Proceedings of ICMMGH 2026 Symposium: Cell and Molecular Biology: Understanding Immune Cell Functions
DOI: 10.54254/2753-8818/2026.PJ31308

Unraveling Benign and Malignant Tumor Heterogeneity Through Spatial Transcriptomics

Haosheng Lyu

South China Agricultural University, Guangzhou, China
rock18963070162@163.com

Abstract. Breast cancer is one of the most common malignant tumors among women worldwide, and its high heterogeneity poses significant challenges for early diagnosis and precise treatment. Despite the continuous development of molecular typing and targeted therapy, how to precisely distinguish the benign and malignant regions of tumors at the spatial level and analyze their molecular characteristics remains an urgent problem to be solved. This study is based on spatial transcriptomics technology on the 10x Genomics Visium platform. Systematic transcriptional analyses were conducted on human breast cancer tissue sections. Through differential expression analysis of DCIS/LCIS and IDC regions, a total of 1,143 significantly differentially expressed genes were identified, among which genes such as S100A6, STC2, and TFF3 were significantly upregulated in breast cancer progression, suggesting a key role in malignant transformation of tumors. Gene ontology enrichment analysis revealed that the differentially expressed genes were significantly aggregated in key functional modules such as metabolic pathways, MAPK signaling pathways, cell cycle and cancer-related pathways. in particular, the significant enrichment of the Rap1 signaling pathway and Pathways in cancer revealed the synergistic activation mechanism of multi-level biological processes such as cell adhesion, migration, and matrix degradation during the transformation of breast cancer from benign to malignant. In summary, this study revealed the molecular heterogeneity in the benign and malignant regions of breast cancer and the spatial distribution characteristics of key signaling pathways at the tissue level through spatial transcriptomics, providing new theoretical basis and spatial omics evidence for spatial molecular typing of breast cancer, targeted drug development, and optimization of individualized treatment strategies.

Keywords: Breast cancer, spatial transcriptomics, tumor heterogeneity, spatial domain

1. Introduction

Breast cancer is one of the most common malignant tumors among women worldwide, seriously threatening women's lives and health [1]. According to statistics, there were approximately 2.3 million new cases and 670,000 deaths of breast cancer worldwide in 2022 [2], topping the list of deaths related to malignant tumors among women. The World Health Organization predicts that new cases and deaths from breast cancer will increase by 38% and 68% respectively by 2050. The

DOI: 10.54254/2753-8818/2026.PJ31308

situation is not optimistic. In China, there is also a high incidence of breast cancer, with more than 400,000 new cases each year, and it is characterized by younger onset.

Breast cancer research and clinical practice have undergone a profound transformation over the past few years, evolving from traditional "extensive" treatment to "ultra-precision" individualized medical care. In this process, the study of the tumor immune microenvironment has deepened, revealing new mechanisms of immune escape in breast cancer and providing a theoretical basis for the optimization of immunotherapy strategies. At the same time, the discovery and application of targeted molecular markers (such as hormone receptors, HER2 protein, etc.) have greatly promoted the refinement of treatment modalities. However, there is significant heterogeneity within breast cancer, with significant differences in treatment response and prognosis among different patients, different types, and even different regions of the same tumor. Clinical data show that the five-year survival rate for patients with early-stage breast cancer can be over 90%, while for those with advanced or highly malignant breast cancer, it is less than 30%. Therefore, how to accurately distinguish between benign and malignant breast tumors and identify molecular markers that can be used for prognosis assessment and typing diagnosis has become a key issue in clinical diagnosis and treatment as well as basic research.

Although traditional genetic testing and histological methods provide important information in the diagnosis and classification of breast cancer, due to the lack of spatial resolution, mechanical tissue sections or cell dissociation are often required, resulting in the destruction of the original spatial structure of the tissue and making it difficult to reveal the precise distribution of benign and malignant regions within the tumor and their complex heterogeneous patterns. In recent years, the development of spatial transcriptomics has provided a revolutionary solution to this problem. This technique enables high-throughput capture and quantitative analysis of gene expression in tissue sections while maintaining the integrity of the tissue's spatial structure, and closely integrates gene-level expression information with histological context, thus not only precisely positioning the spatial boundaries of benign and malignant tumor regions at the molecular level, It can also fully depict the transcriptional characteristics and functional states of cell populations in different regions and reveal the interactions among various cell types in the tumor microenvironment, including immune cells, fibroblasts and vascular endothelial cells. Compared with traditional methods, spatial transcriptomics integrates morphological and molecular information on pathological sections, allowing researchers to visually observe the distribution of gene expression and avoid the bias caused by relying solely on morphological judgment. Studies have shown that the technique can not only clearly distinguish between benign and malignant regions in breast cancer research, but also further reveal its underlying molecular mechanisms and functional pathways, thus providing new ideas and possibilities for finding new molecular markers, improving diagnostic methods, and promoting individualized treatment.

Based on this, spatial transcriptome technology was used in this study to systematically analyze the transcriptional information characteristics of benign and malignant regions of breast tumors. By combining the advantages of histopathology and spatial transcriptomics, we expect to more accurately define the benign and malignant regions of tumors, deeply compare the differences in their expression profiles and functional states, and then screen out potential molecular markers to provide new ideas and theoretical basis for the early diagnosis, molecular typing and individualized treatment of breast cancer.

DOI: 10.54254/2753-8818/2026.PJ31308

2. Spatial transcriptomics

With the rapid development of high-throughput sequencing technology, transcriptomics has been widely applied in tumor research. The earliest Bulk RNA sequencing technology [3], by extracting and sequencing the overall RNA of tissues, provides a comprehensive average information on gene expression levels, providing an important theoretical basis for molecular typing, signaling pathway studies, and molecular mechanism exploration of tumors such as breast cancer. Bulk sequencing technology has advantages such as high throughput, low cost and stable data, and has played an irreplaceable role in the early stage of tumor research. However, because the sequencing results are mixed signals from all cell populations within the tissue, it is impossible to distinguish the expression differences between different cell types. This "averaging effect" masks the heterogeneity characteristics between cells, especially in highly heterogeneous tumors such as breast cancer, where it is often difficult to identify the molecular characteristics of a few key cell populations.

To overcome the shortcomings of Bulk sequencing, scRNA-seq technology emerged and rapidly developed into a powerful tool for analyzing the heterogeneity of tumor cells [4]. scRNA-seq can obtain gene expression profiles at the single-cell level, helping researchers reveal the transcriptional characteristics, functional states of different cell subpopulations and the presence of rare cell populations. The technology has greatly advanced the understanding of tumor heterogeneity, tumorigenesis and development mechanisms, and the immune microenvironment in breast cancer research. Single-cell sequencing, however, often relies on complete tissue dissociation, a process that not only disrupts the spatial positioning relationship of cells but may also introduce technical biases, resulting in the loss of some cell subpopulations. Due to the lack of spatial background information, scRNA-seq has difficulty reflecting the complex tissue structure and intercellular interactions within tumors, limiting its application in spatial biology research.

In this context, the emergence of spatial transcriptomics techniques represents a new stage in the development of transcriptomics [5]. The technology is capable of integrating gene expression profiles with histological morphology while maintaining the integrity of tissue spatial structure, combined with high-throughput transcriptome sequencing. Compared with Bulk sequencing, ST avoids the "averaging" effect and can precisely analyze the expression characteristics of different regions on two-dimensional pathological sections; Compared with single-cell sequencing, ST does not require complete dissociation of the tissue and thus can preserve the original spatial position of the cells in the tissue, thereby revealing the interactions between cell populations and their functional roles in the tumor microenvironment. Studies have shown that spatial transcriptomics in breast cancer research can not only clearly define the spatial distribution of benign and malignant regions, but also reveal the differences in transcriptional characteristics and functional pathways, providing new possibilities for identifying molecular markers, analyzing tumor heterogeneity, and optimizing individualized treatment. Therefore, spatial transcriptomics, as a cutting-edge high-resolution technique, compensates for the shortcomings of traditional sequencing methods and highlights the advancement and necessity of the experimental data used in this study.

3. Dataset

The data used in this study originated from the processing of Space Ranger count on 10x Genomics' Visium platform [6]. The specific samples were taken from human breast cancer tissue numbered V1- breast cancer blocking part -1 (Block A, segment 1), which has a complex pathological composition Covering ductal carcinoma in situ (DCIS), lobular carcinoma in situ (LCIS), and invasive ductal carcinoma, A variety of types including IDC. The space size of the entire tissue

DOI: 10.54254/2753-8818/2026.PJ31308

section is 3 inches, with a large coverage and representativeness. The original sequencing data, provided directly by 10x Genomics, has undergone standardized quality control and processing, and is highly reliable and consistent. Breast cancer is a typical heterogeneous tumor, and its complexity is mainly reflected in both intratumoral heterogeneity and intratumoral heterogeneity, which can be further subdivided into spatial heterogeneity and temporal heterogeneity. in this context, this study chose in situ transcriptome capture technology, which not only enables high-resolution detection of gene expression in local tissue regions, but also combines non-in situ sequencing results to obtain global transcriptional information, this enables a comprehensive analysis and unbiased statistics of the transcriptome in tumor samples.

In terms of data preprocessing and quality control, the overall quality of the sequencing data obtained in this study was relatively high. The total number of capture spots covering the tissue was 3,813, and the total number of detected genes was 22,968. The statistics showed that the true expression rate of the original gene molecule was as high as 99.9%, and the proportion of Q30 bases reached 96.7%, indicating that the accuracy of base identification was extremely high, and the sequencing depth was sufficient, the barcode identification error rate was extremely low, and the overall data quality was fully in line with the requirements for subsequent analysis. In the partition analysis of pathological types, we systematically screened and annotated different lesion regions based on the spatial expression patterns of the spatial transcriptome. Among them, 435 capture points were identified in the DCIS region, 1,928 capture points were identified in the IDC region, and the remaining capture points were mainly distributed in benign or background tissue regions. The high-quality spatial transcriptome data laid a solid foundation for subsequent differential expression analysis and tumor heterogeneity studies.

4. Differential gene analysis

To gain a deeper understanding of the molecular differences between different pathological regions of breast cancer, this study conducted a systematic analysis of differentially expressed genes in transcriptome data of DCIS/LCIS and IDC. The screening of differentially expressed genes not only reveals the differences in expression between benign and malignant regions at the molecular level, but also provides important clues for the discovery of potential molecular markers and clinical classification.

Before conducting the differential analysis, we first carried out strict quality control and gene filtering on the original spatial transcriptome expression matrix. The specific steps included: eliminating genes that were not expressed in all cells to avoid interference from invalid data; Further remove low-abundance genes expressed in less than 5% of cells to reduce the impact of random noise; Then calculate the coefficient of variation based on the mean and variance of the genes, and retain the top 10,000 genes with the highest expression fluctuations for subsequent analysis. This multi-level filtering strategy ensures high reliability of the input data and makes the variance analysis more statistically robust.

The difference analysis employed the DEsingle statistical framework commonly used in single-cell and spatial transcriptomics, which can model transcriptome data with a zero-expansion distribution to accurately identify significant differences in gene expression between different regions. We used the DCIS and IDC regions as two comparison groups and selected 1,143 key differenced genes with a false positive detection rate of less than 0.05 as the significance threshold. Among them, several key genes are known to be associated with the occurrence and development of breast cancer. For example, S100A6 belongs to the S100 calcium-binding protein family, which is widely involved in processes such as cell cycle regulation, proliferation and migration. Studies have

DOI: 10.54254/2753-8818/2026.PJ31308

shown that S100A6 is highly expressed in a variety of malignancies, and its overactivation is closely associated with cytoskeletal remodeling and enhanced invasiveness of tumor cells. In this study, S100A6 was significantly upregulated in the IDC region, suggesting that it may play a facilitating role in the progression of breast cancer from in situ to invasive [7]. STC2 (Stanniocalcin 2) is a glycoprotein hormone with the function of regulating cellular metabolism and stress response [8]. Multiple studies have found that STC2 is often used as a molecular marker of poor prognosis in breast cancer, associated with the adaptability of tumor cells to hypoxic environments and their ability to invade and metastasize. In our analysis, STC2 was significantly highly expressed in the IDC region and may promote tumor cell survival and spread by activating hypoxia-associated signaling pathways. TFF3 (Trefoil Factor 3) is a member of the trefoil factor family that is primarily involved in epithelial repair and mucosal barrier maintenance [9]. In oncology studies, TFF3 is often associated with the proliferation and anti-apoptotic ability of tumor cells and can promote the survival and drug resistance of breast cancer cells by activating signaling pathways such as PI3K/AKT. In this study, TFF3 showed differential expression in both DCIS and IDC regions, suggesting that it may play an important role in both early breast cancer lesions and malignant progression.

5. Results

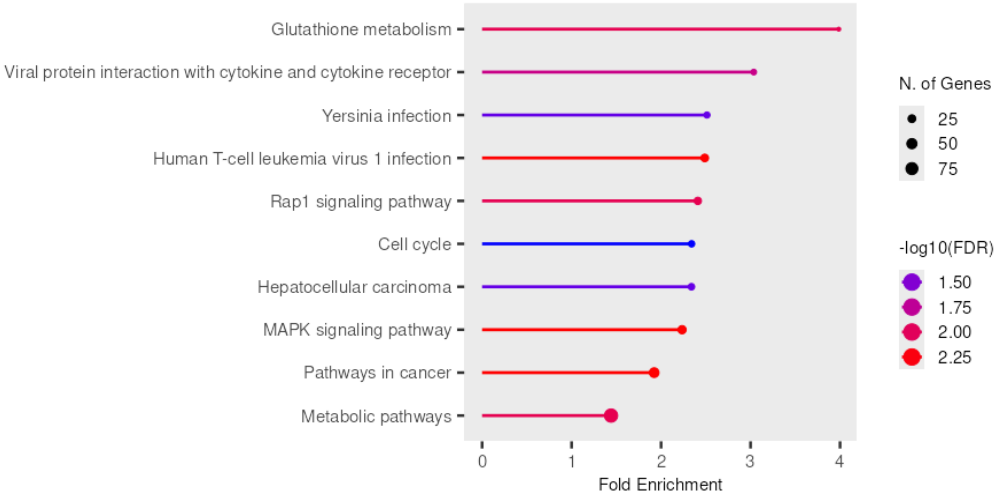

Figure 1. GO results for 1143 genes

Subsequently, we imported these differentially expressed genes into the ShinyGO platform for Gene Ontology (GO) enrichment analysis to systematically reveal their potential biological functions and pathway distributions. The analysis results showed that the differentially expressed genes were significantly enriched in multiple pathways closely related to tumorigenesis and development (as shown in the figure). Among them, Metabolic pathways, MAPK signaling pathway, Cell cycle, Rap1 signaling pathway and Pathways in cancer and others showed significant enrichment, suggesting that these pathways may play important roles in cellular functional remodeling and signal transduction in different tumor types. In addition, Glutathione metabolism and Viral protein interaction with cytokine and cytokine receptor Significant enrichment of immune-related processes suggests a potential interaction between tumor cell metabolism and the immune microenvironment.

DOI: 10.54254/2753-8818/2026.PJ31308

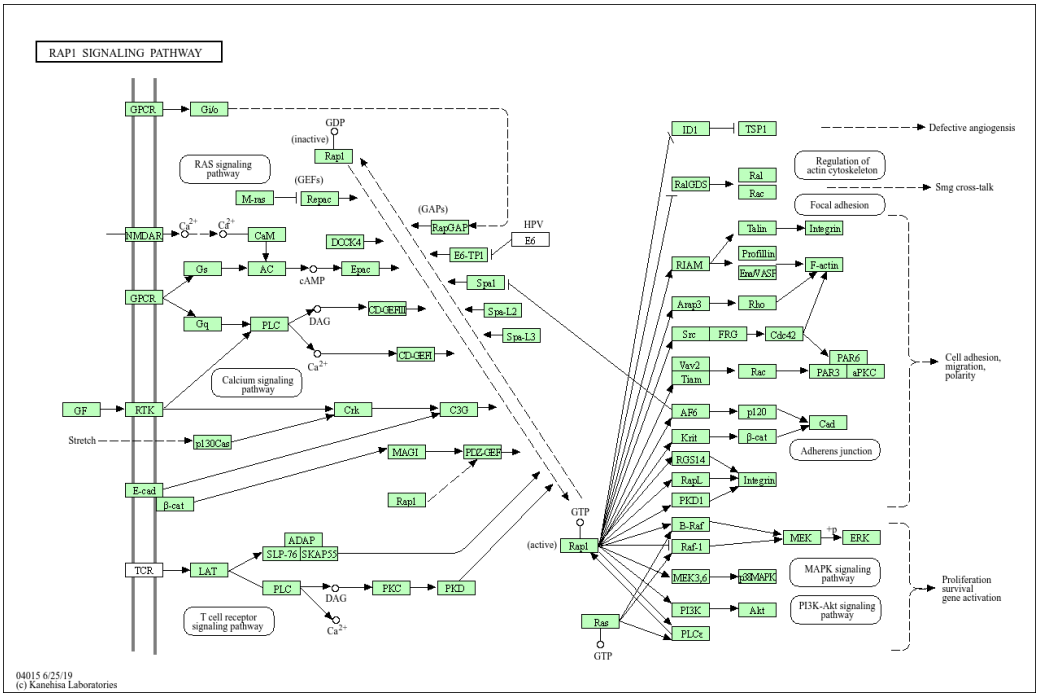

Figure 2. Rap1 signaling pathway

After GO enrichment analysis of differentially expressed genes, we found that the Rap1 signaling pathway was significantly enriched in the process of benign and malignant transformation of breast cancer, suggesting that it may play a key role in tumor progression [10]. Rap1 is a small GTPase, a member of the Ras family, which mainly affects the invasion and metastasis of tumor cells by regulating processes such as cell adhesion, polarization, migration, and cytoskeleton remodeling. Studies have shown that Rap1 plays a core regulatory role in the acquisition of migration ability in tumor cells by regulating cell adhesion to the extracellular matrix through integrin and cadherin signaling. In breast cancer, abnormal activation of Rap1 signaling can promote integrin β1 activation through the JAM-A/AF-6/PDZ-GEF2 complex, enhancing cell adhesion and motility; Meanwhile, SIPA1 downstream, as a GTP-activating protein of Rap1 (RapGAP), promotes extracellular matrix degradation and tumor cell invasion by regulating the integrin β1/FAK/Akt/MMP9 signaling axis [11]. In addition, SIPA1 can directly activate the transcription of the integrin β1 gene in the nucleus, thereby further enhancing the invasion and metastasis potential of breast cancer cells. Overall, the Rap1 signaling pathway plays a bridging and amplifying role in the transformation of breast cancer from benign to malignant by multi-level regulation of cell adhesion, cytoskeleton remodeling and matrix degradation, providing important biological evidence for subsequent exploration of tumor spatial heterogeneity and molecular progression mechanisms.

DOI: 10.54254/2753-8818/2026.PJ31308

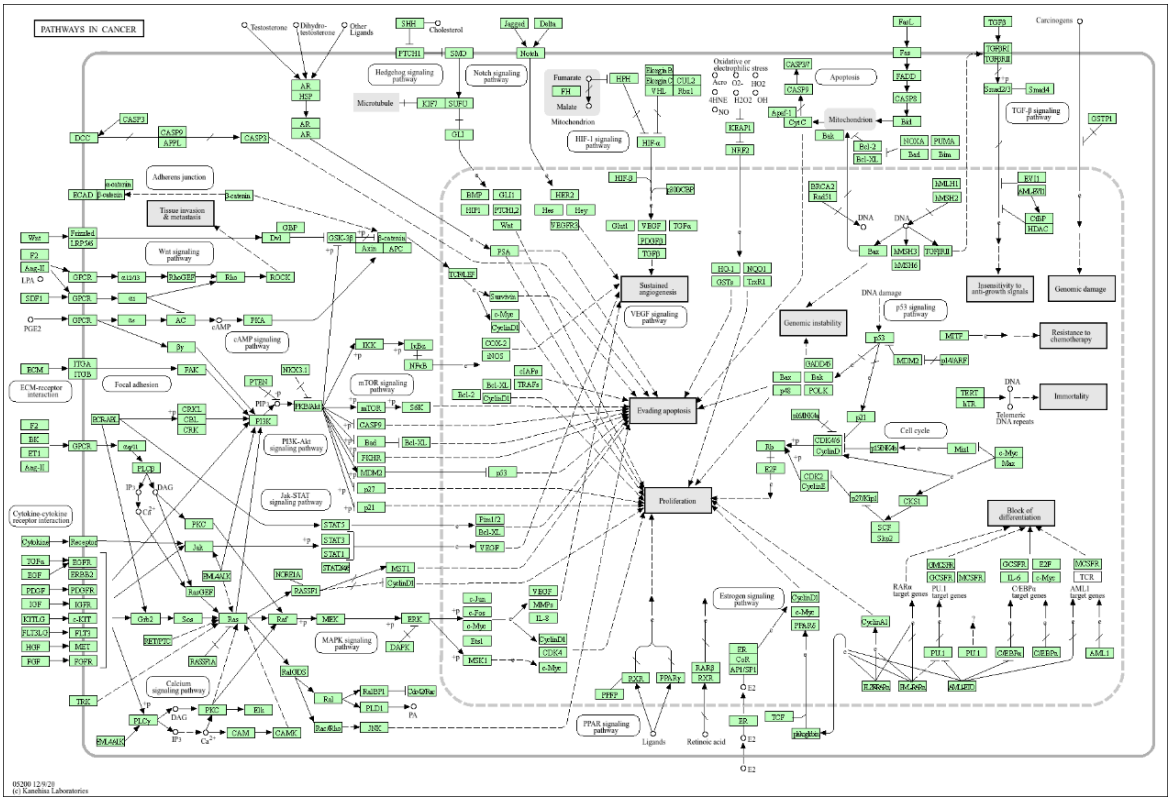

Figure 3. Pathways in cancer

In our GO enrichment analysis, "Pathways in cancer" were significantly enriched, suggesting that multiple classical tumor-associated signaling networks may be synergistic in the transformation of breast cancer from benign to malignant [12,13]. The so-called "Pathways in cancer" often include a collection of multiple core oncogenic pathways (such as PI3K/Akt, MAPK, Wnt/β-catenin, TGF-β, Notch, cell cycle regulation, apoptosis regulation, etc.) It is the "upper-level" pathway category that the high-throughput functional enrichment platform integrates multiple signaling pathways together. In breast cancer, the PI3K/Akt pathway is often abnormally activated to promote the survival and proliferation of tumor cells while inhibiting apoptosis. Studies suggest that upstream (such as growth factor receptors, ERBB family receptors) or downstream (such as mTOR, Bad, FOXO, GSK3β) nodes of the PI3K/Akt pathway in breast cancer may be regulated, allowing cancer cells to evade intracellular apoptotic stress.

6. Discussion and conclusion

Based on spatial transcriptomics techniques, this study systematically analyzed the differences in molecular expression characteristics and functional pathways between benign and malignant regions of breast cancer. By analyzing differentially expressed genes in high-quality spatial transcriptome data of breast cancer tissues, we screened out a total of 1,143 significantly differentially expressed genes and revealed the potential role of multiple key pathways in the transformation of benign and malignant breast cancer through GO enrichment analysis. Among them, the significant enrichment of the Rap1 signaling pathway and the Pathways in cancer pathway suggests that multiple classical oncogenic signaling networks cooperatively drive key biological events such as cell adhesion, migration, invasion, and metabolic remodeling during the transformation of breast cancer from orthotopic lesion to invasive cancer. Further analysis revealed that the Rap1 pathway regulates

DOI: 10.54254/2753-8818/2026.PJ31308

cytoskeleton and matrix degradation through the integrin-FAK /Akt/MMP9 signaling axis, while Pathways in cancer integrates a network of multi-pathway synergisms such as PI3K/Akt, MAPK, Wnt/β-catenin, TGF-β, etc. Revealing the multi-dimensional regulatory mechanism of malignant progression of breast cancer.

The results of this study provide new spatial transcriptional evidence for the molecular stratification of benign and malignant regions of breast cancer. Compared with traditional transcriptome analysis methods, spatial transcriptome techniques not only retain the spatial location information of tissue structure and tumor microenvironment, but also reveal the transcriptional characteristics and interactions of different regional cell populations at the molecular level, thereby analyzing the biological mechanisms of tumor heterogeneity and regional specificity at the spatial scale. This research strategy holds significant clinical translational potential and application prospects in the future. First, in terms of early diagnosis and risk stratification, the spatial transcriptome can precisely determine the benign and malignant regions of tumors by identifying the spatial expression patterns of characteristic genes and signaling pathways. For example, by combining the spatial activation features of Rap1 signaling and the PI3K/Akt pathway, it can assist in the precise distinction between carcinoma in situ (DCIS) and early invasive carcinoma (IDC) in pathological diagnosis, making up for the deficiency of traditional pathology in identifying areas with blurred boundaries. At the same time, by establishing a risk stratification model based on spatial transcriptional features, it is expected to be used for screening early high-risk populations and provide molecular evidence for early intervention of breast cancer. Secondly, in terms of drug target screening and precision treatment, spatial transcriptomics can reveal the activation status of specific signaling pathways in different regions, thereby helping to identify highly active driving molecular networks. For example, the abnormal upregulation of Rap1 signaling can enhance the activity of the integrin-FAK /Akt/MMP9 axis, promoting cell invasion and matrix degradation; Therefore, specific small molecule inhibitors or RNA interference strategies targeting Rap1 or its downstream effector factors, such as FAK or MMP9, can be designed to spatially block the invasion and spread of tumor cells. Likewise, the spatial co-activation of PI3K/Akt and MAPK signaling in the Pathways in cancer pathway suggests a dual-target combined blocking strategy, such as the combination of PI3K inhibitors and MEK inhibitors, to enhance therapeutic effects and overcome resistance to single-target therapy. In the future, by integrating spatial omics data with drug response databases, individualized drug response prediction and precise treatment plan design can also be achieved. Again, in terms of predicting recurrence and metastasis risk, spatial transcriptomics can identify cell subpopulations with high invasive potential and their spatial distribution characteristics. By analyzing the signaling networks characterized by these cell populations (such as Rap1 and the co-activation of Wnt/β-catenin), molecular spatial models can be established to predict the likelihood of local tumor recurrence or distant metastasis. In addition, long-term tracking of spatial expression patterns helps monitor tumor residue or the regeneration of drug-resistant clones after treatment, providing new approaches for dynamic assessment of treatment response and recurrence risk.

Overall, this study, through spatial dimension gene expression analysis, reveals the molecular mechanisms and signaling network basis of benign and malignant transformation of breast cancer, providing a new perspective for understanding tumor heterogeneity. Future research can be further expanded in the following directions: First, by combining single-cell sequencing with spatial multi-omics technology to achieve a fine correspondence between cell types and spatial functional states; The second is to integrate spatial transcriptome results with clinicopathological features, immunohistochemical and prognostic data to construct spatial molecular diagnostic models that can be used for clinical applications; Third, verify the specific role of key pathways in tumor progression

DOI: 10.54254/2753-8818/2026.PJ31308

through functional experiments and animal models to provide a solid theoretical and experimental basis for the early diagnosis, prognosis assessment and targeted therapy of breast cancer.

References

[1] A. N. Giaquinto et al., “Breast cancer statistics 2024, ” CA. Cancer J. Clin., vol. 74, no. 6, pp. 477–495, Nov. 2024, doi: 10.3322/caac.21863.
[2] A. N. Giaquinto et al., “Breast Cancer Statistics, 2022, ” CA. Cancer J. Clin., vol. 72, no. 6, pp. 524–541, Nov. 2022, doi: 10.3322/caac.21754.
[3] F. Ozsolak and P. M. Milos, “RNA sequencing: advances, challenges and opportunities, ” Nat. Rev. Genet., vol. 12, no. 2, pp. 87–98, Feb. 2011, doi: 10.1038/nrg2934.
[4] Z. Zhang et al., “Advancements in the Application of scRNA-Seq in Breast Research: A Review, ” Int. J. Mol. Sci., vol. 25, no. 24, p. 13706, Jan. 2024, doi: 10.3390/ijms252413706.
[5] L. Moses and L. Pachter, “Museum of spatial transcriptomics, ” Nat. Methods, vol. 19, no. 5, pp. 534–546, May 2022, doi: 10.1038/s41592-022-01409-2.
[6] J. Hu et al., “SpaGCN: Integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network, ” Nat. Methods, vol. 18, no. 11, pp. 1342–1351, Nov. 2021, doi: 10.1038/s41592-021-01255-8.
[7] M. Sadeghi et al., “Integrative Analysis of Breast Cancer Cells Reveals an Epithelial-Mesenchymal Transition Role in Adaptation to Acidic Microenvironment, ” Front. Oncol., vol. 10, p. 304, Mar. 2020, doi: 10.3389/fonc.2020.00304.
[8] X. Niu, Y. Zhan, S. Zhang, Z. Liu, and C. Qu, “Research progress of STC2 in breast cancer, ” Biophys. Rep., vol. 7, no. 3, pp. 185–192, June 2021, doi: 10.52601/bpr.2021.210002.
[9] A. R. H. Ahmed, A. B. Griffiths, M. T. Tilby, B. R. Westley, and F. E. B. May, “TFF3 Is a Normal Breast Epithelial Protein and Is Associated with Differentiated Phenotype in Early Breast Cancer but Predisposes to Invasion and Metastasis in Advanced Disease, ” Am. J. Pathol., vol. 180, no. 3, pp. 904–916, Mar. 2012, doi: 10.1016/j.ajpath.2011.11.022.
[10] Y.-L. Zhang, R.-C. Wang, K. Cheng, B. Z. Ring, and L. Su, “Roles of Rap1 signaling in tumor cell migration and invasion, ” Cancer Biol. Med., vol. 14, no. 1, pp. 90–99, Feb. 2017, doi: 10.20892/j.issn.2095-3941.2016.0086.
[11] S. Shah et al., “Downregulation of Rap1Gap: A Switch from DCIS to Invasive Breast Carcinoma via ERK/MAPK Activation, ” Neoplasia N. Y. N, vol. 20, no. 9, pp. 951–963, Sept. 2018, doi: 10.1016/j.neo.2018.07.002.
[12] D. Song et al., “Pathway-based analysis of breast cancer, ” Am. J. Transl. Res., vol. 6, no. 3, pp. 302–311, May 2014.
[13] E. K. Konstantinou, A. Gioxari, M. Dimitriou, G. I. Panoutsopoulos, and A. A. Panagiotopoulos, “Molecular Pathways of Genistein Activity in Breast Cancer Cells, ” Int. J. Mol. Sci., vol. 25, no. 10, p. 5556, Jan. 2024, doi: 10.3390/ijms25105556.