

Classification of triple-negative breast cancer based on immune cell infiltration

Wenji Shi

Shanghai Jiaotong University School of Medicine Xinhua Hospital

Dan Wang

People's Hospital of Rong Chang

Luz Angela Torres-de la Roche

Carl von Ossietzky Universitat Oldenburg

Rui Zhuo

Guilin Traditional Chinese Medicine Hospital

Rudy Leon De Wilde (✉ rudy-leon.dewilde@pius-hospital.de)

Carl von Ossietzky Universitat Oldenburg <https://orcid.org/0000-0001-5759-9745>

Research article

Keywords: Triple-negative breast cancer, Immune infiltration, Targeted therapy

Posted Date: July 30th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-43664/v1>

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

[Read Full License](#)

Abstract

Background

Although the classification system of triple-negative breast cancer (TNBC) has become more and more perfect, there is no report on the immune subtype of triple-negative breast cancer based on immune cell infiltration.

Results

According to immune infiltrating cells, data from 360 patients were divided into cluster A (subtype 1 and subtype 3) and cluster B (subtype 2; with poorly immune phenotype). Expression of memory B cells, naïve B cells, M0 macrophages, M1 macrophages, CD4 memory activated T cells, and CD4 naïve T cells were significantly higher in cluster A ($P < 0.05$). In contrast, the expression of M2 macrophages and resting mast cells were higher in cluster B ($P < 0.05$). GSEA results show that B cell receptor pathway and JAK-STAT pathway are activated and more frequently altered in cluster A ($P < 0.05$). mTOR pathway alterations usually appear in cluster B ($P < 0.05$). Compared with cluster A, the risk of recurrence in cluster B patients is significantly increased ($P < 0.05$).

Conclusions

This analysis of tumor microenvironment revealed the multifaceted nature of TNBC and its impact on patient prognosis, being recurrence more often in those with poorly immune phenotype. These results provide a reference for further exploration of the heterogeneity of TNBC.

Background

Despite the advances in screening and diagnosis, breast cancer remains a leading cause of death in women worldwide, especially triple-negative breast cancer (TNBC), which accounts for 12–17% of all breast cancers [1, 2]. Because the tumor itself lacks specific receptors (ER, PR, and HER2), it is unlikely to respond to the traditional targeted therapies. Therefore, finding a valid effective tumor target is critical for the treatment of triple-negative breast cancer [3, 4]. Unfortunately, although a part of TNBC-related targeting drugs has been discovered, the clinical benefits of patients are very different due to the heterogeneity of this disease [5]. In 2010 Perou et al. [6] identified the claudin-low molecular subtype of TNBC, based on gene expression profiling, and in 2011 Lehmann et al. [7] further reported that TNBC could be divided into six subtypes including two basal-like (BL1 and BL2), an immunomodulatory (IM), a mesenchymal (M), a mesenchymal stem-like (MSL), and a luminal androgen receptor (LAR) subtype. The expression profiles of these six subtypes are significantly different and each subtype is associated with different signal pathway. Both studies mentioned above were based on the transcriptome data for TNBC classification. In recent years, with extensive research on the genomic and biological changes of cancer

cells, cumulative evidence indicates that breast tumor immune infiltration is closely associated with the clinical prognosis of the patients [8]. Therefore, based on the proportion of immune cells in the tumor microenvironment, we used bioinformatics methods to identify TNBC subgroups with different prognosis and try to explore the heterogeneity of this disease.

Results

Immunophenotyping and prognosis

We used the CIBERSORT algorithm to infer the proportion of the 22 immune infiltrating cells in the tissue, and used K-means to cluster the immune infiltrating cells of the enrolled patients. We found that when the clustering number κ value was 3, classifier model shows well stability. Therefore, in this study, $\kappa = 3$ was used to classify the 360 TNBC sample data into three subtypes (**Fig. 1A**). When we plotted the Kaplan-Meier curve for this three subtypes of TNBC patients, there was no significant difference between three subtypes ($P = 0.11$), but careful observation revealed that the curve of subtype 1 and subtype 3 were almost coincide, and the curve of subtype 2 was far away from then, so we tried to merge subtypes 1 and 3 into a new subtype and named cluster A, while subtype 2 was redefined as cluster B. We evaluated the recurrence-free survival of patients between cluster A and B. The results showed that patients in cluster B had significantly lower recurrence-free survival compared to cluster A. Log-rank analysis showed a statistically significant difference in recurrence-free survival between the two groups ($P = 0.015$) (**Fig. 1B**). This result suggests that the TNBC subtype has a predictive effect, with subtype 3 showing the better prognosis.

TME cell expression and aberrant pathway activations

To verify the differential expression of the 22 immune cells in the different clusters, we used R package ggplot2 to visually analysis the differences in the expression matrices. The box plot results showed that the expression of memory B cells, naïve B cells, M0 macrophages, M1 macrophages, CD4 memory activated T cells, and CD4 naïve T cells were significantly higher in cluster A than in cluster B, and the difference in expression was significant ($P < 0.05$). In contrast, the expression of M2 macrophages and resting mast cells were significantly higher in cluster B than in cluster A. ($P < 0.05$) (**Fig. 2A**). In addition, GSVA results showed that the following GSVA signatures were upregulated in Cluster A, including the B cell receptor pathway and JAK-STAT pathway ($P < 0.05$). Tumors with poorly immune phenotype (Cluster B) tended to have frequent alterations in the mTOR pathway ($P < 0.05$) (**Fig. 2B**). GSEA results also support that the B cell receptor pathway and JAK-STAT pathway are activated in cluster A. However, mTOR pathway was not significantly enriched in cluster B (**Fig. 2C**).

Discussion

Compared to the other types of breast cancer, TNBC still lacks clear therapeutic targets. Therefore, finding valid and effective therapeutic targets for TNBC has been the focus of numerous and intense research. In

this study, TNBC was subgrouped using unsupervised clustering method to explore the molecular characteristics of the different immune-phenotypes in TNBC, to provide a reference for the realization and development of an effective treatment strategy for TNBC.

Tumor immune infiltration refers to the migration of immune cells from the blood to the tumor tissue, which then begins to exert their function [13]. Attempts to determine the landscape of the immune infiltration cells using bioinformatics methods can provide valuable information on genotypes of genes and cells with similar characteristics and functions. It has been widely used in a variety of tumors. Zeng et al. [14] evaluated the tumor microenvironment infiltration pattern of 1,524 gastric cancer patients, and integrated patient clinical information to define the tumor microenvironment (TME) phenotype of three gastric cancers, providing a new strategy for targeted therapy in gastric cancer. In addition, Xiong et al. [15] used unsupervised clustering to analyze the proportion of colorectal cancer immune cells and identified five tumor subgroups in colorectal cancer and the differences of five subgroups may be important determinants for prognosis and response to treatment. It was a breakthrough result in the treatment of colorectal cancer. As for breast cancer, M2 macrophages, also named tumor-associated macrophage (TAM) was an important part of immune infiltrating cells in breast cancer tissues. Studies have shown that it is involved in the process of breast tumor occurrence, growth, invasion and metastasis, and high expression of TAM is usually associated with poor prognosis of TNBC patients [16, 17]. This conclusion is consistent with our analysis. We found that when compared with the cluster A, the expression of M2 macrophages in cluster B subtype was significantly increased, and patients with high expression of M2 macrophages meant a shorter relapse-free survival time. The above conclusion suggests that the TNBC subtype we tried to explore has potential clinical application value. After patients enter the clinical diagnosis and treatment procedure, we use this classifier to stratify the patient's risk (Low-risk cluster A vs. High-risk cluster B). In consequence, according to the characteristics of cluster B patients overexpressing M2 macrophages, M2 targeted therapy could be selected to treat this group of patients, which might be a meaningful attempt to achieve precise medicine of TNBC patients.

Pathway analysis plays an important role in the study of tumor development. In this study, we used GSVA and GSEA to perform pathway variation analysis on grouped samples to explore pathway variations that may occur in the different subtypes of TNBC. GSVA results showed that the cluster A subtype is significantly enriched in the B cell receptor pathway and JAK-STAT pathway. GSEA enrichment results also revealed this phenomenon. This means that the activation of both two pathways may be an important factor to promote the occurrence and development of cluster A subtype, and are closely related to the biological functions of proliferation, differentiation and apoptosis of breast cancer cells. As for cluster B, it is often accompanied by variations in the mTOR pathway, which is closely related to tumor development, by GSVA analysis. This result implies that in the cluster B subtype misalignment of the mTOR pathway may lead to abnormal proliferation of breast cancer cells, which may affect the prognosis of patients. Although, this result did not appear in the GSEA analysis, this above conclusion had been supported by many other researchers. Zhang et al. [18] confirmed that Aur-A and mTOR are potential therapeutic targets in the TNBC subtype. Ueng et al. [19] evaluated the expression of p-mTOR in 172 TNBCs by immunohistochemistry and reported that in early stage TNBC patients, the overall and

recurrence-free survival was significantly worse in patients with p-mTOR expression compared to patients without p-mTOR expression. This conclusion suggests that a detailed mechanism of action of mTOR pathway in TNBC will be a new breakthrough for immune-targeting therapy. Based on above findings, we could infer that tumor immune infiltration and aberrant pathway activations may be important factors that lead to cluster A and cluster B, but this hypothesis needs further verification.

Conclusion

Our study revealed the multi-faceted features of TNBC, from the perspective of immune infiltration of the tumor cells, by performing TNBC immune-phenotyping based on the tumor microenvironment and provided a reference for further exploring the heterogeneity of TNBC.

Materials And Methods

Obtaining TNBC datasets and preprocessing

We downloaded the clinical information of 360 TNBC patients and the RNA-seq raw data, corresponding to each primary tumor samples, from the sequence read archive database (SRA). The clinical information includes age, relapse-free survival (RFS), and clinical stage.

Construction of immune cell expression matrix

To assess the relative proportion of the 22 types of infiltrating immune cells, we used CIBERSORT (<http://cibersort.stanford.edu/>), a deconvolution-based online tool [9]. This tool enables the use of gene expression data to estimate the cellular composition of complex tissues and to quantify the abundance of immune cells and other specific cell types. In addition, we also used the LM22 characteristic matrix to calculate (after removing the low-abundance immune cells with a median ratio of less than 1%) and perform subsequent analysis only on the remaining immune cells.

K-means clustering for TME-infiltrating cells

K-means clustering is often used in datasets where the category attributes are unclear, and it is desirable to be able to mine or automatically classify objects with similar characteristics through data mining [10]. In this study, to identify the heterogeneity of different individual patients, the patient's tissue subpopulation was initially analyzed by K-means clustering, which was done using the R package `factoextra`.

Pathway enrichment analysis

To understand the cause of the phenotypic differences from a bioinformatics perspective, we used the gene set variation analysis (GSVA) and gene-set enrichment analysis (GSEA) [11, 12]. GSVA is a method of enrichment of gene sets that estimates changes in pathway activity in different samples in an unsupervised manner. Here, we estimated the difference in pathways between the cluster A and cluster B

of TNBC. First, we obtained the reference gene set from the Broad Institute's MSigDB database (software.broadinstitute.org/gsea/msigdb/), and then used the limma package to set the difference standard as $P < 0.05$, the false discovery rate (FDR) was greater than 0.2 to screen the differential genes [13]. Next, the GSVA package was used to carry out the differential pathway enrichment. Finally, GSEA analysis by cluster Profiler R package was used to explore whether the aberrant pathway activations are also consistent with GSVA results.

Statistical analysis

The immune infiltration ratio was calculated using the online tool CIBERSORT. Patients with a P value of < 0.05 were used for subsequent analysis. The Kaplan-Meier survival curves of the TNBC subtype and recurrence-free survival data (RFS) were plotted using the "survival" function, and the differences between the two subtypes of RFS were evaluated using log-rank test analysis. The Wilcoxon rank-sum test was used to detect the expression of different clusters in the 22 immune cells. The GSVA package was used to annotate the possible pathway differences between different clusters. All data were statistically analyzed using R (version 3.6.1) software. $P < 0.05$ was considered statistically significant, and all the statistical tests performed were two-sided.

Abbreviations

TNBC

Triple-negative breast cancer

ER

Estrogen receptor

PR

Progesterone receptor

HER2

Human Epidermal Growth Factor Receptor 2

BL1

Basal-like 1

BL1

Basal-like 2

IM

Immunomodulatory

MSL

Mesenchymal stem-like

LAR

Luminal androgen receptor

SRA

Sequence read archive database

RFS

Relapse-free survival
TME
Tumor microenvironment
GSVA
Gene set variation analysis
GSEA
Gene-set enrichment analysis
FDR
False discovery rate

Declarations

Ethics approval and consent to participate:

Not applicable.

Consent for publication:

Not applicable.

Availability of data and materials:

All of this data could download from public database (<https://www.ncbi.nlm.nih.gov/sra>).

Competing interests:

The authors declare that they have no competing interests.

Funding:

No external funding was involved.

Authors' contributions:

RDW, RZ: Conceptualization. WS and DW: Data curation, formal analysis. WS and DW: writing original study report. WS, DW and LATR: data analysis and writing final manuscript. All authors read and approved the final manuscript.

Acknowledgements :

We thank Professor Zhimin Shao for uploading the original sequencing data in public database. Especially, we sincerely thank Luo Peng, who is an excellent researcher at Zhu Jiang Hospital of Southern Medical University, for his guidance and selfless help in our data processing.

References

1. Kumar P, Aggarwal R. An overview of triple-negative breast cancer. *Arch Gynecol Obstet*. 2016;293:247–69.
2. Cleator S, Heller W, Coombes RC. Triple-negative breast cancer: therapeutic options. *Lancet Oncol*. 2007;8:235–44.
3. Isakoff SJ. Triple-negative breast cancer: role of specific chemotherapy agents. *Cancer J*. 2010;16:53–61.
4. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res*. 2007;13:4429–34.
5. Masuda H, Baggerly KA, Wang Y, Zhang Y, Gonzalez-Angulo AM, Meric-Bernstam F, et al. Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. *Clin Cancer Res*. 2013;19:5533–40.
6. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res*. 2010;12:R68.
7. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest*. 2011;121:2750–67.
8. Fridman WH, Galon J, Dieu-Nosjean MC, Cremer I, Fisson S, Damotte D, et al. Immune infiltration in human cancer: prognostic significance and disease control. *Curr Top Microbiol Immunol*. 2011;344:1–24.
9. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods Mol Biol*. 2018;1711:243–59.
10. Steinley D. K-means clustering: a half-century synthesis. *Br J Math Stat Psychol*. 2006;59:1–34.
11. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.
12. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics*. 2012;16:284–7.
13. Jiang B, Mason J, Jewett A, Liu ML, Chen W, Qian J, et al. Tumor-infiltrating immune cells: triggers for tumor capsule disruption and tumor progression? *Int J Med Sci*. 2013;10:475–97.
14. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, et al. Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures. *Cancer Immunol Res*. 2019;7:737–50.

15. Xiong Y, Wang K, Zhou H, Peng L, You W, Fu Z. Profiles of immune infiltration in colorectal cancer and their clinical significant: A gene expression-based study. *Cancer Med.* 2018;7(9):4496–508.
16. Sica A, Larghi P, Mancino A, Rubino L, Porta C, Totaro MG, et al. Macrophage polarization in tumour progression. *Semin Cancer Biol.* 2008;18:349–55.
17. Moorman AM, Vink R, Heijmans HJ, van der Palen J, Kouwenhoven EA. The prognostic value of tumour-stroma ratio in triple-negative breast cancer. *Eur J Surg Oncol.* 2012;38:307–13.
18. Zhang W, Xia D, Li Z, Zhou T, Chen T, Wu Z, et al. Aurora-A/ERK1/2/mTOR axis promotes tumor progression in triple-negative breast cancer and dual-targeting Aurora-A/mTOR shows synthetic lethality. *Cell Death Dis.* 2019;10:606.
19. Ueng SH, Chen SC, Chang YS, Hsueh S, Lin YC, Chien HP, et al. Phosphorylated mTOR expression correlates with poor outcome in early-stage triple negative breast carcinomas. *Int J Clin Exp Pathol.* 2012;5(8):806–13.

Figures

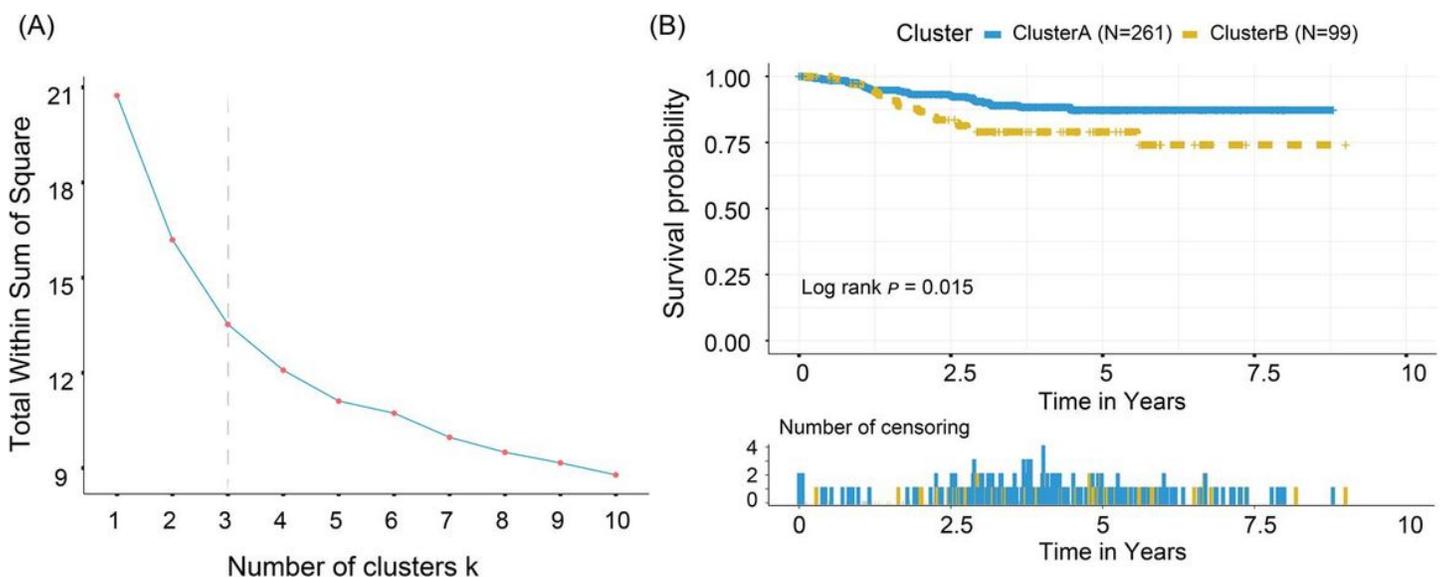


Figure 1

Immune subtypes and prognosis analysis of TNBC. Figure. 1A. Optimal number of clusters, select the most optimal number of clusters where the slope change is not obvious, and determine the best K=3 for this batch of samples for the first time. Figure 1B. The recurrence-free survival curves of cluster A and cluster B, blue for cluster A and yellow for cluster B. It can be seen from the graph that there is a significant difference in postoperative recurrence-free survival curves between the two groups

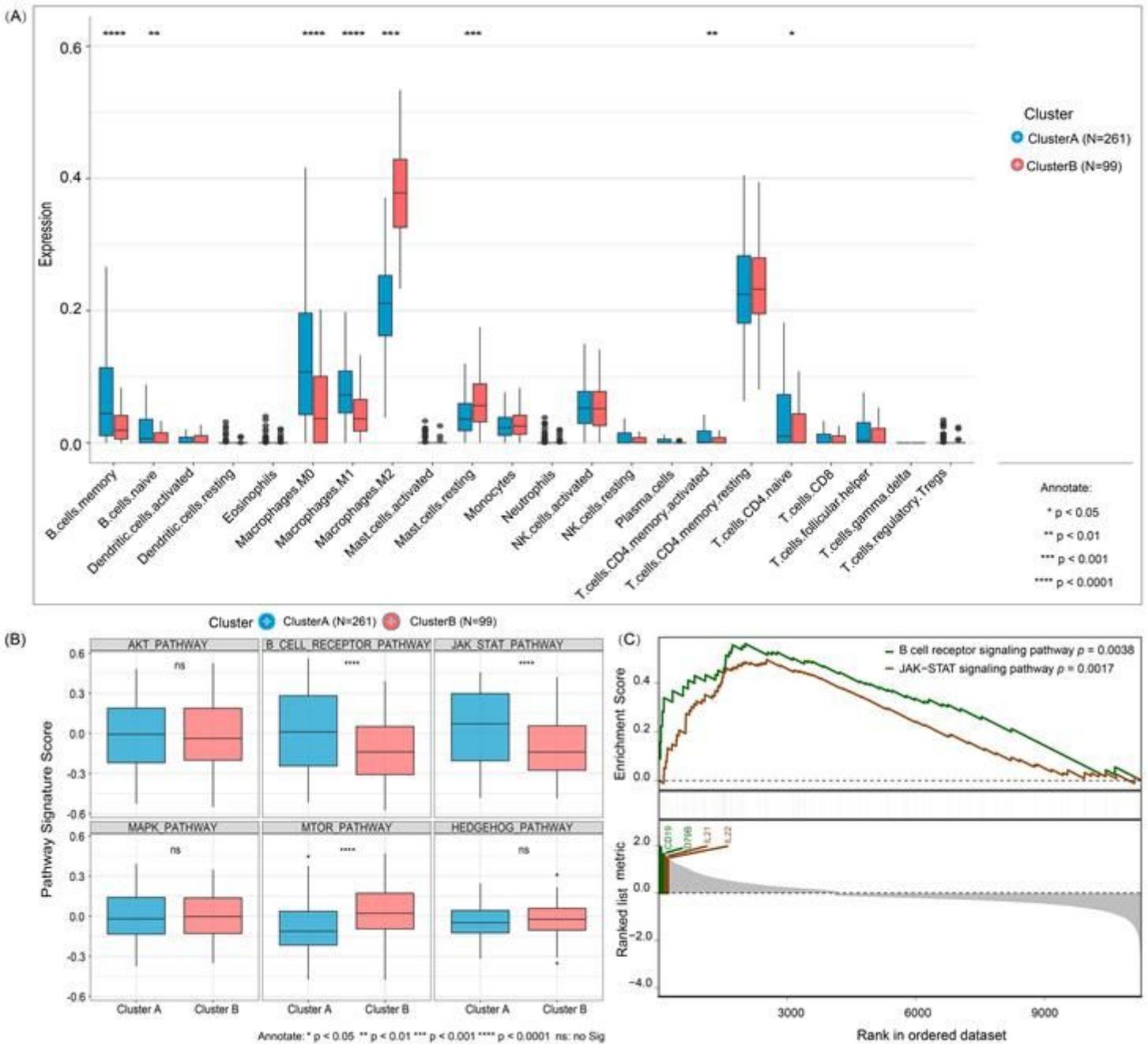


Figure 2

TME cell expression and aberrant pathway activations between two clusters Figure 2. A. Box plot of different immune cells expressed in different clusters, blue represents cluster A and red represents cluster B. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. Figure 2B. Indicates the aberrant pathway activations between cluster A and cluster B by GSVA; Blue and red represent cluster A and B cluster, respectively. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, ns: no significant Figure 2C. Explore the significantly aberrant pathway activations between cluster A and cluster B subtypes by GSEA; Dark green represents B cell receptor signaling pathway (P = 0.0038), Chocolate4 represents JAK-STAT signaling pathway(P = 0.0017).