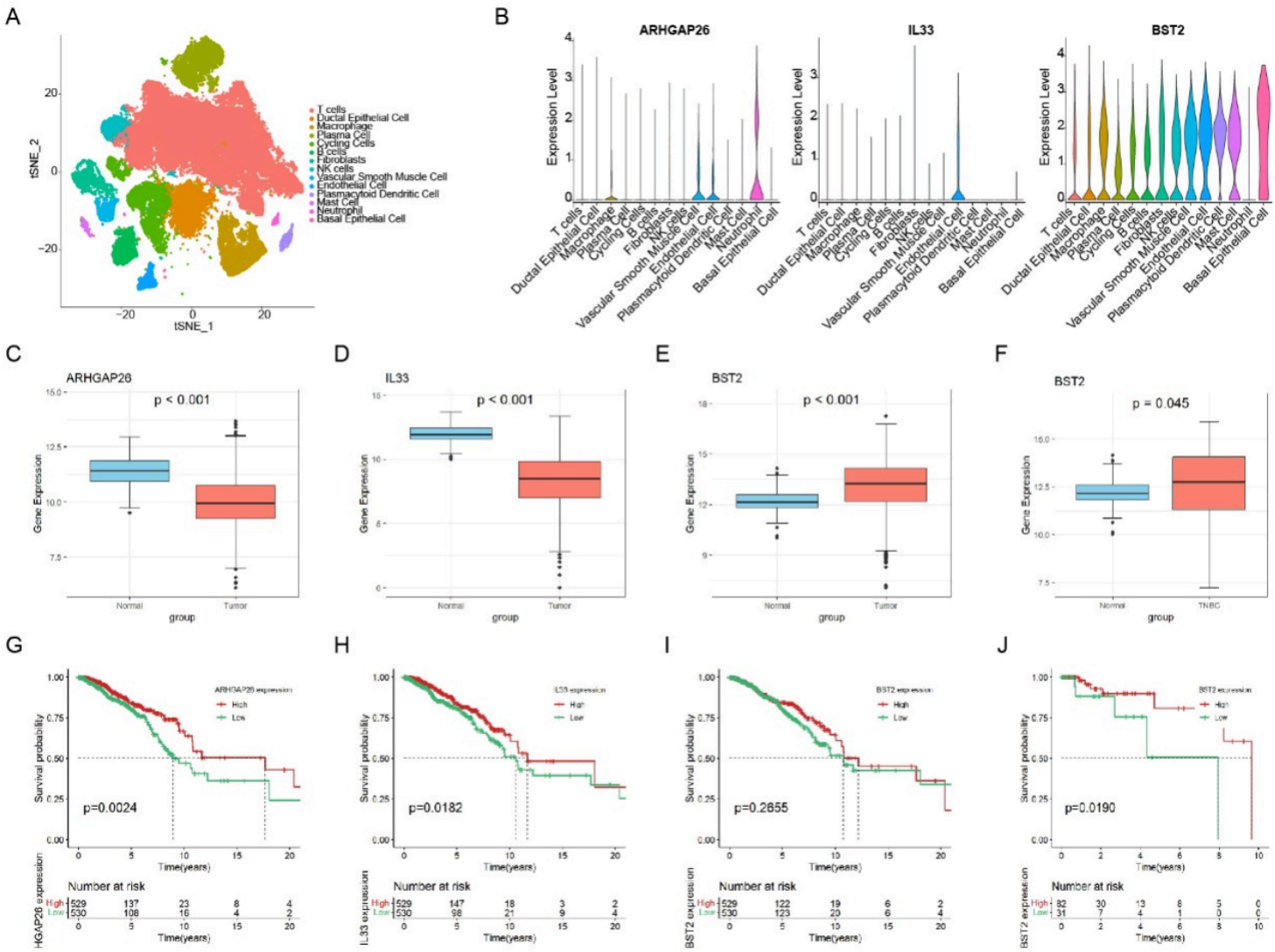


Figure 6

Single-cell expression profiles and prognostic significance of ARHGAP26, IL33, and BST2 in breast cancer. (A) t-SNE visualization of diverse cell subsets within the tumor microenvironment. (B) Violin plots showing the expression levels of ARHGAP26, IL33, and BST2 across distinct cell populations. (C–E) Box-and-whisker plots comparing mRNA expression of ARHGAP26 (C), IL33 (D), and BST2 (E) between normal breast and breast cancer tissues. (F) Box-and-whisker plot comparing BST2 expression between normal breast and triple-negative breast cancer (TNBC) tissues. (G–I) Kaplan-Meier survival analyses for breast cancer patients stratified by high versus low expression of ARHGAP26 (G), IL33 (H), and BST2 (I). (J) Kaplan-Meier survival analysis for TNBC patients.

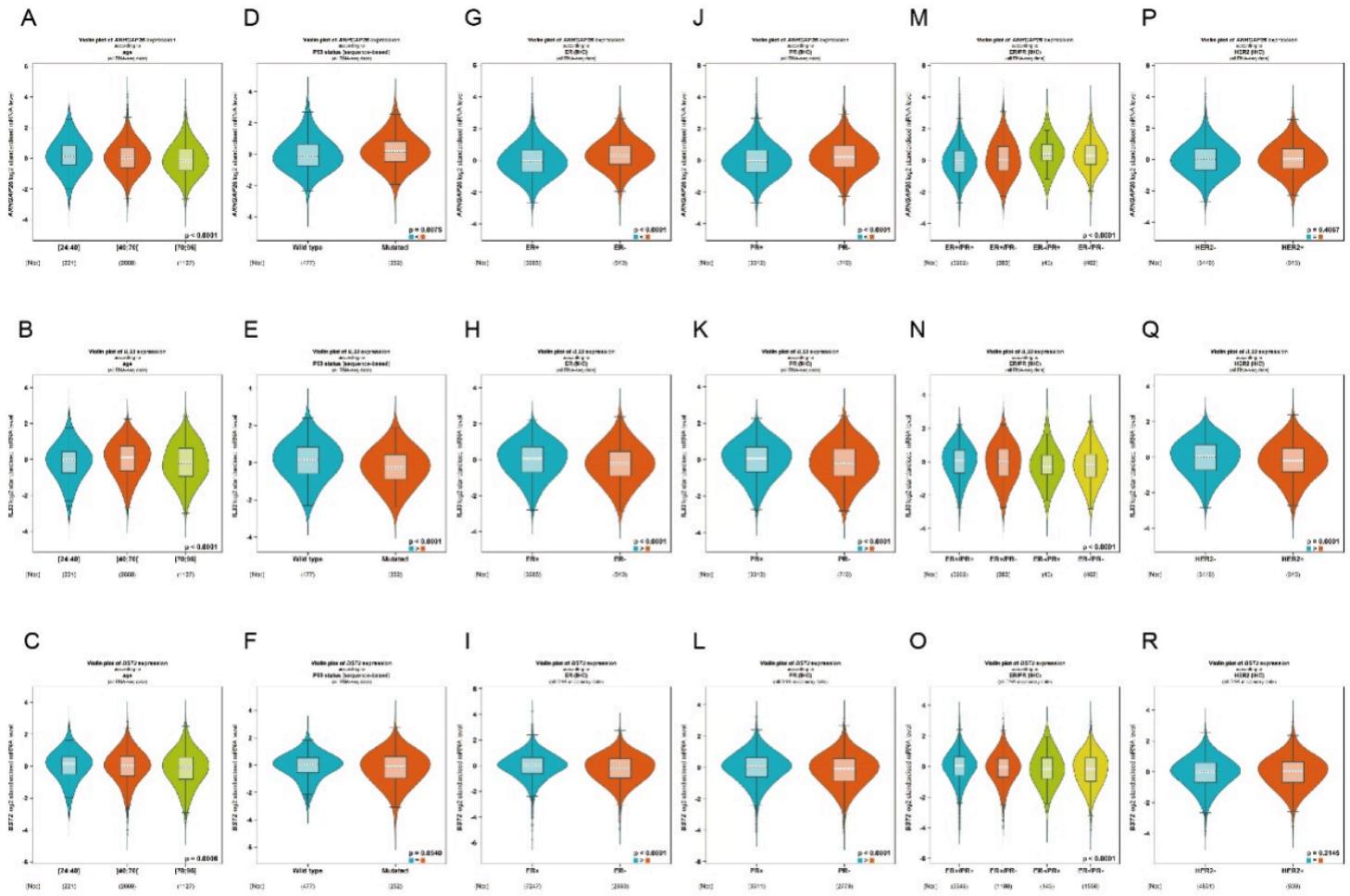


Figure 7

The target gene's expression correlates with key clinical characteristics in breast cancer patients. (A–C) Association of ARHGAP26 (A), IL33 (B), and BST2 (C) expression with patient age. (D–F) Association of ARHGAP26 (D), IL33 (E), and BST2 (F) expression with P53 mutation status. (G–I) Association of ARHGAP26 (G), IL33 (H), and BST2 (I) expression with ER status. (J–L) Association of ARHGAP26 (J), IL33 (K), and BST2 (L) expression with PR status. (M–O) Association of ARHGAP26 (M), IL33 (N), and BST2 (O) expression with combined ER/PR status. (P–R) Association of ARHGAP26 (P), IL33 (Q), and BST2 (R) expression with HER2 status.

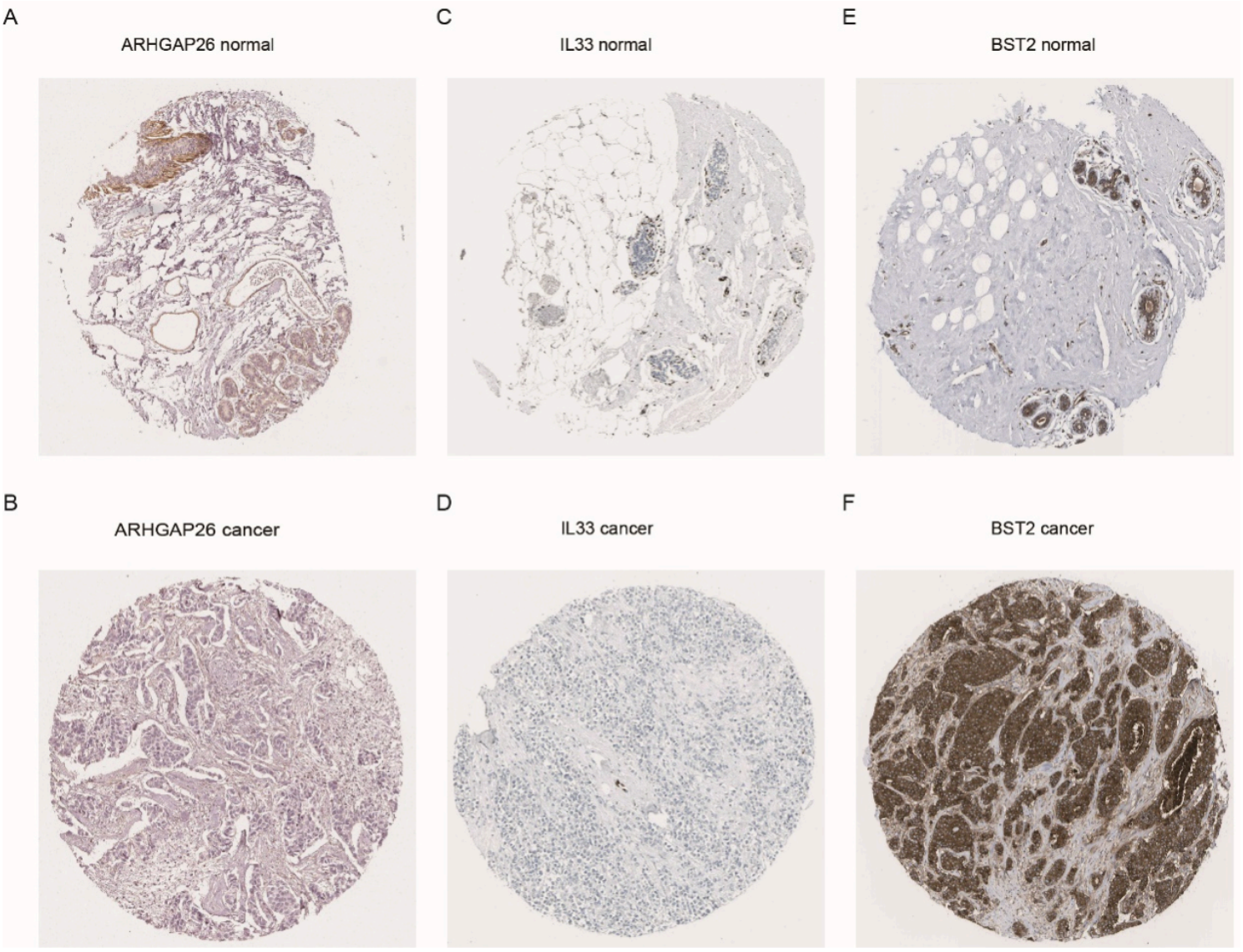


Figure 8

Immunohistochemical staining. (A–F) Immunohistochemical staining of ARHGAP26 (A, B), IL33 (C, D), and BST2 (E, F) in normal versus breast cancer tissues.