Dissecting the Single-Cell Transcriptome Network in Superficial/Deep Tumor Tissues of Diffuse Gastric Cancer

Yili Ren
Affiliated Hospital of Shaoxing University
Beibei Zhang
Zhejiang Hospital
Chenkai Xu
Zhejiang Hospital
Lei Zhang (zhanglei19831008@163.com)
Affiliated Hospital of Shaoxing University

Research Article

Keywords: Gastric cancer, single-cell sequencing, GNG11, bioinformatics, microenvironment

Posted Date: December 3rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-1130074/v1

Abstract

Background and purpose: Gastric cancer is a type of highly heterogeneous malignant tumor and the prognosis of gastric cancer is hard to be improved due to limited knowledge on the molecular mechanism of heterogeneity. Single-cell sequencing technology is recently widely used for the investigation of both inter-tumoral heterogeneity and intra-tumoral heterogeneity. The present study aims to explore the potential oncogene by analyzing the single-cell data in the GSE167297 dataset.

Methods: The GSE167297 dataset was downloaded from the GEO database, followed by quality control to remove data with lower quality. The division on cell subtypes was determined by the characteristic marker expressed in each cell subpopulation. Wilcoxon rank-sum test was used to screen out differentially expressed genes. Survival analysis was performed to evaluate the prognostic value of G-protein subunit g 11 (GNG11) gene which was significantly overexpressed in deep tumor tissues of diffuse gastric cancer.

Results: In both normal tissues and tumor tissues, subtypes of immune cells and stromal cells were identified, with a higher proportion of infiltrated macrophages observed in deep tumor tissues. EPCAM was found significantly highly expressed in a cell subpopulation from gastric tumor tissues. 515 differentially expressed genes (| log2FC | > 2 and FDR < 1e-5) were screened out between normal tissues and tumor tissues. 86 differentially expressed genes (| log2FC | > 1 and FDR < 0.01) were screened out between superficial and deep tumor tissues, in which GNG11 was most highly expressed in deep tumor tissues (mean expression value: 0.1247, FC value: 52.2109). Disease-specific survival analysis on GNG11 results showed that the HR [95%CI] in the constructed univariate Cox proportional risk model was 4.419 [1.399-13.96] and the P-value in the log-rank test was 0.0056.

Conclusion: Differentially expression profiles were provided both extratumorally and intratumorally, indicating a higher infiltration of macrophages in deep tumor tissues. Additionally,GNG11 was screened out to be a significant risk factor in STAD patients.

Introduction

Gastric cancer is one of the most common malignant tumors in the world, and the morbidity and mortality of gastric cancer rank among the top in China [1]. Gastric cancer is a highly heterogeneous tumor, whose heterogeneity is not only reflected in different races, regions and populations, but also the same individual [2]. Due to gene instability and differences in the tumor microenvironment, offspring cell populations with completely different biological behaviors will be formed during the progression of gastric cancer and such cell populations show different proliferation rates, invasion ability, and sensitivity to therapies, which significantly affect the prognosis of patients [2]. Additionally, the heterogeneity of gastric cancer is observed between primary and metastatic foci, as well as between primary and recurrent foci [1]. Therefore, claiming the molecular mechanism underlying the heterogeneity of gastric cancer

using a novel sequencing method is of great significance to improving the prognosis of gastric cancer patients.

Single-cell sequencing technology was first presented by Tang et al in 2009 and in 2011 the journal of Nature Methods listed single-cell sequencing as one of the most anticipated technologies of the year. Since 2015, drop sequencing, microwell sequencing, split-pool ligation-based transcriptome sequencing, and other emerging technologies and platforms have lowered the threshold of single-cell sequencing, which is widely applicated in both basic scientific research and clinical research. On the basis of the next-generation sequencing (NGS), genome, transcriptome, and other multi-omics information can be obtained through single-cell sequencing, which is utilized to investigate the relationship between differences in cell population and cell evolution to explore the changes of biofunctions [3]

Tumor microenvironment (TME) is defined as the internal and external environment of tumor cells, in which large numbers of inhibitory immune cells are located [4]. Recently, it is confirmed that characteristics of cells in TME play an important role in the process of tumor proliferation and metastasis, which impact the therapeutic response of tumor cells to immune checkpoint inhibitors. Sathe et al used single-cell sequencing technology to describe characteristics of TME in gastric cancer at the single-cell level [5]. By single-cell transcriptome sequencing on tumors, adjacent tissues, and peripheral blood mononuclear cells from 7 gastric cancer and 1 intestinal metaplasia patients, they found reprogramming of immune cells and interstitial cells in TME compared with normal mucosa. Through deep sequencing of different cell components in TME, the unique characteristics of each cell type will be claimed by single-cell sequencing to clarify the interaction between gastric cancer cells, immune cells, and interstitial cells, which is of great importance for the development of novel potential therapeutic targets.

The present study conducted the single-cell processing following quality control on the single-cell data sets downloaded from the GEO database. By identifying the specific biomarkers on single cells, cell division was achieved both in norma gastic tissues and tumor tissues. The Wilcoxon's rank sum test was utilized to obtain differentially expressed genes, followed by analysis and functional annotation. The present study provided potential scrrened optimized oncogene for the development of treatment strategy against gastric cancer.

Materials And Methods

Data

Single-cell data sets of gastric cancer were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), the number of which was GSE167297. In GSE167297, approximately 30365 single cells collected from three tissue types, normal gastric tissue, superficial tumor tissues of diffuse gastric cancer, and deep tumor tissues of diffuse gastric cancer, were detected. Additionally, the data sets of TCGA Stomach Cancer (STAD) were downloaded from the database of

UCSC Xena (http://xena.ucsc.edu/), including the data of gene expression and the follow-up information on the prognosis of patients.

Quality control for the single-cell data sets

To ensure the quality of the data sets and obtain high-quality results in the subsequent analysis, we carried out preliminary quality control for the single-cell data. Firstly, the cell number and the captured gene number in 14 samples were counted. Subsequently, Subsequently, the number of captured genes in each cell in each sample and the proportion of cells expressing each gene were counted, followed by dividing the thresholds according to the distribution in all samples. Single cells expressed with lower number of genes and genes that only expressed in small number of cells were then eliminated. Finally, the proportion of mitochondrial gene expression in each single cell was counted to exclude single cells with the proportion of mitochondrial gene expression higher than 10%. Through above steps, the high-quality single-cell datasets were achieved.

Integration and standardized data processing on the single-cell data sets

Firstly, the 'Seurat' package in the R package was utilized to establish the data objects for each sample individually, followed by being standardized using the function 'NormalizeData' in the 'Seurat' package and screening highly variable genes using the function of 'FindVariableFeatures', among which default parameters were applied. Then, the integrated data functions of 'FindIntegrationAnchors' and 'integration data' were utilized to integrate multiple data sets collected from superficial or deep tumor tissues of diffuse gastric cancer into one data object for downstream analysis, respectively. Considering the inherent differences between the cells from normal tissue and tumor tissues, the data analysis was processed separately.

Single cell data processing

Series of standard data analysis procedures in the 'Seurat' (Version: 4.0.3) was conducted on achieved data objects (derived from normal tissues and integrated tumor tissues), including the data standardization, PCA dimension reduction, and K-Nearest Neighbors (KNN) classifier. All parameters were set as default except that the resolution of the "FindClusters" function was set to 0.5.

The division of cell subtypes

In the gastric tissue, cells are generally divided into 2 types: immune cells and stromal cells. For immune cells, CD45 (PTPRC+) was firstly used for localization and then subpopulation was divided into $CD4^+$ T cells (CD4), $CD8^+$ T cells ($CD8^+$), NK cells ($NKG7^+$), B cells (MS4A1+), and macrophages ($CD68^+$) according to the characteristic marker expressed in each cell subpopulation. For stromal cells, the identification on cell types was conducted according to the expressed characteristic markers, including vascular endothelial cells (ENG^+) [6], chief cells (PGC^+) [7], mucous cells ($TFF1^+$)[8], fibroblasts ($PDGFRA^+$ $COL1A1^+$) [9], and plasma cells (SLAMF7[CD138]$^+$, TNFRSF17[CD269]$^+$, SDC1[CD319]$^+$)[10].

Analysis and functional annotation of differentially expressed genes

Cancerous chief cells and mucous cells were identified from single-cell subsets in superficial and deep tumor tissues of diffuse gastric cancer based on the expression level of EPCAM, followed by calculating the mean expression level and fold change of each gene in normal and cancerous chief cells and mucous cells, respectively. The Wilcoxon's rank sum test was used to screen out genes with |log2FC| > 2 and false discovery rate (FDR) adjusted p value < $1*e^{-5}$. Additionally, differentially expressed genes that invaded into superficial and deep tumor tissues were screened according to the following standard: |log2FC| > 1 and FDR adjusted p value <0.01. Finally, the functional enrichment on differentially expressed genes was conducted using the DAVID tool (https://david.ncifcrf.gov/).

The survival analysis of GNG11 gene in gastric cancer patients

Survival analysis was performed to evaluate the prognostic value of GNG11 gene significantly overexpressed in deep tumor tissues of diffuse gastric cancer. According to the follow-up information on the patient prognosis recorded in the TCGA STAD dataset, patients were divided into 2 groups: the GNG11 highly expressed group and the GNG11 lowly expressed group. Then a univariate Cox proportional analysis model was constructed using the 'Survival' package (Version: 3.2-11)in the R, followed by performing the log-rank test.

Results

Results of the quality control for the single-cell data sets

Diffuse growth is reported in tumor cells in diffuse gastric cancer and the adhesion ability between cells is disrupted, with the invasive characteristic. To explore the molecular mechanism underlying differences on invasion of gastric cancer cells and the degree of infiltration, GSE167297 datasets were downloaded from the GEO database, which contained analyzed data of single cells collected from normal gastric tissues, superficial, and deep tumor tissues of diffuse gastric cancer. Totally 14 samples (30365 single cells) derived from 5 patients were collected to analyze the expression level of 32738 genes, among which normal gastric tissues were not available in one patient. Through sample statistics, single cell samples from 4298 normal tissues, 13,986 superficial tumor tissues of diffuse gastric cancer, and 12,081 deep tumor tissues of diffuse gastric cancer were obtained (Table 1 and Figure 1).

Subsequently, quality control was conducted on single cell samples extracted from the above tissues to eliminate samples and genes with low quality. Firstly, all single cell samples were integrated and the number of detected genes in each single cell was counted. According to the overall distribution, single cells with less than 500 detected genes were considered as low-quality samples and removed (Figure 2). Subsequently, the percentage of samples in which each gene was expressed was counted. According to the overall distribution, genes with a percentage less than 0.5% were considered as low-quality genes and removed (Figure 2).

Lastly, the proportion of mitochondrial gene expression in each sample was counted and single cells with the proportion of mitochondrial gene expression higher than 10% was eliminated. After the completion of quality control on collected samples, the number of genes and single cells was recounted in 14 samples, among which large amounts of cells were excluded in GSM5101025 (Table 2), indicating a poor sequencing output in the sample of GSM5101025.

Characterization on the microenvironment in normal gastric tissues

Considering inherent differences between the cells from normal tissue and tumor tissues, we firstly integrate data collected from 4 normal tissues (GSM5101013, GSM5101018, GSM5101021, and GSM5101024) and finally obtained the expressional profile data of 2296 single cells and 9926 genes. Subsequently, standard data analysis procedures in the 'Seurat' package was conducted on achieved data objects, including the data standardization, PCA dimension reduction, and KNN classifier. Single cells were divided into 12 subgroups, followed by identification on detailed cell types for each subgroup (Table 3, Figure 3).

Firstly, for immune cells, 6 subpopulations were extracted from 12 cell populations according to the expression level of CD45(PTPRC), including Cluster 0,1,3,4,6, and 11. As MS4A1 gene was highly expressed in Cluster 0, 3, and 6, we suspected that these three cell subpopulations were B cells. Subsequently, as CD4 and CD8 were highly expressed in Cluster 1 and 4, we considered these two cell subsets to be T cells. Additionally, the expression level of NKG7 was evaluated and we found that large amounts of single cells expressed with CD8 was mixed with NK cells. Lastly, as CD68 was highly expressed in Cluster 11, we considered Cluster 11 to be macrophages (Figure 4). By identifying the proportion of different immune cells, we found that a large number of immune cells were infiltrated in normal adjacent tissues collected from gastric cancer patients, among which $CD8^+$ T cells and NK cells might play a predominant role and macrophages might play a secondary role due to a small proportion.

Then, the cell type of remaining 6 cell subsets were identified. Firstly, mucous cells with high expression level of TFF1 and chief cells with high expression level of PGC were identified in Cluster 5, the cell types in which were considered as main cells that constituted gastric tissues. As ENG was highly expressed in Cluster 7, we suspected that cells in Cluster 7 were vascular endothelial cells. Cluster 9 was considered to be fibroblasts due to high expression level of COL1A1 and PDGFRA. Lastly, as the expression of SLAMF7, TNFRSF17, and SDC1 in Clusters 2 and 8 was relatively high, Clusters 2 and 8 were identified as plasma cells (Figure 5).

Characterization on TME in gastric tumor tissues

Data collected from 10 samples of superficial and deep tumor tissues of diffuse gastric cancer (Table 1) were integrated using the 'Seurat' package to obtain the expressional profile data of 16912 single cells and 9926 genes. Subsequently, standard data analysis procedures in the 'Seurat' package was conducted on achieved data objects, including the data standardization, PCA dimension reduction, and KNN

classifier. Single cells were divided into 15 subgroups, followed by identification on detailed cell types for each subgroup (Table 3, Figure 3).

Firstly, 9 subpopulations were extracted from 15 cell populations according to the expression level of CD45(PTPRC), including Cluster 0,1,2,5,6,10,11,13, and 14. Cluster 0 and 14 were considered as B cells due to high expression level of MS4A1. As CD4 and CD8 were highly expressed in Cluster 1,2,5,6, and 10, we considered Cluster 1,2,5,6, and 10 to be T cells. Additionally, similar to normal tissues, large amounts of single cells expressed with CD8 were mixed with NK cells. Lastly, as CD68 was highly expressed in Cluster 3,11, and 13, Cluster 3,11, and 13 were considered to be macrophages (Figure 4). We found that the number of macrophages in deep tumor tissues (n =1696) was much higher than that in superficial tumor tissues (n = 593), indicating that more macrophages were infiltrated in deep tumor tissues of diffuse gastric cancer. (Table 4, Figure 6).

Subsequently, stromal cell types in gastric tumor tissues were identified. According to the abnormal expression level of TFF1, MUC5AC, and PGC, we found that main cell types in Cluster 9 included abnormal chief cells and mucous cells. Cluster 7 was considered as vascular endothelial cell due to high expression level of ENG. Additionally, as COL1A1 and PDGFRA were highly expression in Cluster 8, Cluster 8 was identified as fibroblasts. Lastly, Cluster 4 and 12 were considered as plasma cells due to high expression level of SLAMF7, TNFRSF17, and SDC1.

EPCAM is regarded as an important biomarker of tumor stem cells in multiple types of malignant tumor [11]. In single cells derived from normal tissues and tumor tissues in the dataset of GSE167297, EPCAM was found significantly highly expressed in Cluster 9, which was a cell subpopulation from gastric tumor tissues. Addition, in Cluster 5, which was a cell subpopulation from normal gastric tissues, the expression level of EPCAM was also observed (Figure 9). Subsequently, screening on differentially expressed genes was conducted on 157 single cells in normal gastric tissues (Cluster 5, EPCAM low) and 421 single cells in gastric tumor tissues (Cluster 9, EPCAM high). The mean expression of each gene in the two types of single cells was calculated and the FC value was calculated by the Wilcoxon rank sum test. 515 differentially expressed genes ($|\log2FC| > 2$ and $FDR < 1 * e^{-5}$) were screened out, which was analyzed using DAVID to annotate the biological functions. As shown in Table 5 and Figure 10, we found that 515 differentially expressed genes were mainly closely associated with the following KEGG pathways: hsa04064 (NF-κB pathway), hsa04662 (B cell receptor signaling pathway), and hsa04142 (lysosome pathway).

1. ***protein subunit γ 11 (GNG11) was associated with the prognosis in gastric cancer patients***

To further explore the difference on single cells extracted from superficial and deep tumor tissues of diffuse gastric cancer, screening on differentially expressed genes was conducted on 296 single cells in superficial tumor tissues and 131 single cells in deep tumor tissues. The mean expression of each gene in the two types of single cells was calculated and the FC value was calculated by the Wilcoxon rank sum test. 86 differentially expressed genes ($|\log2FC| > 1$ and $FDR < 0.01$) were screened out, which was

further analyzed using DAVID to annotate the biological functions. As shown in Table 6 and Figure 11, 86 differentially expressed genes were mainly closely associated with the following KEGG pathways: hsa04510 (focal adhesion), hsa04672 (intestinal immune network for IgA production), and hsa05200 (cancer related pathway).

GNG11, which was most highly expressed in deep tumor tissues (mean expression value: 0.1247, FC value: 52.2109), compared to superficial tumor tissues (mean expression value: 0.0024), was picked out for analysis. According to the expression level of GNG11, 450 patients were divided into 2 groups: the GNG11 highly expressed group (mean expression value > 48.5364) and the GNG11 lowly expressed group (mean expression value <48.5364). The HR [95%CI] in the constructed univariate Cox proportional risk model was 1.811[1.308-2.508] and the P value in the log-rank test was 0.00029(Figure 12).

Disease-specific survival analysis was further performed on patients, which were divided into the GNG11 highly expressed group (mean expression value > 39.4713) and the GNG11 lowly expressed group (mean expression value <39.4713) according to the expression of GNG11. The HR [95%CI] in the constructed univariate Cox proportional risk model was 4.419 [1.399-13.96] and the P value in the log-rank test was 0.0056 (Figure 13), indicating that GNG11 was a significant risk factor in STAD patients.

Discussion

Gastric cancer is a type of cancer with high incidence in China. Highly heterogeneous biological characteristics are derived from the evolution of polyclonal selection during the development of malignant tumors and under the pressure of environmental treatments, which contribute to the molecular variations on cancer genes [12-14]. Currently, both inter-tumoral heterogeneity and intra-tumoral heterogeneity are observed in malignant tumors. Undoubtedly, the challenges in clinical treatments are significantly induced by the complex and heterogeneous characteristics of malignant tumors, which finally contributes to the main inducements of relapse, metastasis, and drug resistance. As the most promising biotechnology defined by Nature Methods in 2013, single-cell sequencing shows superior advantages in the identification of inter-tumoral heterogeneity and intra-tumoral heterogeneity recent years [15, 16]. Actually, single-cell sequencing has already been applied to investigate the heterogeneity in gastric cancer and specific genes have been identified to be responsible for the malignance of gastric cancer, such as OR51E1 [17], Gkn3 [18], and Krt7 [19].

In the present study, both inter-tumoral heterogeneity and intra-tumoral heterogeneity in the GSE167297 datasets were analyzed to identify the specific gene involved in the heterogeneity of gastric cancer. For the comparation between normal gastric tissues and gastric tumor tissues, EPCAM, a recognized oncogene, was used to test the accuracy of the gene screening in single-cell sequencing and we found that EPCAM was found significantly upregulated in tumor tissues, while rarely expressed in normal gastric tissues, indicating that the gene screening system was relatively reliable. 515 differentially expressed genes were screened out to be possibly associated with the tumorigenesis of gastric cancer.

For the investigation on intra-tumoral heterogeneity in GSE167297 datasets, 86 differentially expressed genes were identified to be possibly related to the heterogeneity in deep tumor tissues, among which GNG11 was identified as greatest difference in expression. GNG11 is a member of the family of heteromeric G-protein and plays an important role in regulating the activity of GTPase by transforming GDP into GTP. Ras pathway is reported to be regulated by the GTPase-activating protein through binding with the Ras protein [20]. In normal cells, GNG11 is previously identified as an inducer for cellular senescence in SUSM-1 cells, a human fibroblast cell line [21, 22]. In malignant tumor cells, GNG11 is found to be a differentially expressed gene (DEG) in triple-negative breast cancer [23] and has been proved to inhibit the progression of lung cancer [24]. However, in the present study, GNG11 was identified to be significantly upregulated in deep tumor tissues, compared to superficial tumor tissues. Additionally, the high expression level of GNG11 was positively correlated with the poor prognosis of gastric cancer patients, indicating that GNG11 might be an oncogene in the development and processing of gastric cancer. In our future work, the specific functions of GNG11 in gastric cancer will be further investigated in GNG11 overexpressed gastric cancer cell lines by knocking down GNG11.

Studies on tumor pathogenesis mostly focus on cancer cells, oncogene-related mechanisms, and important signaling pathways [25]. However, it is recently reported that the metastasis and invasion of tumor cells are largely dependent on the interaction between tumor cells and the surrounding microenvironment [26-28]. The migration of tumor cells can be promoted by mesenchymal cells in TME, which facilitates the development of novel therapeutic targets and antitumor strategies [28]. Under the recruitment of tumor-related signals, a variety of immune cell components in the immune microenvironment interact closely with tumor cells, which further evolve with each other to jointly promote the development of tumors [29, 30]. Therefore, fully understanding of the effects of corresponding components in tumor immune microenvironment on tumor cells is conducive to explore potential prognostic factors for gastric cancer [31, 32]. In the present study, cellular components in the microenvironment of normal gastric tissues and tumor tissues were investigated and compared, respectively. The greatest difference between normal gastric tissues and tumor tissues was that more infiltrated macrophages were found in tumor tissues, suggesting the important role of tumor associated macrophages (TAMs) in the development of malignant tumor, which provided a fundamental basis for the research on the potential therapeutic strategies by targeting TAMs [33, 34].

Taken together, our data provided a differentially expression profile between normal gastric tissues and diffuse gastric cancer tissues, as well as between superficial and deep tumor tissues of diffuse gastric cancer. Additionally, GNG11 was screened out to be a significant risk factor in STAD patients.

Declarations

Acknowledgements

Not applicable

Authors' contributions

Lei Zhang conceived and designed the study. Yili Ren is responsible for analysis and writing. Beibei Zhang and Chenkai Xu contributed in collecting and interpreting data. All authors approved the final version of the manuscript and are accountable for the accuracy and integrity in all aspects of the study.

Authors' information

[a] Department of Geriatrics, Affiliated Hospital of Shaoxing University, Zhejiang 312000, China

[b] Department of Oncology, Zhejiang Hospital, Zhejiang 312000, China

Funding

None

Availability of data and materials

The data analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

References

1. Smyth EC, Nilsson M, Grabsch HI, van Grieken NC, Lordick F: **Gastric cancer**. *Lancet* 2020, **396**(10251):635-648.
2. Pectasides E, Stachler MD, Derks S, Liu Y, Maron S, Islam M, Alpert L, Kwak H, Kindler H, Polite B *et al*: **Genomic Heterogeneity as a Barrier to Precision Medicine in Gastroesophageal Adenocarcinoma**. *Cancer Discov* 2018, **8**(1):37-48.
3. Grun D, van Oudenaarden A: **Design and Analysis of Single-Cell Sequencing Experiments**. *Cell* 2015, **163**(4):799-810.
4. Kaymak I, Williams KS, Cantor JR, Jones RG: **Immunometabolic Interplay in the Tumor Microenvironment**. *Cancer Cell* 2021, **39**(1):28-37.
5. Sathe A, Grimes SM, Lau BT, Chen J, Suarez C, Huang RJ, Poultsides G, Ji HP: **Single-Cell Genomic Characterization Reveals the Cellular Reprogramming of the Gastric Tumor Microenvironment**. *Clin*

Cancer Res 2020, **26**(11):2640-2653.

6. Kopczynska E, Makarewicz R: **Endoglin - a marker of vascular endothelial cell proliferation in cancer**. *Contemp Oncol (Pozn)* 2012, **16**(1):68-71.
7. Korstanje A, den Hartog G, Biemond I, Lamers CB: **The serological gastric biopsy: a non-endoscopical diagnostic approach in management of the dyspeptic patient: significance for primary care based on a survey of the literature**. *Scand J Gastroenterol Suppl* 2002(236):22-26.
8. Kouznetsova I, Peitz U, Vieth M, Meyer F, Vestergaard EM, Malfertheiner P, Roessner A, Lippert H, Hoffmann W: **A gradient of TFF3 (trefoil factor family 3) peptide synthesis within the normal human gastric mucosa**. *Cell Tissue Res* 2004, **316**(2):155-165.
9. Ivey MJ, Kuwabara JT, Riggsbee KL, Tallquist MD: **Platelet-derived growth factor receptor-alpha is essential for cardiac fibroblast survival**. *Am J Physiol Heart Circ Physiol* 2019, **317**(2):H330-H344.
10. O'Connell FP, Pinkus JL, Pinkus GS: **CD138 (syndecan-1), a plasma cell marker immunohistochemical profile in hematopoietic and nonhematopoietic neoplasms**. *Am J Clin Pathol* 2004, **121**(2):254-263.
11. Terris B, Cavard C, Perret C: **EpCAM, a new marker for cancer stem cells in hepatocellular carcinoma**. *J Hepatol* 2010, **52**(2):280-281.
12. Califano A, Alvarez MJ: **The recurrent architecture of tumour initiation, progression and drug sensitivity**. *Nat Rev Cancer* 2017, **17**(2):116-130.
13. Shibue T, Weinberg RA: **EMT, CSCs, and drug resistance: the mechanistic link and clinical implications**. *Nat Rev Clin Oncol* 2017, **14**(10):611-629.
14. McGranahan N, Swanton C: **Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future**. *Cell* 2017, **168**(4):613-628.
15. Shapiro E, Biezuner T, Linnarsson S: **Single-cell sequencing-based technologies will revolutionize whole-organism science**. *Nat Rev Genet* 2013, **14**(9):618-630.
16. Wang Y, Navin NE: **Advances and applications of single-cell sequencing technologies**. *Mol Cell* 2015, **58**(4):598-609.
17. Zhang P, Yang M, Zhang Y, Xiao S, Lai X, Tan A, Du S, Li S: **Dissecting the Single-Cell Transcriptome Network Underlying Gastric Premalignant Lesions and Early Gastric Cancer**. *Cell Rep* 2020, **30**(12):4317.
18. Bockerstett KA, Lewis SA, Noto CN, Ford EL, Saenz JB, Jackson NM, Ahn TH, Mills JC, DiPaolo RJ: **Single-Cell Transcriptional Analyses Identify Lineage-Specific Epithelial Responses to Inflammation and Metaplastic Development in the Gastric Corpus**. *Gastroenterology* 2020, **159**(6):2116-2129 e2114.
19. Dixon K, Brew T, Farnell D, Godwin TD, Cheung S, Chow C, Ta M, Ho G, Bui M, Douglas JM *et al*: **Modelling hereditary diffuse gastric cancer initiation using transgenic mouse-derived gastric organoids and single-cell sequencing**. *J Pathol* 2021, **254**(3):254-264.

20. Schaber MD, Garsky VM, Boylan D, Hill WS, Scolnick EM, Marshall MS, Sigal IS, Gibbs JB: **Ras interaction with the GTPase-activating protein (GAP)**. *Proteins* 1989, **6**(3):306-315.
21. Hossain MN, Sakemura R, Fujii M, Ayusawa D: **G-protein gamma subunit GNG11 strongly regulates cellular senescence**. *Biochem Biophys Res Commun* 2006, **351**(3):645-650.
22. Takauji Y, Kudo I, En A, Matsuo R, Hossain MN, Nakabayashi K, Miki K, Fujii M, Ayusawa D: **GNG11 (G-protein subunit gamma 11) suppresses cell growth with induction of reactive oxygen species and abnormal nuclear morphology in human SUSM-1 cells**. *Biochem Cell Biol* 2017, **95**(4):517-523.
23. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, Pietenpol JA: **Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies**. *J Clin Invest* 2011, **121**(7):2750-2767.
24. Hsu YL, Hung JY, Lee YL, Chen FW, Chang KF, Chang WA, Tsai YM, Chong IW, Kuo PL: **Identification of novel gene expression signature in lung adenocarcinoma by using next-generation sequencing data and bioinformatics analysis**. *Oncotarget* 2017, **8**(62):104831-104854.
25. Polyak K, Haviv I, Campbell IG: **Co-evolution of tumor cells and their microenvironment**. *Trends Genet* 2009, **25**(1):30-38.
26. Page-McCaw A, Ewald AJ, Werb Z: **Matrix metalloproteinases and the regulation of tissue remodelling**. *Nat Rev Mol Cell Biol* 2007, **8**(3):221-233.
27. Turley SJ, Cremasco V, Astarita JL: **Immunological hallmarks of stromal cells in the tumour microenvironment**. *Nat Rev Immunol* 2015, **15**(11):669-682.
28. Hanahan D, Weinberg RA: **Hallmarks of cancer: the next generation**. *Cell* 2011, **144**(5):646-674.
29. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD *et al*: **The prognostic landscape of genes and infiltrating immune cells across human cancers**. *Nat Med* 2015, **21**(8):938-945.
30. Zhan HX, Zhou B, Cheng YG, Xu JW, Wang L, Zhang GY, Hu SY: **Crosstalk between stromal cells and cancer cells in pancreatic cancer: New insights into stromal biology**. *Cancer Lett* 2017, **392**:83-93.
31. Hu WQ, Fang M, Zhao HL, Yan SG, Yuan JP, Peng CW, Yang GF, Li Y, Li JD: **Tumor invasion unit in gastric cancer revealed by QDs-based in situ molecular imaging and multispectral analysis**. *Biomaterials* 2014, **35**(13):4125-4132.
32. Peng CW, Tian Q, Yang GF, Fang M, Zhang ZL, Peng J, Li Y, Pang DW: **Quantum-dots based simultaneous detection of multiple biomarkers of tumor stromal features to predict clinical outcomes in gastric cancer**. *Biomaterials* 2012, **33**(23):5742-5752.
33. Pan Y, Yu Y, Wang X, Zhang T: **Tumor-Associated Macrophages in Tumor Immunity**. *Front Immunol* 2020, **11**:583084.
34. Cassetta L, Pollard JW: **Tumor-associated macrophages**. *Curr Biol* 2020, **30**(6):R246-R248.

Tables

Table 1. Tissue source and data statistics of single cells in the dataset of GSE167297

Type	Sample ID	Cell number
Normal	GSM5101013,GSM5101018, GSM5101021,GSM5101024	4298
Superficial layer	GSM5101014,GSM5101016, GSM5101019,GSM5101022,GSM5101025	13986
Deep layer	GSM5101015,GSM5101017,GSM5101020, GSM5101023,GSM5101026	12081

Table 2. Sample statistics of GSE167297 dataset after the quality control.

Sample ID	Gene number	Cell number	Gene number after QC	Cell number after QC
GSM5101013	32738	874	14138	256
GSM5101014	32738	2315	14789	1117
GSM5101015	32738	2234	15690	1446
GSM5101016	32738	2757	14782	1116
GSM5101017	32738	2932	14882	1965
GSM5101018	32738	1704	13962	1056
GSM5101019	32738	4753	13887	3427
GSM5101020	32738	3309	14902	2822
GSM5101021	32738	660	14114	359
GSM5101022	32738	2831	13837	2454
GSM5101023	32738	2214	13654	1965
GSM5101024	32738	1060	13326	719
GSM5101025	32738	1330	11363	261
GSM5101026	32738	1392	13640	856

Table 3. Markers used to identify cell types

Cell type	Markers
CD4+ T cells	PTPRC+ CD4+
CD8+ T cells	PTPRC+ CD8A+
NK cells	PTPRC+ NKG7+
Macrophages	PTPRC+ CD68+
B cells	PTPRC+ MS4A1+
Vascular endothelial cell	ENG+
Chief cells	PGC+
Mucous cells	TFF1+
Fibroblasts	PDGFRA+ COL1A1+
Plasma cells	SLAMF7 + TNFRSF17+ SDC1 +

Table 4. Statistical analysis of immune microenvironment cells in superficial and deep diffuse gastric tumor tissues

Cell number	Deep layer	superficial layer
T cells	4075	3529
B cells	1258	2256
Macrophages	1696	593

Table 5. KEGG pathway enriched with significantly differentially expressed genes in 515 normal and cancerous chief cells and mucous cells.

Term	Count	PValue
hsa05152:Tuberculosis	15	2.06E-04
hsa05162:Measles	12	6.84E-04
hsa04144:Endocytosis	16	0.001548973
hsa04064:NF-kappa B signaling pathway	9	0.00188605
hsa05164:Influenza A	13	0.001946987
hsa04662:B cell receptor signaling pathway	6	0.033914938
hsa04722:Neurotrophin signaling pathway	8	0.037487538
hsa05169:Epstein-Barr virus infection	8	0.040441405
hsa05134:Legionellosis	5	0.052077457
hsa04931:Insulin resistance	7	0.063314929
hsa05203:Viral carcinogenesis	10	0.088319884
hsa04142:Lysosome	7	0.097023672
hsa04810:Regulation of actin cytoskeleton	10	0.099680537

Table 6. KEGG pathway enriched with significantly differentially expressed genes in 86 chief cells and mucous cells from superficial and deep diffuse gastric tumor tissues

Term	Count	PValue
hsa05205:Proteoglycans in cancer	6	0.002098788
hsa05133:Pertussis	4	0.004920399
hsa05410:Hypertrophic cardiomyopathy (HCM)	4	0.005491451
hsa05323:Rheumatoid arthritis	4	0.007676167
hsa05143:African trypanosomiasis	3	0.010118118
hsa05146:Amoebiasis	4	0.012754218
hsa04510:Focal adhesion	5	0.014564198
hsa04672:Intestinal immune network for IgA production	3	0.019895641
hsa05144:Malaria	3	0.0215199
hsa05150:Staphylococcus aureus infection	3	0.025812516
hsa05134:Legionellosis	3	0.025812516
hsa05200:Pathways in cancer	6	0.033386616
hsa05321:Inflammatory bowel disease (IBD)	3	0.035336456
hsa04610:Complement and coagulation cascades	3	0.040536341

Figures

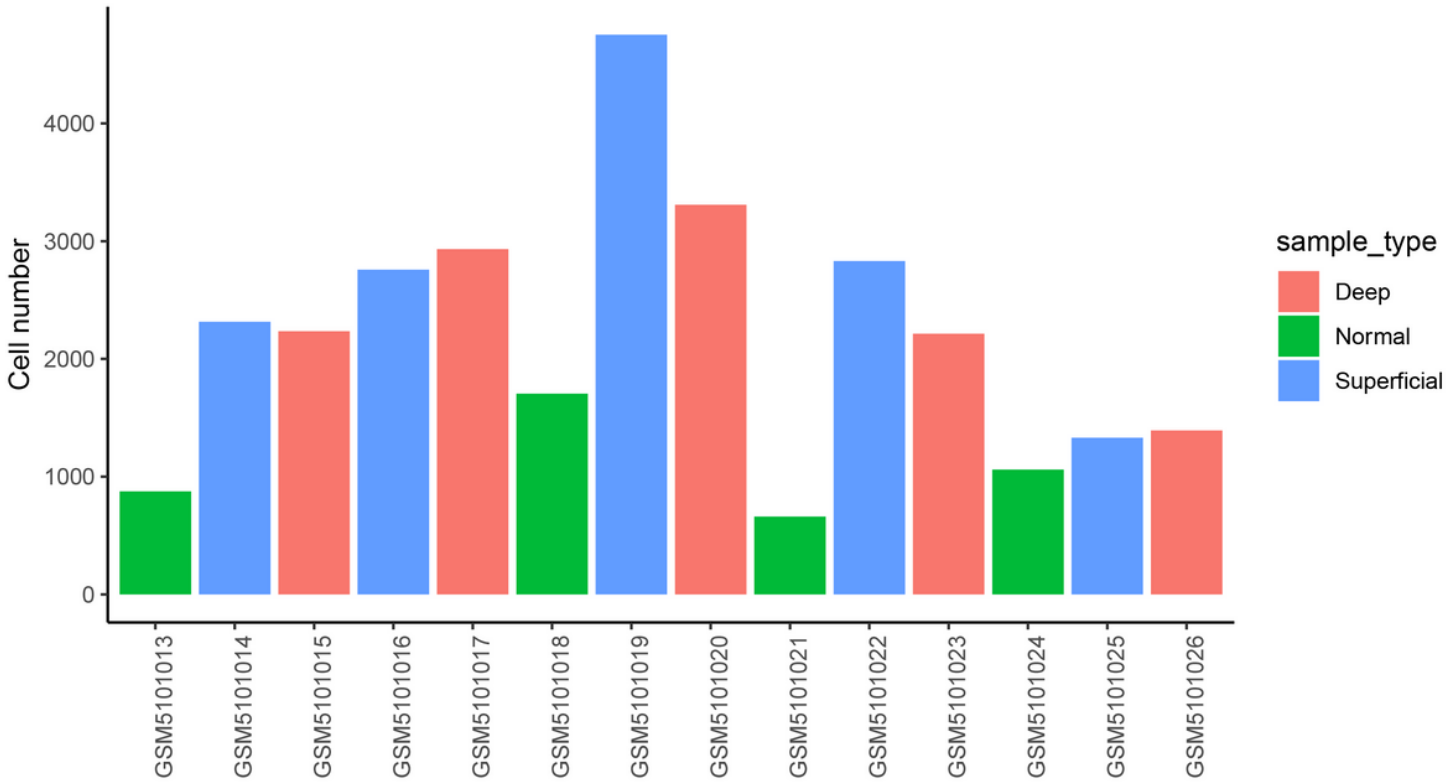

Figure 1

Tissue source and data statistics of single cells in the dataset of GSE167297.

Figure 2

Quality control and threshold selection on the single-cell data.

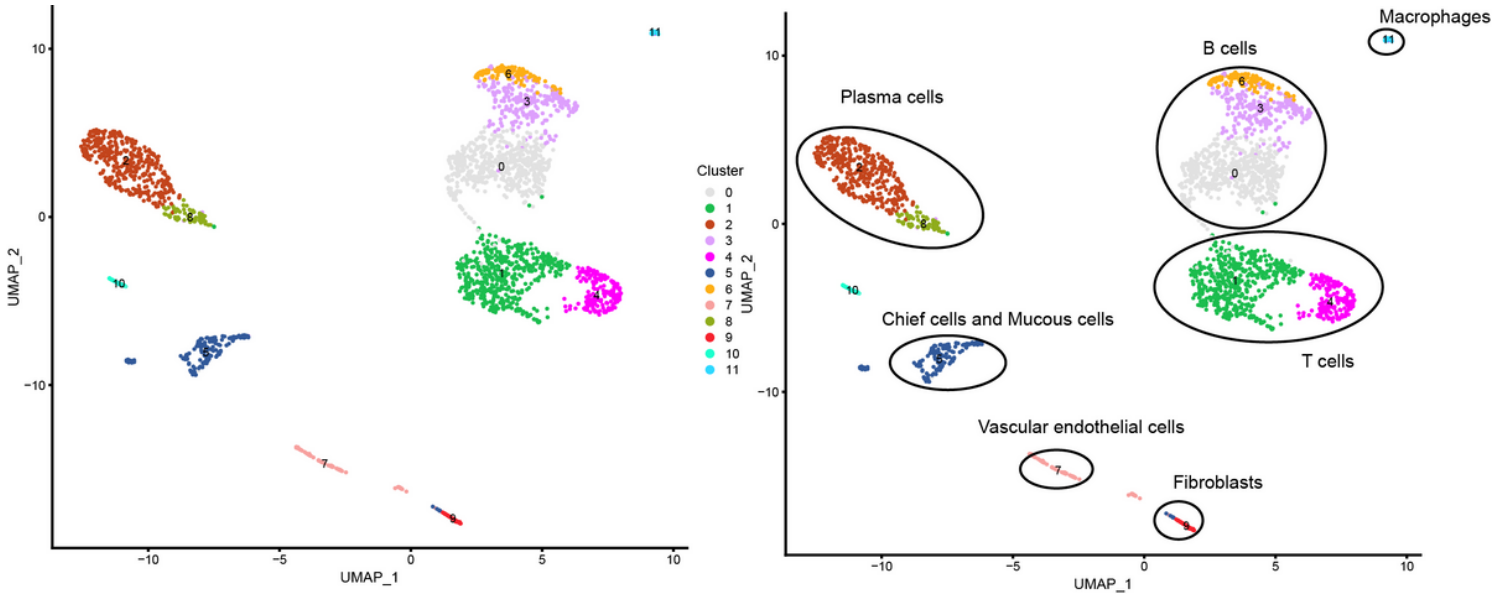

Figure 3

Cell type division of single cell samples isolated from normal gastric tissues.

Figure 4

Distribution of the expression levels of immune cell markers in the gastric microenvironment.

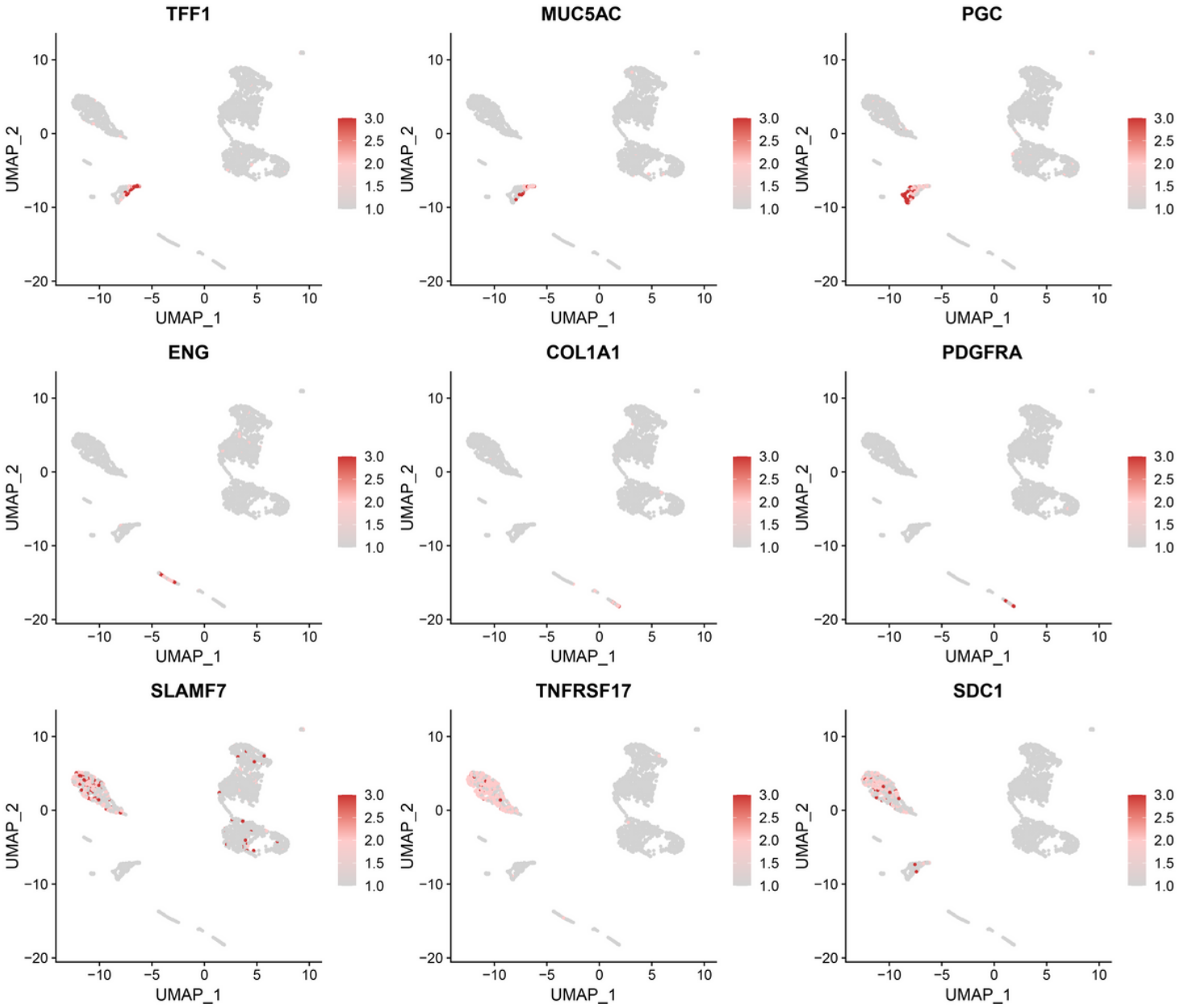

Figure 5

Distribution of the expression levels of stroma cell markers in the gastric microenvironment.

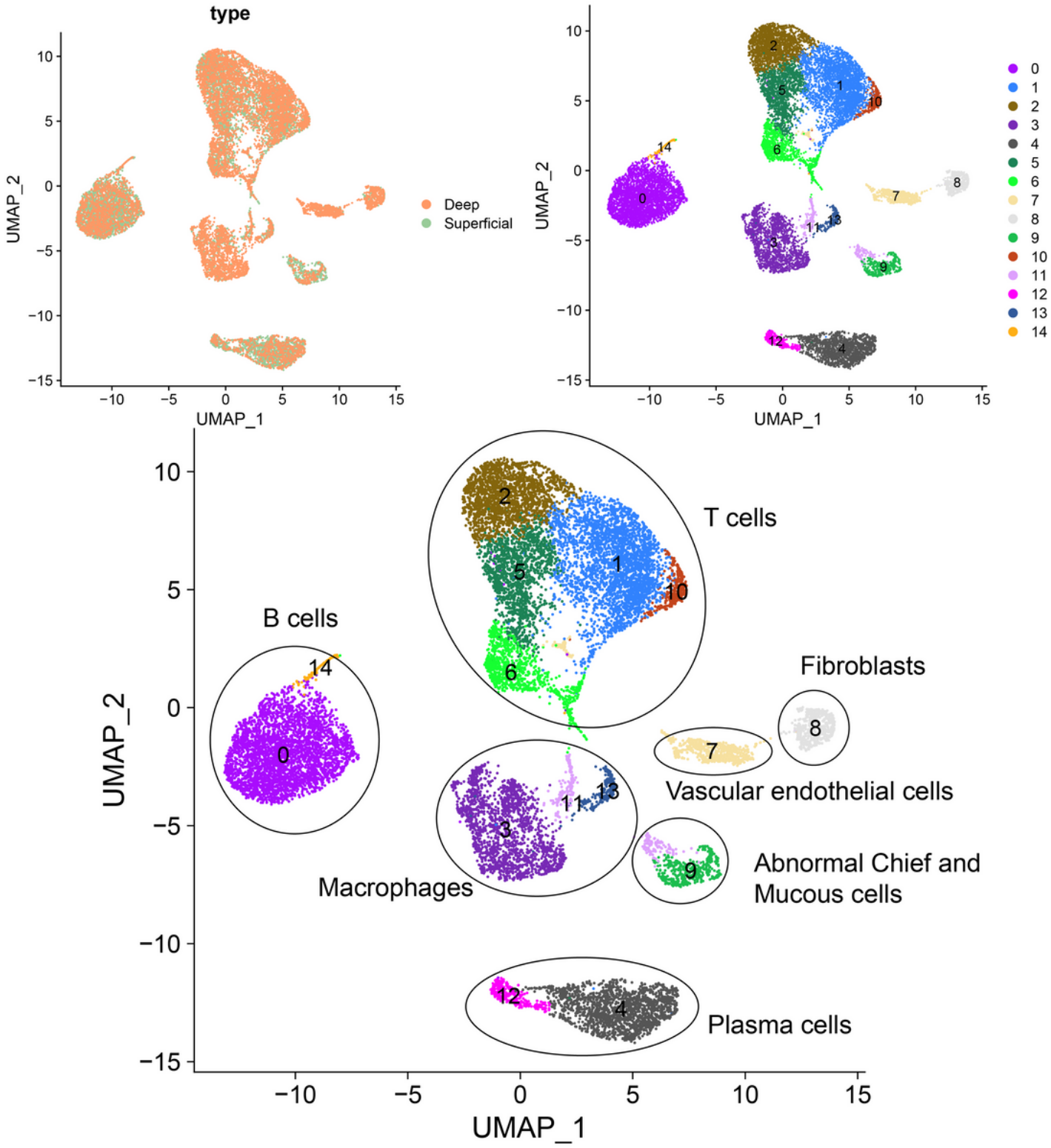

Figure 6

Cell type division of single cell samples isolated from diffuse gastric cancer tissue.

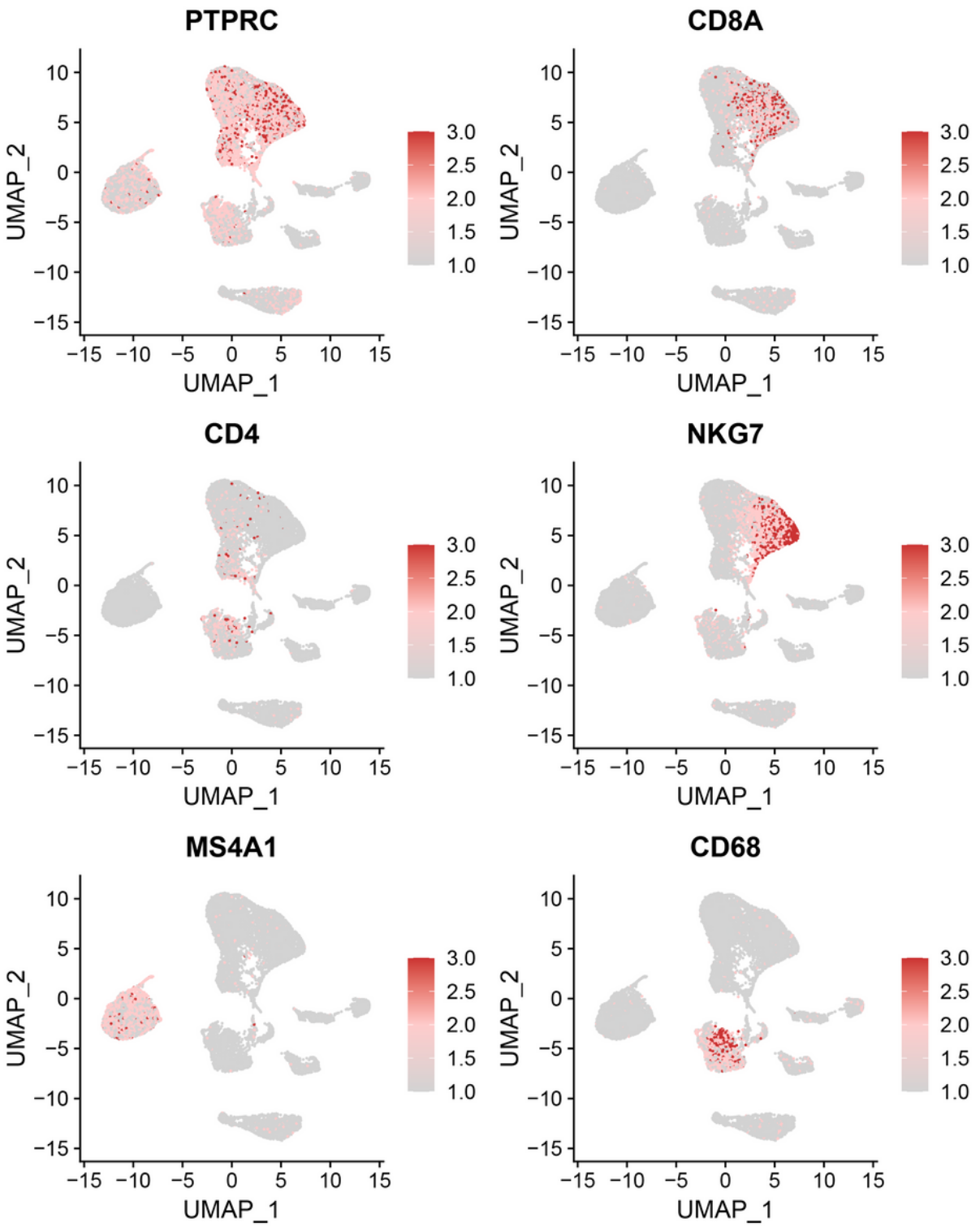

Figure 7

Distribution of the expression levels of immune cell markers in the microenvironment of gastric cancer.

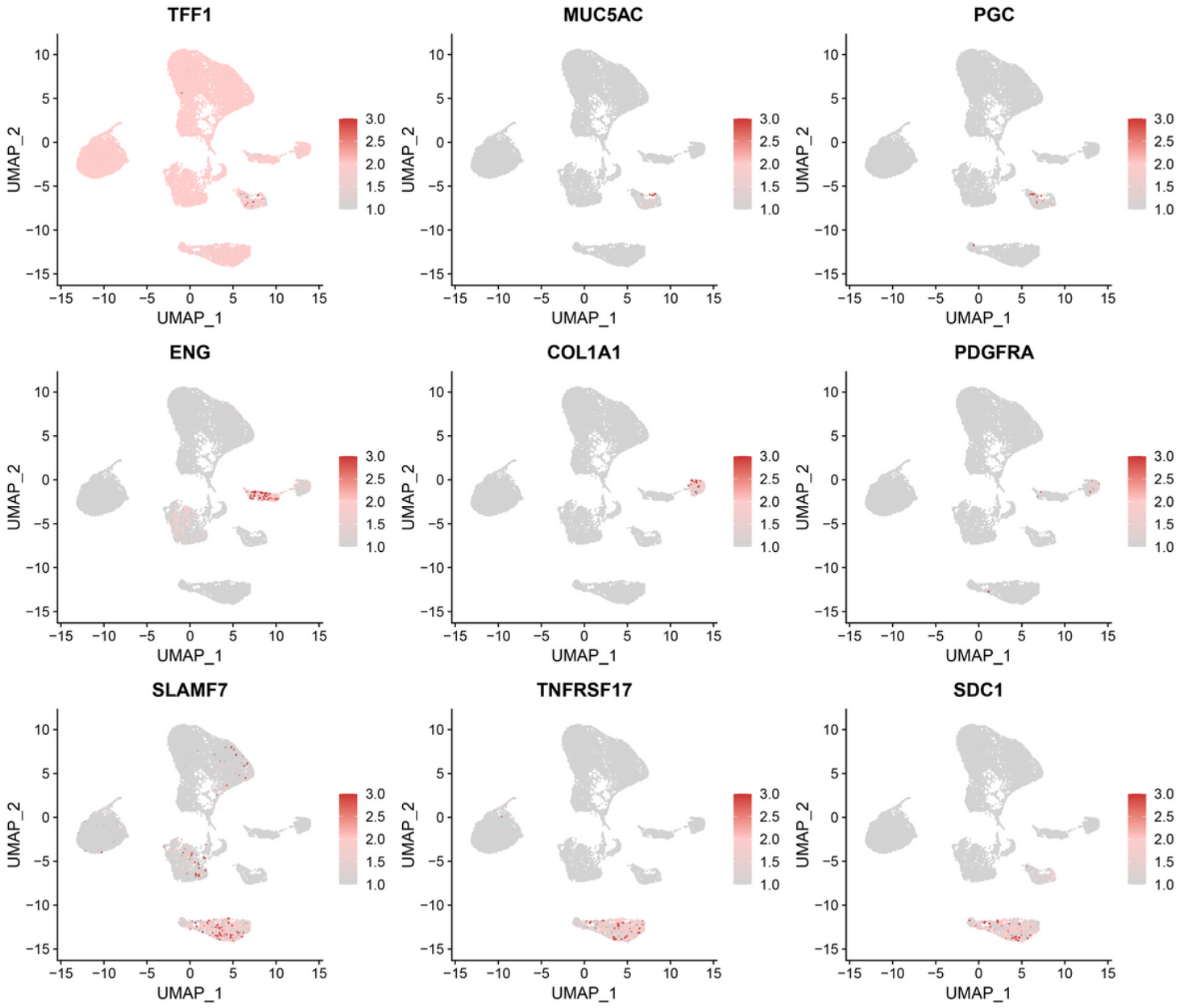

Figure 8

Distribution of the expression levels of stroma cell markers in the microenvironment of gastric cancer.

Figure 9

Expression pattern of EPCAM in normal and gastric cancer tissues.

Figure 10

KEGG pathway enriched with significantly differentially expressed genes in 515 normal and cancerous chief cells and mucous cells.

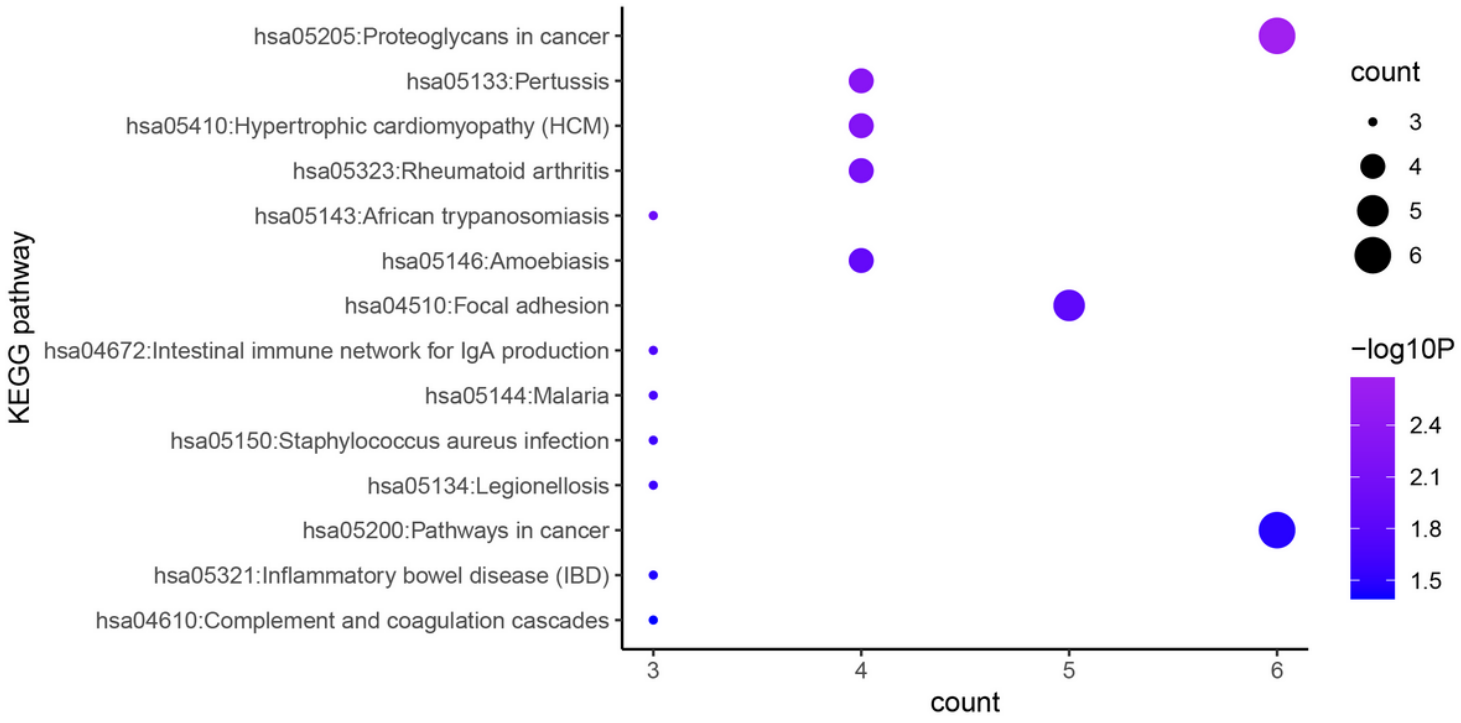

Figure 11

KEGG pathway enriched with significantly differentially expressed genes in 86 chief cells and mucous cells from superficial and deep diffuse gastric tumor tissues.

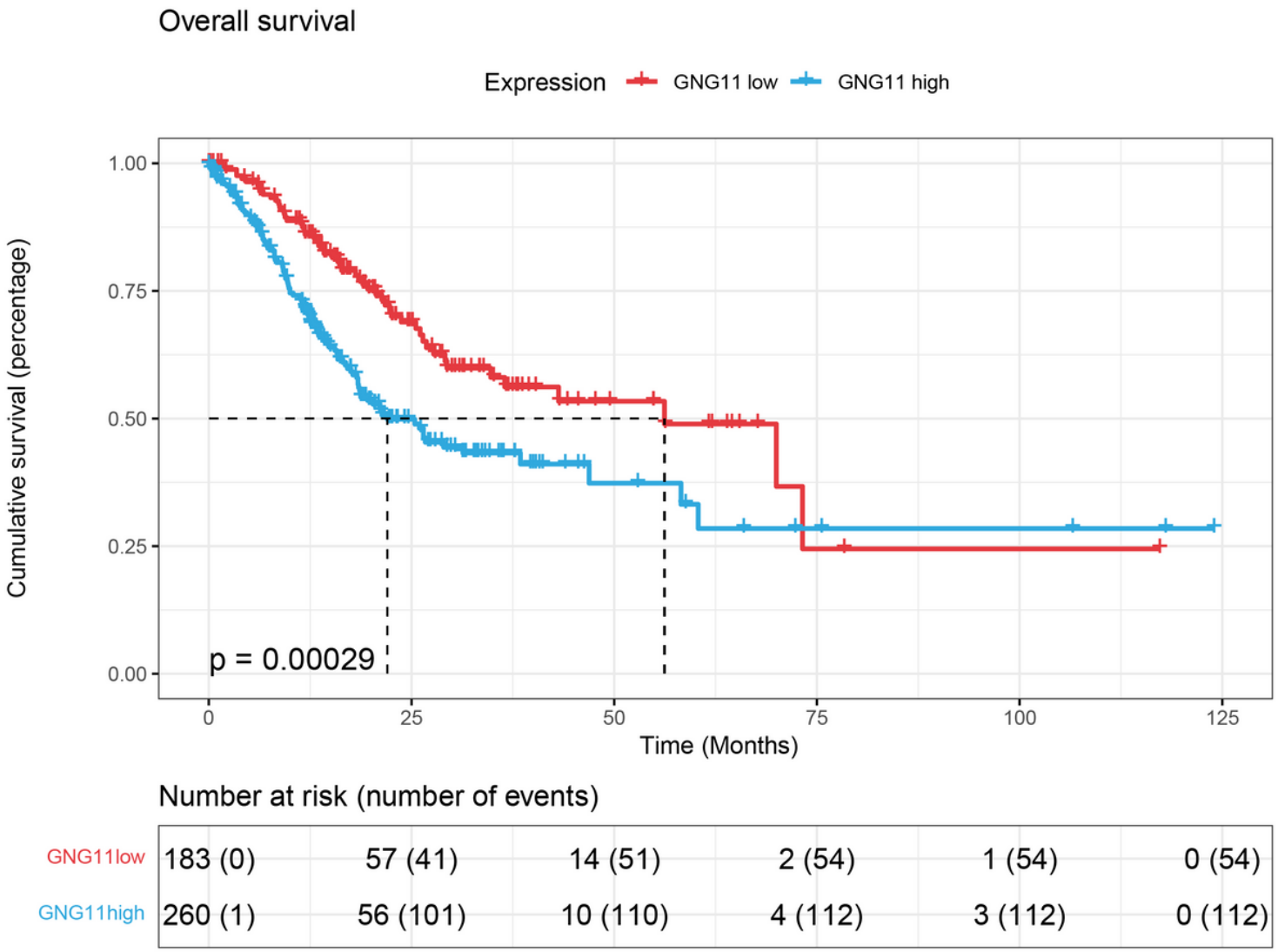

Figure 12

Overall survival K-M curve of GNG11 in TCGA STAD patients.

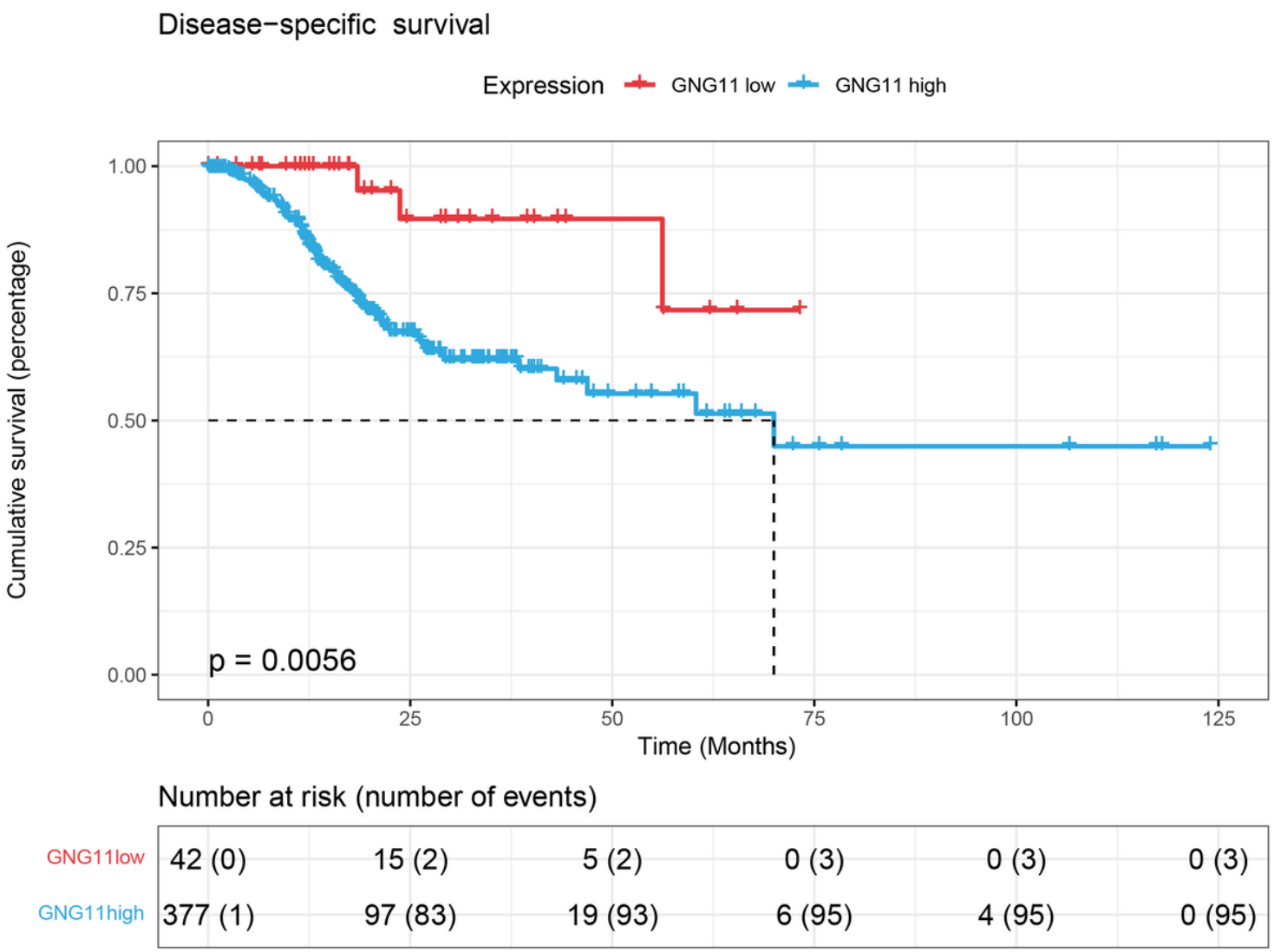

Figure 13

Disease specific survival K-M curve of GNG11 in TCGA STAD patients.