

Complement C1QC as a potential prognostic marker and therapeutic target in colon carcinoma based on single-cell RNA sequencing and immunohistochemical analysis

Huiming Deng^{1,2†}, Yan Chen^{3†}, Yong Liu¹, Li Liu¹, Ronghua Xu^{1*}

ABSTRACT

Immune cell infiltration plays an essential role in the occurrence and development of colon cancer. However, the main tumor-associated immune cell infiltration and its gene regulation in colon cancer still need to be further clarified to provide a new perspective for diagnosing and treating this disease. For this study, single-cell RNA sequencing (scRNA-seq) expression profiles and TCGA colon cancer data sets were first acquired from the GEO database. Then, Seurat, Monocle, LIMMA, Clusterprofile, GSVA, and GSEABase algorithms were used to systematically examine the data. Potential target drugs corresponding to target genes were analyzed in the Drugbank database and detected by molecular docking. Immunohistochemistry was used to assess the level of C1QC expression in the tissue microarray. Single cell analysis suggested that neutrophil activation might be the critical regulatory pathway in colon cancer and that macrophages were the main cell population involved. Subsequent functional enrichment analysis on differential genes in macrophages suggested that C1QC may be a critical regulatory factor in the occurrence and progression of colon cancer and was closely related to the survival of patients. According to the drug target prediction, palivizumab is a targeted drug for C1QC, and molecular docking demonstrated that palivizumab binds to C1QC. In addition, tissue-microarray-based immunohistochemical analysis showed that C1QC was highly expressed in colon cancer tissue, and the prognosis of colon cancer patients with high C1QC expression was worse, closely related to age, lymphatic metastasis, and the TNM stage (Tumor, Nodes, and Metastases). Our findings suggest that C1QC may regulate the macrophages in colon cancer immune infiltration, which is expected to be a potential immunotherapy target for colon cancer, and beneficial for the diagnosis and prognosis of colon cancer patients.

KEYWORDS: C1QC; colon carcinoma; prognostic marker; single-cell RNA sequencing

INTRODUCTION

With more than 1.85 million new cases and 850,000 fatalities per year [1,2], colorectal cancer has emerged as one of the fastest-growing cancers as well as the third leading cause of death globally, following the diet changes and aging of the population. In recent years, surgery, radiation and chemotherapy, targeted therapy, immunotherapy, and other treatment techniques have increased the overall survival (OS) of patients

with advanced colon cancer. For patients with colon cancer, survival has been impacted by early diagnosis, post-operative recurrence, chemotherapy drug resistance, and severe adverse reactions [3,4]. Therefore, it is crucial to advance understanding of the pathogenesis of colon cancer and search for potential biomarkers to further improve the diagnosis and survival rates of patients with colon cancer.

Studies have shown that immune cells in the tumor microenvironment participate in the biological development of colon cancer cells in a variety of ways, including tumor angiogenesis, metastasis, and immune escape. These roles have a significant impact on the pathological evaluation, treatment decision, and prognosis of colon cancer patients [5-7]. Analyzing the immune cell subsets regulating the functions of colon cancer cells with significant differences will be beneficial for the diagnosis and proper treatment of colon cancer. As a result of advances in biotechnology, researchers have developed the cutting-edge single-cell-based transcriptome sequencing approach (scRNA-seq) for analyzing the tumor microenvironment and clearly distinguishing cell diversity and heterogeneity in the tumor [8].

In recent years, scRNA-seq has revealed the complexity of tumor-infiltrating cells in different cancer types, including tumor-associated macrophages [9-11]. Macrophages are

¹Department of Gastrointestinal Surgery, Huazhong University of Science and Technology Union Shenzhen Hospital, Shenzhen, China
²Molecular Pharmacology Research Center, School of Pharmaceutical Science, Wenzhou Medical University, Wenzhou, China
³Department of Ultrasound, Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University, Haikou, China
[†]These authors equally contributed

*Corresponding author: Ronghua Xu, Department of Gastrointestinal Surgery, Huazhong University of Science and Technology Union Shenzhen Hospital, No. 89 Taoyuan Road, Shenzhen 518000, China. E-mail: hyfyweichang@163.com

DOI: <https://doi.org/10.17305/bjbms.2022.7309>

Submitted: 02 April 2022/Accepted: 15 June 2022/
Published online: 23 June 2022

Conflict of interest: The authors declare no conflicts of interest.

Funding: This study was supported by Nanshan District of Shenzhen Science and Technology Project (2019057).



©The Author(s) (2022). This work is licensed under a Creative Commons Attribution 4.0 International License

heterogeneous cell types that contribute to the development and progression of cancer by generating vascular growth factors, promoting extracellular matrix remodeling and inhibiting immunity [12]. Clinicians have utilized immunotherapy to inhibit tumor-associated macrophages, including disruption of macrophage expansion and differentiation by blocking the interaction between the colony-stimulating factor 1 receptor and its ligand colony-stimulating factor 1 and interleukin-34 [13-15]. Moreover, monotherapy was found to only have a limited impact on cancer due to compensatory inhibition of body mechanisms [16,17]. Therefore, searching actively for genes that regulate macrophages involved in the development of colon cancer is crucial for immunotherapy.

The present study analyzed changes in cell populations in single-cell data of colon cancer and paracancer tissue and identified five cell populations. Using functional enrichment analysis, macrophage populations that significantly correlated with neutrophil activity were obtained. We also screened out the differentially expressed gene C1QC, which regulates the infiltration of macrophages in colon cancer and affects the prognosis of colon cancer patients. C1QC was confirmed to be highly expressed in colon cancer tissue and the prognosis of colon cancer patients with high expression of C1QC is worse. Palivizumab was identified as a possible target drug for C1QC by drug target predictions and a molecular docking assay. Our study revealed the specific immune infiltrating subsets of macrophages in colon cancer, as well as the related target genes and targeted drugs that regulate macrophages involved in colon cancer development. The findings have provided new methods for colon cancer diagnosis and treatment.

MATERIALS AND METHODS

scRNA sequencing analysis

GSE163974 expression profile data by high-throughput scRNA-seq using 10× Genomics including barcodes, genes, and matrix files were downloaded from the GEO database. The samples included three normal tissues (GSM4994385) and three cancer tissues (GSM4994386). The data analysis was as follows: (1) Expression spectrum acquisition: A matrix package was used to read barcodes, genes, and matrix files, and each single cell cycle was acquired by single cell experiment, scan, and a scatter function. The cell cycle was defined as the G₁ phase: G₁ score > 0.5 and G₂/M score < 0.5; G₂/M phase: G₂/M score > 0.5 and G₁ score < 0.5; S phase: G₁ score < 0.5 and G₂/M score < 0.5; unknown: G₁ score > 0.5 and G₂/M score > 0.5. (2) Data quality control: The Seurat function was used to filter data, and parameters were set as min. cells = 3 and min. features = 200, the proportion of mitochondrial genes < 0.05, and the number of cell genes ranged from 200 to 20,000. (3) Unsupervised clustering was used to select

appropriate components to construct the cell cluster: (a) Data normalization was carried out using the global-scaling method; (b) the candidate genes with high expression variation were screened based on an average value algorithm; (c) PCA analysis was used to observe the batch effect between samples; (d) the Seurat random sampling method was used to construct a background distribution of correlation values between characteristic genes and principal components, and the JackStraw algorithm was employed to screen suitable principal components for subsequent cell subgroup analysis; (e) the expression profile of each cell was converted into a highly correlated cell population using the KNN algorithm to complete cell cluster identification, resolution = 0.6; and (f) the genes with certain differential expression folds and detectable in most cells in each cluster were determined using a log-scale as subsequent markers, and the criteria were: min.pct = 0.25, logfc.threshold = 0.25; (3) Cell subsets definition: Cell subsets were reannotated by SingleR and scCATCH.

Pseudotime analysis

The Monocle algorithm was used to analyze the developmental trajectories of genes with large changes. Genes with high expression variability in developmental trajectories were screened for subsequent analysis. The criteria were status = ok, family = tobit, $p < 0.05$, and order = true.

Functional enrichment analysis of pseudotime regulatory genes

DEGs were obtained by pseudotime analysis and the cluster profile package was used for the GO analysis. The differential pathway was corrected by the BH method and the adjusted $p < 0.05$.

Pathway enrichment and GSVA analysis

After the critical pathway, most related to the pseudotime-difference genes was obtained, the gene set of a pathway was downloaded from the MsigDB database (<https://www.gsea-msigdb.org/gsea/msigdb/>), and a standardized score was obtained for each cell using GSVA and GSEABase algorithms.

TCGA analysis

RNA-seq Recompute TPM sample data of GDC TCGA-HNCC, clinical characteristics, and survival follow-up data were downloaded from the XENA database (<https://xenabrowser.net/datapages/>). DESeq algorithm was used to normalize the expression profile and filter low expression genes. The selection criteria for differential genes were a log₂ fold change ≥ 1.0 , and a p -value after the Benjamini-Hochberg procedure < 0.05 . In addition, cluster profile packages were used for GSEA analysis

of differential genes, and differential pathways were corrected by the BH method, with an adjusted $p < 0.05$.

Hub genes identification and functional enrichment analysis

The development-related differential genes obtained by scRNA pseudotime analysis were divided into different modules according to different developmental states. Functional enrichment analysis was performed for each gene in different modules in the Toppgene database (<https://toppgene.cchmc.org/>), with screening criteria of FDR adjusted to $p < 0.05$.

Hub genes screening and survival analysis

After obtaining the cell population with the most significant activation of the critical pathway through GSVA analysis, the genes with the most remarkable differences and the most significant p -value in the cell population were selected as candidate genes. In addition, the expression values of hub genes were used to analyze OS by the survival and survminer algorithm in each TCGA patient.

Functional enrichment analysis of hub genes and prediction of drug targets

After obtaining target genes, functional enrichment analysis was conducted in the Toppgene database (<https://toppgene.cchmc.org/>). Potential target drugs corresponding to target genes were analyzed in the Drugbank database (<https://go.drugbank.com/>), and the screening criterion was $p < 0.05$ after FDR correction.

Tissue microarrays and IHC assays

Colon cancer tissue microarray (TMA, HCoA18oSu18-XT17-025) was obtained from Shanghai Outdo Biotech, including 93 cases of colon cancer tissue and 87 cases of non-cancer tissue, with complete case and follow-up information. IHC assays and IHC scores were performed with a previously described protocol [18]. Briefly, the colon cancer tissue microarray was dewaxed and rehydrated using gradient alcohol after being baked in an oven for 1 h at 63°C. The endogenous peroxidase was blocked and inactivated by 3% H₂O₂. The tissue microarray was subject to antigen thermal repair with citric acid and was incubated overnight in a refrigerator at 4°C with C1QC primary antibody (Abcam, MA, US) and continued for 15 min at 37°C with secondary antibody (Zhong ShanGolden Bridge Biological Technology Co., Ltd, Beijing, China). The colon cancer tissue microarray was stained with DAB (Beyotime Biotechnology, Shanghai, China) and restained with hematoxylin (Sigma, CA, US). After dehydration and transparency, the samples were sealed and scanned.

The scoring system for IHC is as follows: The degree of immunostaining (A): 0 point for no obvious coloring, 1 point for mild, 2 points for moderate, or 3 points for strong; percentage of positive cells (B): <5%, 0 point; 6~25%, 1 point; 26~50%, 2 points; 51~75%, 3 points; and >75%, 4 points. The total score for each examined slide was the summed score of A+B. A high expression level was defined as the sum of scores ≥ 4 points, while a low expression level was considered as the sum of scores <4 point.

Molecular docking

The protein structures of C1QC and palivizumab were downloaded from the PDB database and saved in the pdb format. Docking was performed by the zDock module from Discovery Studio 2019 to select the conformation with the highest score for subsequent analysis.

Statistical analysis

Statistical analysis and chart preparation were performed by GraphPad Prism software 8.0 (GraphPad Software, Inc., San Diego, CA, USA). Student's t-test was used for the comparison between the two groups. A Chi-squared test was used to analyze clinicopathological characteristics, and the Kaplan-Meier method and log-rank test for survival analysis. $p < 0.05$ was considered statistically significant.

RESULTS

Single-cell analysis results

After strict quality control, single-cell transcriptome data of 1211 cells were obtained, including 322 cells from cancer tissue and 889 cells from normal tissue (Figure 1A). After data standardization and correction, unsupervised clustering was performed on the combined dataset, and the cell types were annotated using known marker genes for specific cell types. 4,997 variable genes were identified, including STX3, ZG16, HERC4, NMNAT1, CCL18, and TMEM2 (Figure 1B). In addition, t-SNE and UMAP dimensionality reduction algorithms were used (two algorithms should correspond to two types of plots, when using two algorithms are redundant). The data set identified five cell populations, including CD8+ T-cells, neutrocytes, macrophages, B-cells, and epithelial cells. The distribution of cells is shown in two-dimensional space, with each point representing a cell (Figure 1C). We found a series of immune cells involved in gastric cancer by single-cell analysis.

Pseudotime analysis and enrichment analysis

The expression profile data were projected into low-dimensional space to construct the differentiation trajectory among cells. Each dot represents a cell and cells in a similar cell state are grouped together. This study found that two

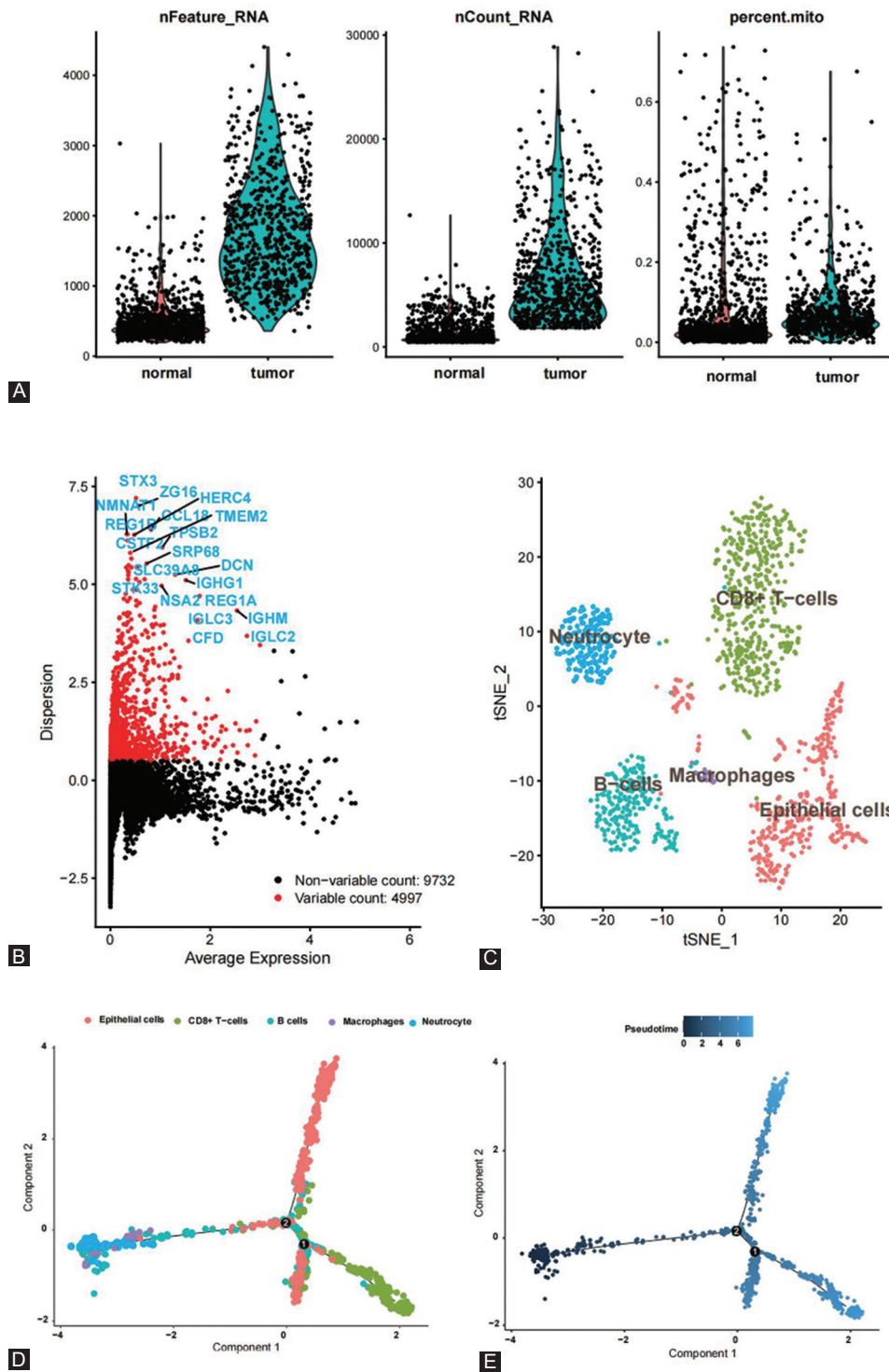


FIGURE 1. scRNA-seq clustering and pseudotime analysis in the GSE163974 dataset. (A) The sample quality control and proportion of mitochondria; (B) the highly variable genes are presented and the top 25 genes of each dataset are visualized; (C) UMAP map showing the scRNA cell clusters and annotation in each dataset; and (D and E) pseudotime analysis of the GSE163974 dataset.

branching points in the differentiation trajectory represent potential decision-making points in cellular biological processes (Figure 1D and E). Figure 1D is color-coded according to cell clustering and annotation information, while Figure 1E is color-coded according to the differentiation state.

Developmental trajectory analysis was conducted based on the gene sets with great changes. Log standardization

and DDRTree dimension reduction were carried out and B-H corrected differential regulatory genes were selected for subsequent functional enrichment analysis after pseudotime analysis. Figure 2A shows that the BP pathway is associated with cancer differentiation and progression, including neutrophil activation ($p_{\text{adjust}} = 2.63 \times 10^{-41}$; enriched genes = 158), neutrophil-mediated immunity

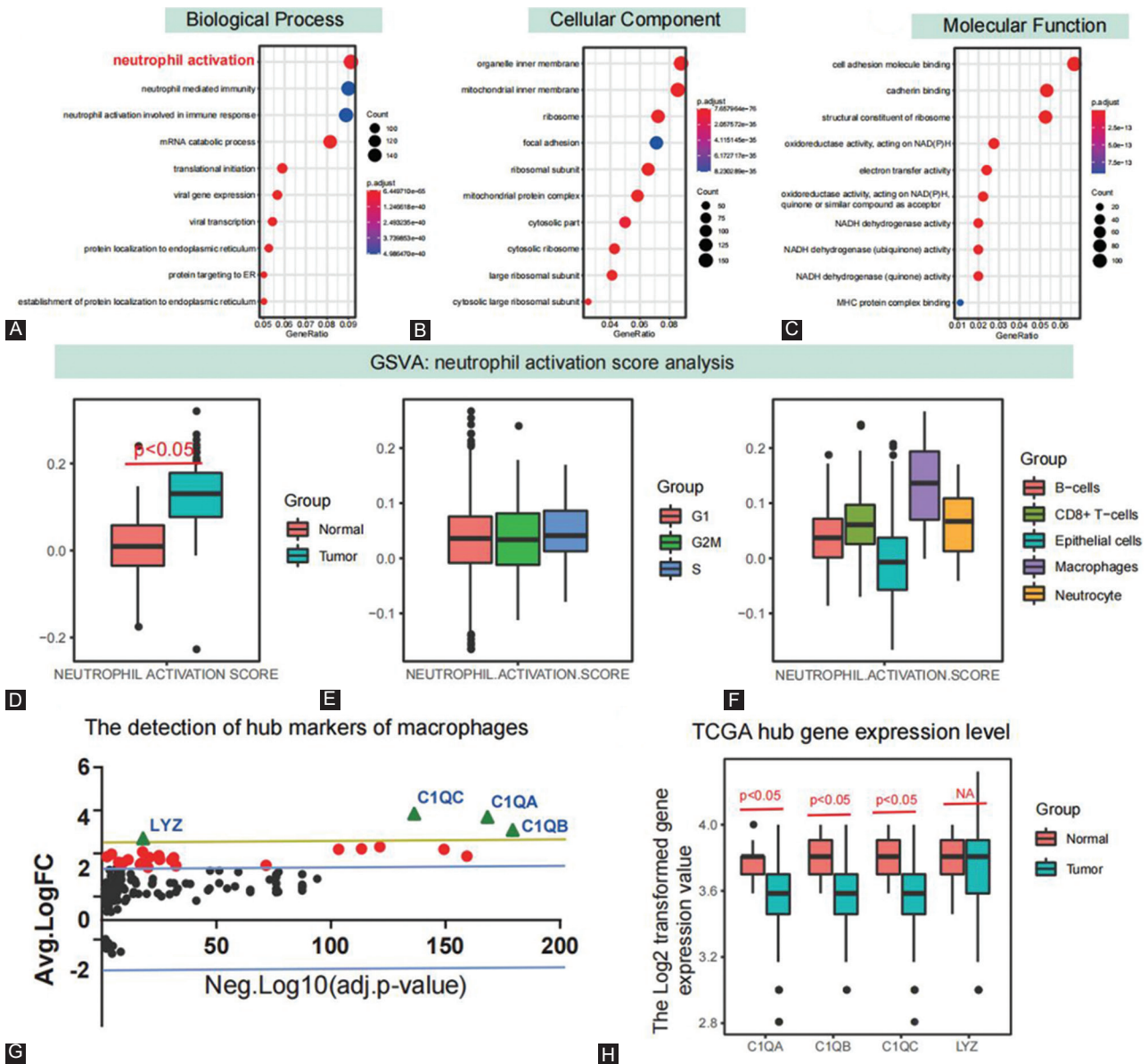


FIGURE 2. Functional enrichment analysis of pseudotime related differential expressed genes. (A) The GO terms analysis in response to pseudotime related differential expressed genes; (B) the tendency of GSEA calculated neutrophil activation scores involved in cancer samples, cell cycle, and different cell populations; (C) identification of the hub genes in response to development and progress of CRC involving macrophages; (D) the hub gene's expression level in response to the TCGA-CRC dataset; (E and F) neutrophil activation score changes in different cell cycles and cell subsets in single cells; (G) distribution of differentially expressed genes in macrophages; and (H) differential expression changes of differential hub genes C1QA, C1QB, C1QC, and LYZ in TCGA tissue samples.

($p.adjust = 4.99 \times 10^{-40}$; enriched genes = 156), and neutrophil activation involved in immune response ($p.adjust = 4.99 \times 10^{-40}$, enriched genes = 154).

Moreover, organelle inner membrane ($p.adjust = 2.72 \times 10^{-43}$, enriched genes = 158), mitochondrial inner membrane ($p.adjust = 2.71 \times 10^{-46}$, enriched genes = 154), and ribosome ($p.adjust = 1.11 \times 10^{-62}$, enriched genes = 130) were significantly enriched in cellular components (Figure 2B). Cell adhesion molecule binding ($p.adjust = 5.27 \times 10^{-46}$, enriched genes = 90), cadherin binding ($p.adjust = 4.34 \times 10^{-21}$, enriched genes = 91), and structural constituent of ribosome ($p.adjust = 1.34 \times 10^{-19}$, enriched genes = 114) were mainly

enriched in molecular function (Figure 2C). Pseudotime analysis indicated that neutrophil mediated immune pathway may play an essential role in the development and progression of gastric cancer.

GSEA analysis of neutrophil activation and changes in various cell populations

Using the ssGSEA algorithm, we calculated neutrophil activation pathway scores for each cell. Differential analysis revealed that neutrophil activation GSEA scores were significantly different between cancer and normal tissue ($p < 0.05$; Figure 2D). We also found no significant differences

in neutrophil activation and the cell cycle ($p > 0.05$; Figure 2E). Macrophages exhibited the most significant change in the neutrophil activation GSVA score among all cell populations (Figure 2F). Therefore, we selected macrophages as the critical regulatory cells. Based on the quantitative difference analysis of GSVA pathway, we further confirmed the important role of neutrophil mediated immunity pathway in gastric cancer, and the change of this pathway may be caused by macrophages.

Identification of hub genes in the macrophage population

Differentially expressed genes in macrophages were selected and analyzed by expression differential folds and corrected p -values. Genes with large expression differences and the most significant p -value were selected as candidate regulatory genes, namely, C1q A-peptide (C1QB), C1q B-peptide (C1QB), C1QC, and LYZ (Figure 2G). There were significant differences in the expression of C1QA, C1QB, and C1QC in the TCGA colon cancer dataset ($p < 0.05$), but no significant differences in LYZ (Figure 2H). C1QA, C1QB, and C1QC may be the hub genes for macrophages regulating neutrophil activation.

Pseudotime regulation gene enrichment analysis

Through pseudotime analysis, the most significant differential expression gene sets in different developmental states were obtained and enrichment analysis of each gene set was carried out. In the first stage of development, the main expressed genes were LILRB2, C1QC, and KB-1125A3.11, which were enriched in GO:0002521 leukocyte differentiation ($p = 1.53 \times 10^{-06}$; FDR = 1.62×10^{-03}), GO:0002250 adaptive immune response ($p = 4.60 \times 10^{-06}$; FDR = 1.93×10^{-03}), and GO:0030098 lymphocyte differentiation ($p = 5.45 \times 10^{-06}$; FDR = 1.93×10^{-03}) pathways. In the second stage, the main expressed genes were IDO1, PLA2G7, C1QB, and C1QA, and enriched in GO:0002682 regulation of immune system process (P -value = 3.07×10^{-06} ; FDR = 9.49×10^{-04}), GO:0050729 positive regulation of the inflammatory response ($p = 3.53 \times 10^{-06}$; FDR = 9.49×10^{-04}), and GO:0050727 regulation of the inflammatory response ($p = 4.11 \times 10^{-06}$; FDR = 9.49×10^{-04}) pathways. GPR171, LAG3, and NEDD9 were the main expressed genes in the third stage, which were enriched in GO:0042110 T cell activation (p value = 1.42×10^{-05} ; FDR = 1.18×10^{-02}), GO:0002682 regulation of immune system process ($p = 4.34 \times 10^{-05}$; FDR = 1.23×10^{-02}), and GO:0002250 adaptive immune response ($p = 5.63 \times 10^{-05}$; FDR = 1.23×10^{-02}) pathways. The fourth stage included CYCS, P4HB, RAB1A, and GPR171, which were enriched in GO:0045916 negative regulation of complement activation ($p = 2.83 \times 10^{-05}$; FDR = 2.30×10^{-02}) and GO:0002921

negative regulation of the humoral immune response ($p = 5.40 \times 10^{-05}$; FDR = 2.30×10^{-02}) pathways (Figure 3). Pseudotime analysis showed that C1QC critically regulates macrophages.

GSEA analysis of differential genes in TCGA colon cancer

The count dataset in GDC TCGA Colon Cancer (COAD) was downloaded from the UCSC Xena database (<https://gdc.xenahubs.net>), and 1685 differential genes by DEseq calculation were obtained (Figure 4A and B). GSEA analysis revealed that GSE22886_NEUTROPHIL_VS_DC_UP (ES = -0.46 , NES = -1.45 , p .adjust = 0.02), GSE22886_NAIVE_BCELL_VS_NEUTROPHIL_UP (ES = -0.45 , NES = -1.43 , p .adjust = 0.01) and GSE22886_NEUTROPHIL_VS_MONOCYTE_UP (ES = -0.45 , NES = -1.41 , p .adjust = 0.04) were significantly enriched in the C7 module (Figure 4D). Transcriptome analysis of TCGA gastric cancer tissues suggested that neutrophil activation plays a key role gastric cancer progression.

C1QC survival analysis and functional enrichment analysis

Survival analysis showed that only C1QC was strongly associated with poor long-term survival in colon cancer (Hazard Ratio = 1.49, 95% CI:1.01–2.2, $p = 0.037$) (Figures 4C and Supplementary Figure 1). According to the above results, we chose C1QC for further analysis. Association pathway analysis suggested negative regulation of macrophage differentiation ($p = 4.96 \times 10^{-04}$), negative regulation of granulocyte differentiation ($p = 3.97 \times 10^{-04}$), and synapse pruning ($p = 5.96 \times 10^{-04}$) were mainly enriched in C1QC (Figure 4E). In addition, classical antibody-mediated complement activation ($p = 1.20 \times 10^{-03}$), classical complement ($p = 1.20 \times 10^{-03}$), and complement ($p = 1.60 \times 10^{-03}$) pathways were similarly identified (Figure 4F). Drug target prediction suggested that palivizumab ($p = 4.38 \times 10^{-04}$) was most likely to interact with the C1QC protein structure (Figure 4G), which was also verified by molecular docking experiments. The binding score of palivizumab to C1QC protein was 21.44 (Figure 5A).

C1QC was highly expressed in colon cancer tissue and correlated with its poor prognosis

To verify the high expression of C1QC in colon cancer tissue and its correlation with OS of colon cancer patients, we detected the expression of C1QC in 93 colon cancer tissue and 87 non-cancer tissue. IHC results showed that the expression score of C1QC in colon cancer tissues was higher than that in non-cancer tissue (Figure 5B and C). In addition, the positive expression rate of C1QC in colon cancer tissue (IHC score ≥ 4) was also higher than that in non-cancer tissue (Figure 5D). Kaplan-Meier survival analysis showed

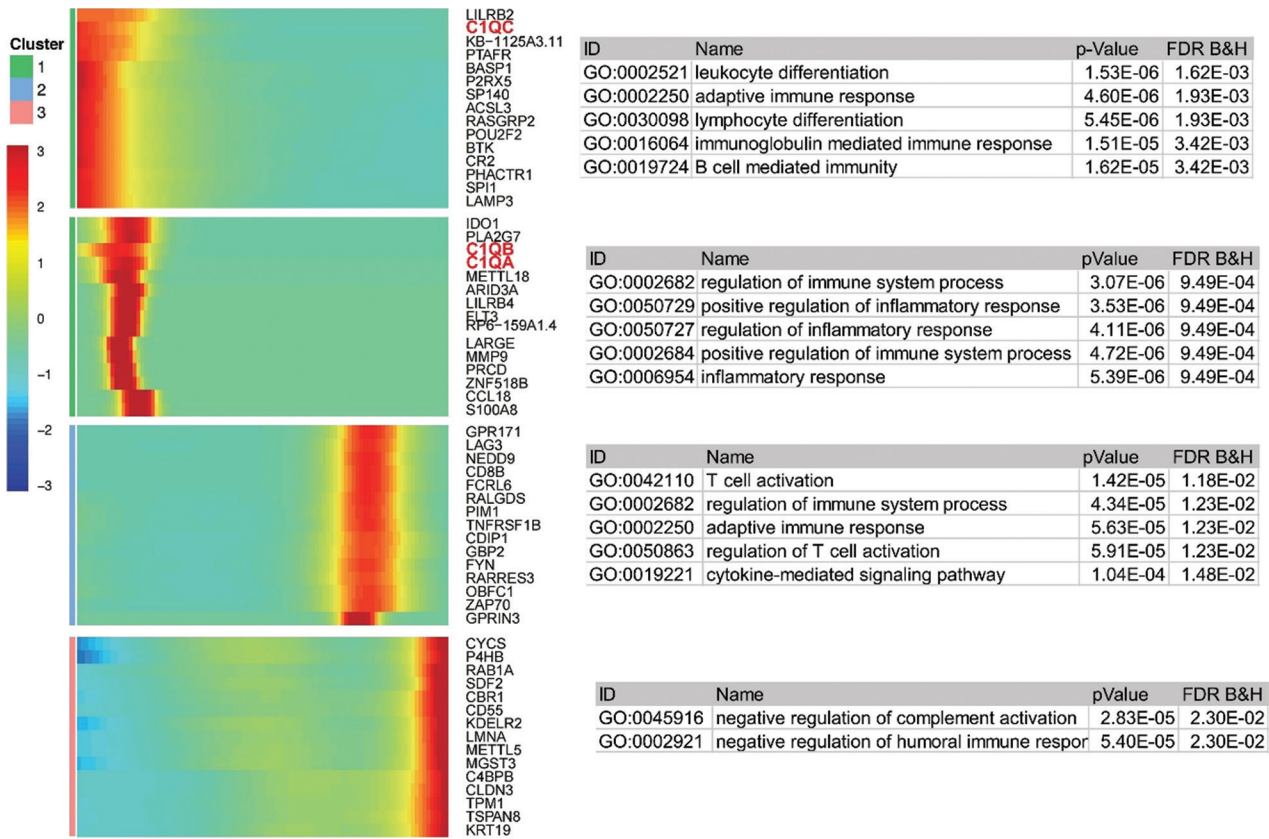


FIGURE 3. The pseudotime-related DEGs and functional enrichment for each cluster. Function enrichment analysis of pseudotime-related DEGs and biological processes.

that high C1QC expression was significantly correlated with poor OS (Figure 5E). As shown in Figure 5F, Table 1, and Supplementary Figure 2, C1QC expression in colon cancer was positively correlated with gender, lymph node metastasis, and the TNM stage. No significant correlation was found between C1QC expression and clinicopathological features, such as age, tumor size, primary tumor location, pathological type, or the positive expression rate of PDL. The expression of C1QC was higher in the colon cancer tissue of males, I+II (TNM stage), and lymph node metastasis patients. C1QC expression is closely related to the long-term prognosis of gastric cancer patients.

DISCUSSION

Immunotherapy has been further developed for different types of solid cancer, including colon cancer, after being successfully applied for melanoma. Patients with metastatic colon cancer can significantly prolong their survival with immunotherapy [19]. Identifying the biological functions of immune cell subsets and the related genes regulating their functional status is the key to cancer immunotherapy. Multiple scRNA-seq analysis studies showed differences in immune cell subsets across cancer types [9,20,21]. The differences in immunotherapy efficacy among different cancer are

related to the diversity of infiltration of immune cell subsets. At present, the function and prognostic value of immune cell subsets are mainly analyzed by combining different biomarkers of immune cells with immunohistochemical spatial localization [22,23].

In the present study, 4997 variable genes were identified in colon cancer cells and paracancer tissue cells. CD8+ T-cells, neutrocytes, macrophages, B-cells, and epithelial cells were identified by t-SNE and the UMAP dimensional reduction algorithm. Then, the differentiation trajectories of these cell populations were constructed, and crucial decision points of biological process were revealed. After the neutrophil activation GSVA analysis, macrophages displayed the greatest change in the neutrophil activation GSVA score. Studies have shown that the characteristics of macrophages depend on the tumor tissue source. Compared with CD8+ T cells, neutrophils, B cells, and epithelial cells, macrophages in colon cancer tissue have a more substantial neutrophil activation effect, but this action was not significantly correlated with the cell cycle [9,10,21].

In line with earlier research, but in contrast to the M1 and M2 dichotomous polarization of macrophages and the different continuity of tumor-associated macrophages in breast and lung cancer [9,24], our results also demonstrated that there was no significant difference between neutrophil activation

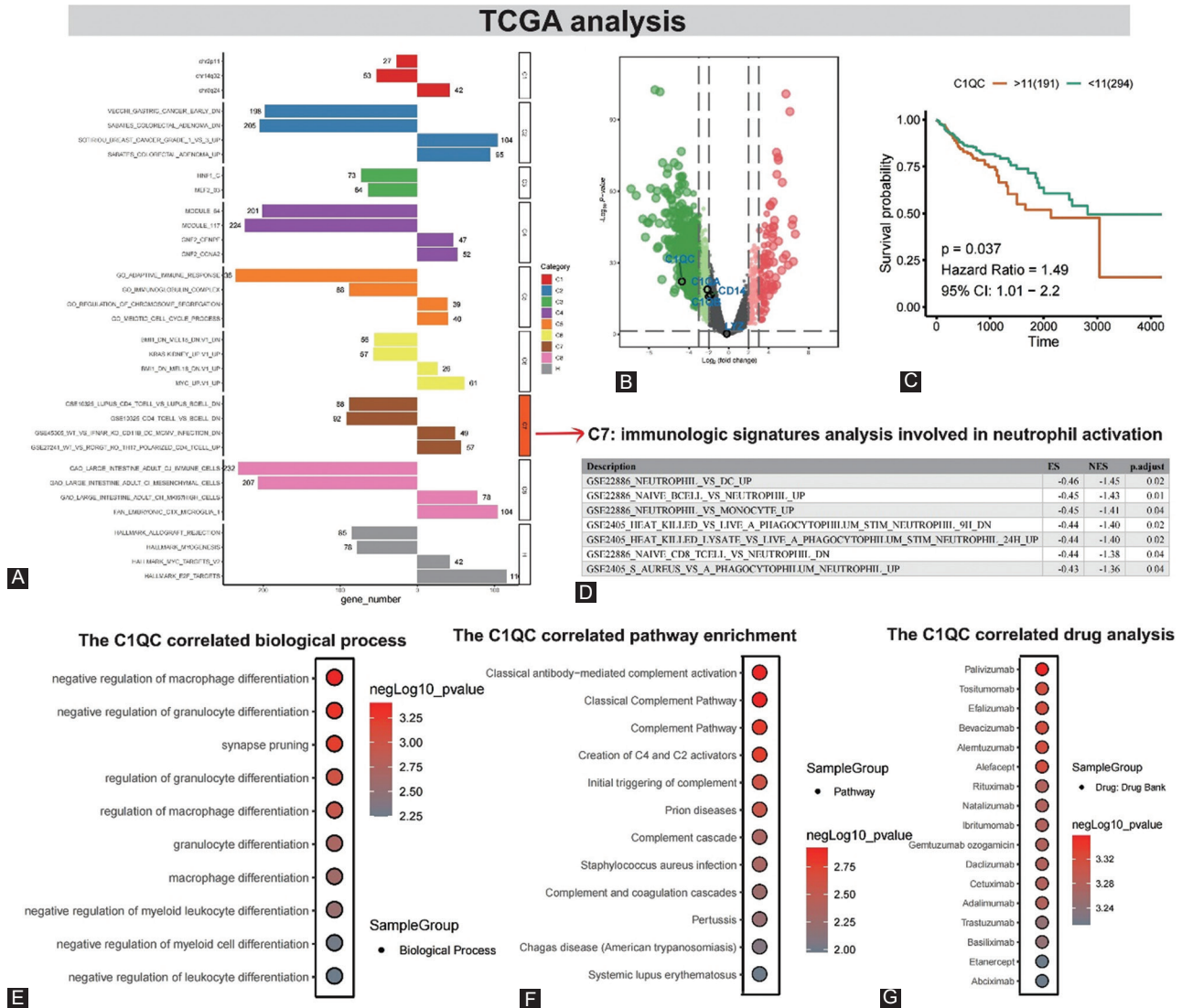


FIGURE 4. Identification of hub genes and their expression patterns in CRC cancer development in TCGA dataset. (A) The bar plot shows results of the GSEA analysis in CRC development; (B) the volcano map presents the CRC-related differential gene distribution; (C) the overall-survival analysis in response to C1QC based on the TCGA database. The survival curve is based on survival and survminer functions. The optimal cutoff values of multiple gene variables are calculated by an automatic cut-off algorithm, and the survival curve is drawn; (D) summary of GSEA analysis of TCGA differential genes: In C7, the immune signature analysis involved in neutrophil activation mainly enriched GSE22886 NEUTROPHIL VS DC UP, GSE22886 NAIVE BCELL VS NEUTROPHIL UP and GSE22886 NEUTROPHIL VS MONOCYTE UP core pathways. Further validation of neutrophil activation in the progression of colon cancer; and (E-G) the biological process, targeted drugs, and pathway enrichment analysis for C1QC.

and the cell cycle in colon cancer cells. Macrophages may have a different functional status in various tumor types, which needs to be clarified in further studies. However, compared with normal colon tissue, macrophages in colon cancer tissue showed obvious immune response functions, suggesting that macrophages are the critical cells among the infiltrated immune cells in colon cancer.

Based on the above results, we analyzed the hub genes that regulate the infiltration of the macrophage population in colon cancer. Finally, we selected complement C1Q-related molecules C1QA, C1QB, C1QC, and LYZ as candidate genes. Among them, the expression of C1QA, C1QB, and C1QC showed significant differences in the TCGA colon cancer dataset, while LYZ exhibited no significant difference.

Complement is a class of peptides with enzyme activity and self-regulation and the complement system regulates both innate and adaptive immune responses [25]. Complement C1q is produced by macrophages, immature dendritic cells, and mast cells. It is a subunit of complement C1 and regulates various immune cells. Under normal physiological conditions, the activation of complement C1q can effectively remove immune complexes, apoptotic cells and residual bodies, and maintain the immune response level in balance [26]. Studies have shown that complement C1q has a dual regulatory effect on macrophage activation and polarization [25] and promote liver metastasis of pancreatic ductal adenocarcinoma and progression of liver cancer [27,28]. In addition, tumor cells can hijack macrophages to produce complement C1q promoting

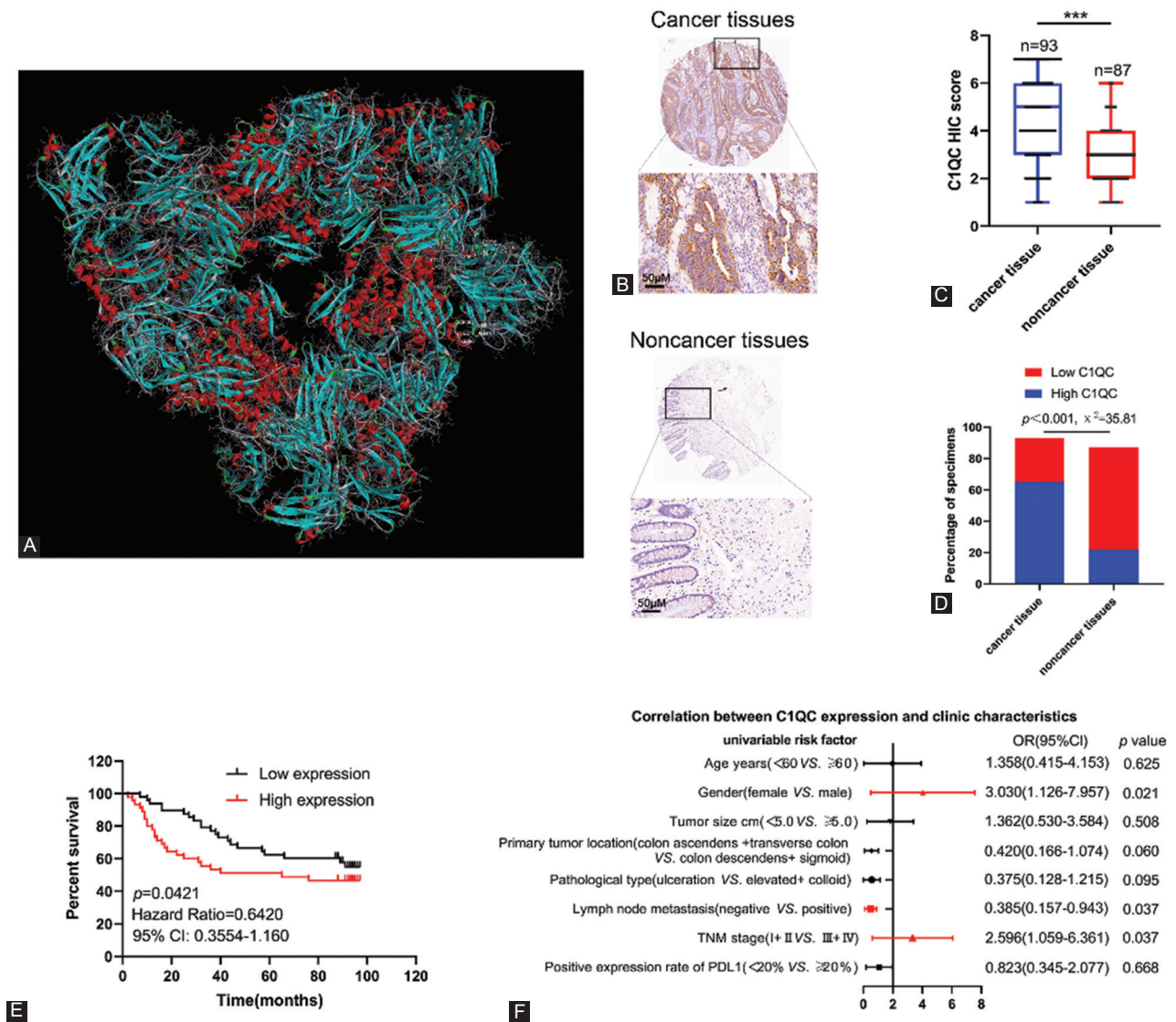


FIGURE 5. Expression levels of C1QC in colon cancer tissue and non-cancer tissue and its relationship with clinical characteristics. (A) Palivizumab docking with C1QC; the three-dimensional diagram shows the binding conformation of palivizumab (red) with C1QC (azure); (B) representative immunohistochemistry staining of C1QC expression in colon cancer tissue and non-cancer tissue; (C and D) statistical analysis of immunohistochemical scores of C1QC, and percentage of high expression and low expression of C1QC in colon cancer tissue and non-cancer tissue; (E) Kaplan-Meier survival analysis of C1QC expression in patients with colon cancer; and (F) correlations of C1QC expression levels in colon cancer tissue and clinicopathological features of colon cancer patients. The statistical significance was determined by the χ^2 -test. Data are expressed as the mean \pm SD.

tumor growth [29]. In terms of C1q-related molecules, C1QC+ tumor-associated macrophages can predict the prognosis, tumor stage, and immune cell infiltration of patients with cervical cancer, which helps treat cervical cancer [30]. The differential expressions of C1QA, C1QB and C1QC were found to be closely related to the survival prognosis, pathological features, and tumor microenvironment of osteosarcoma, which can help improve the clinical prognosis and immunotherapy [31]. The evidence suggests that C1q and its related molecules play an essential roles in macrophage activation and tumor development.

To explore the effects of C1QA, C1QB, and C1QC on macrophages in colon cancer, we conducted pseudotime analysis. The results demonstrated that among the candidate

genes, C1QC was involved in white blood cells and lymphocyte differentiation and the adaptive immune response in the early stages, while C1QA and C1QB regulated immune system processes and inflammatory responses in the second stage. These two stages are also crucial in the biological roles of macrophages in the development of colon cancer. Therefore, it is speculated that the expression levels of C1QA, C1QB, and C1QC are closely related to the development of colon cancer. We also analyzed the relationship between C1QA, C1QB, C1QC and the survival of colon cancer patients, and the results showed that a high expression of C1QC in colon cancer patients led to a poor prognosis ($p < 0.05$). However, the apparent relationship between the high expression of C1QA and C1QB and the prognosis of colon cancer patients still

TABLE 1. Correlation of the expression of C1QC in colon cancer with clinicopathologic features

Characteristics	No. of patients	C1QC		<i>p</i> value
		High expression <i>n</i> (%)	Low expression <i>n</i> (%)	
Overall	93	65 (69.9)	38 (30.1)	
Age				0.625
<60 years, <i>n</i> (%)	16 (17.2)	12 (75.0)	4 (25.0)	
≥60 years, <i>n</i> (%)	77 (82.8)	53 (68.8)	24 (31.2)	
Gender				0.021
female	45 (48.4)	37 (82.3)	8 (17.7)	
male	48 (51.6)	29 (60.4)	19 (39.6)	
Tumor size				0.508
<5.0 cm	38 (58.5)	28 (73.7)	10 (26.3)	
≥26. cm	55 (41.5)	37 (67.3)	18 (32.7)	
Primary tumor location				0.060
Colon ascendens+transverse colon	46 (49.5)	28 (60.9)	18 (39.1)	
Colon descendens+sigmoid	47 (50.5)	37 (78.7)	10 (21.3)	
Pathological type				0.095
Ulceration	69 (74.2)	45 (65.2)	24 (34.8)	
Elevated+colloid	24 (25.8)	20 (83.3)	4 (16.7)	
Lymph node metastasis				0.037
Negative	35 (37.6)	20 (57.1)	15 (42.9)	
Positive	58 (62.4)	45 (77.6)	13 (22.4)	
TNM stage				0.037
I+II	58 (62.4)	45 (77.6)	13 (22.4)	
III+IV	35 (37.6)	20 (57.1)	15 (42.9)	
Positive expression rate of PDL1				0.668
<20%	50 (53.8)	34 (68.0)	16 (32.0)	
≥20%	43 (46.2)	31 (72.1)	12 (27.9)	

needs to be further explored as statistical significance was not reached, which may be due to insufficient sample size. Future studies should expand the sample size to address this issue.

Since the current analysis was based on TCGA bulk transcriptome data, this inevitably led to some limitations of the findings. To further verify the result, we analyzed the gene characteristics of C1QC in 93 clinical specimens of colon cancer patients and 87 non-cancer tissues. IHC results showed that C1QC was highly expressed in colon cancer tissues, and the expression of C1QC in colon cancer was positively correlated with gender, lymph node metastasis, and the TNM stage, but there was no significant correlation with age, tumor size, primary tumor site, pathological type, or the PDL1-positive expression rate. Furthermore, the results illustrated the poor prognosis of colon cancer patients with high expression of C1QC. In addition, the drug target prediction and molecular docking experiments showed that palivizumab is the most likely prospective drug to target C1QC. The drug target docking analysis found that palivizumab can bind effectively to the C1QC protein.

CONCLUSION

scRNA-seq analysis and validation of clinical samples showed that C1QC is a potential diagnostic biomarker and immunotherapy target for colon cancer patients. Clarifying

the role of macrophage subsets characterized by the C1QC gene in the development of colon cancer, as well as the mechanisms of C1QC regulating immune infiltration of macrophage subsets will be essential. Further research is also required to determine the role of palivizumab as a C1QC targeted drug for colon cancer therapy. Although these issues remain to be addressed, our findings help identify new diagnostic and therapeutic approaches for colon cancer patients.

REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68(6):394-424. <https://doi.org/10.3322/caac.21492>
- Kuipers EJ, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, et al. Colorectal cancer. *Nat Rev Dis Primers* 2015;1:15065. <https://doi.org/10.1038/nrdp.2015.65>
- Dekker E, Tanis PJ, Vleugels JL, Kasi PM, Wallace MB. Colorectal cancer. *Lancet* 2019;394(10207):1467-80. [https://doi.org/10.1016/S0140-6736\(19\)32319-0](https://doi.org/10.1016/S0140-6736(19)32319-0)
- Frankenfeld CL, Menon N, Leslie TF. Racial disparities in colorectal cancer time-to-treatment and survival time in relation to diagnosing hospital cancer-related diagnostic and treatment capabilities. *Cancer Epidemiol* 2020;65:101684. <https://doi.org/10.1016/j.canep.2020.101684>
- Cassetta L, Fragkogianni S, Sims AH, Swierczak A, Forrester LM, Zhang H, et al. Human tumor-associated macrophage and monocyte transcriptional landscapes reveal cancer-specific reprogramming, biomarkers, and therapeutic targets. *Cancer Cell* 2019;35(4):588-602.

- <https://doi.org/10.1016/j.ccell.2019.02.009>
- [6] Fridman WH, Pagès F, Sautès-Fridman C, Galon J. The immune contexture in human tumours: Impact on clinical outcome. *Nat Rev Cancer* 2012;12(4):298-306. <https://doi.org/10.1038/nrc3245>
- [7] Zhou R, Zhang J, Zeng D, Sun H, Rong X, Shi M, et al. Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage I-III colon cancer. *Cancer Immunol Immunother* 2019;68(3):433-42. <https://doi.org/10.1007/s00262-018-2289-7>
- [8] Papalexi E, Satiya R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol* 2018;18(1):35-45. <https://doi.org/10.1038/nri.2017.76>
- [9] Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, et al. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. *Cell* 2018;174(5):1293-308.e36. <https://doi.org/10.1016/j.cell.2018.05.060>
- [10] Chevrier S, Levine JH, Zanotelli VR, Silina K, Schulz D, Bacac M, et al. An immune atlas of clear cell renal cell carcinoma. *Cell* 2017;169(4):736-49. <https://doi.org/10.1016/j.cell.2017.04.016>
- [11] Zeng ZL, Li WP, Zhang D, Zhang C, Jiang X, Guo R, et al. Development of a chemoresistant risk scoring model for prechemotherapy osteosarcoma using single-cell sequencing. *Front Oncol* 2022;12:893282. <https://doi.org/10.3389/fonc.2022.893282>
- [12] DeNardo DG, Ruffell B. Macrophages as regulators of tumour immunity and immunotherapy. *Nat Rev Immunol* 2019;19(6):369-82. <https://doi.org/10.1038/s41577-019-0127-6>
- [13] Butowski N, Colman H, De Groot JF, Omuro AM, Nayak L, Wen PY, et al. Orally administered colony stimulating factor 1 receptor inhibitor PLX3397 in recurrent glioblastoma: An Ivy foundation early phase clinical trials consortium phase II study. *Neuro Oncol* 2016;18(4):557-64. <https://doi.org/10.1093/neuonc/nov245>
- [14] Cassier PA, Italiano A, Gomez-Roca CA, Le Tourneau C, Toulmonde M, Cannarile MA, et al. CSF1R inhibition with ematuzumab in locally advanced diffuse-type tenosynovial giant cell tumours of the soft tissue: A dose-escalation and dose-expansion phase 1 study. *Lancet Oncol* 2015;16(8):949-56. [https://doi.org/10.1016/s1470-2045\(15\)00132-1](https://doi.org/10.1016/s1470-2045(15)00132-1)
- [15] Lee JH, Chen TW, Hsu CH, Yen YH, Yang JC, Cheng AL, et al. A phase I study of pexidartinib, a colony-stimulating factor 1 receptor inhibitor, in Asian patients with advanced solid tumors. *Invest New Drugs* 2020;38(1):99-110. <https://doi.org/10.1007/s10637-019-00745-z>
- [16] Papadopoulos KP, Gluck L, Martin LP, Olszanski AJ, Tolcher AW, Ngarmchamnanrith G, et al. First-in-human study of AMG 820, a monoclonal anti-colony-stimulating factor 1 receptor antibody, in patients with advanced solid tumors. *Clin Cancer Res* 2017;23(19):5703-10. <https://doi.org/10.1158/1078-0432.ccr-16-3261>
- [17] Ries CH, Cannarile MA, Hoves S, Benz J, Wartha K, Runza V, et al. Targeting tumor-associated macrophages with anti-CSF-1R antibody reveals a strategy for cancer therapy. *Cancer Cell* 2014;25(6):846-59. <https://doi.org/10.1016/j.ccr.2014.05.016>
- [18] Deng H, Huang L, Liao Z, Liu M, Li Q, Xu R. Itraconazole inhibits the Hedgehog signaling pathway thereby inducing autophagy-mediated apoptosis of colon cancer cells. *Cell Death Dis* 2020;11(7):539. <https://doi.org/10.1038/s41419-020-02742-0>
- [19] Kishore C, Bhadra P. Current advancements and future perspectives of immunotherapy in colorectal cancer research. *Eur J Pharmacol* 2020;893:173819. <https://doi.org/10.1016/j.ejphar.2020.173819>
- [20] Zilionis R, Engblom C, Pfirschke C, Savova V, Zemmour D, Saatioglu HD, et al. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* 2019;50(5):1317-34.e10. <https://doi.org/10.1016/j.immuni.2019.03.009>
- [21] Lambrechts D, Wauters E, Boeckx B, Aibar S, Nittner D, Burton O, et al. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med* 2018;24(8):1277-89. <https://doi.org/10.1038/s41591-018-0096-5>
- [22] Huang YK, Wang M, Sun Y, Di Costanzo N, Mitchell C, Achuthan A, et al. Macrophage spatial heterogeneity in gastric cancer defined by multiplex immunohistochemistry. *Nat Commun* 2019;10(1):3928. <https://doi.org/10.1038/s41467-019-11788-4>
- [23] Schürch CM, Bhatt SS, Barlow GL, Phillips DJ, Noti L, Zlobec I, et al. Coordinated cellular neighborhoods orchestrate antitumoral immunity at the colorectal cancer invasive front. *Cell* 2020;182(5):1341-59.e19. <https://doi.org/10.1016/j.cell.2020.07.005>
- [24] Müller S, Kohanbash G, Liu SJ, Alvarado B, Carrera D, Bhaduri A, et al. Single-cell profiling of human gliomas reveals macrophage ontogeny as a basis for regional differences in macrophage activation in the tumor microenvironment. *Genome Biol* 2017;18(1):234. <https://doi.org/10.1186/s13059-017-1362-4>
- [25] Bohlson SS, O'Conner SD, Hulsebus HJ, Ho MM, Fraser DA. Complement, c1q, and c1q-related molecules regulate macrophage polarization. *Front Immunol* 2014;5:402. <https://doi.org/10.3389/fimmu.2014.00402>
- [26] Bajic G, Degn SE, Thiel S, Andersen GR. Complement activation, regulation, and molecular basis for complement-related diseases. *EMBO J* 2015;34(22):2735-57. <https://doi.org/10.15252/embj.201591881>
- [27] Yang J, Lin P, Yang M, Liu W, Fu X, Liu D, et al. Integrated genomic and transcriptomic analysis reveals unique characteristics of hepatic metastases and pro-metastatic role of complement C1q in pancreatic ductal adenocarcinoma. *Genome Biol* 2021;22(1):4. <https://doi.org/10.1186/s13059-020-02222-w>
- [28] Lee JH, Poudel B, Ki HH, Nepali S, Lee YM, Shin JS, et al. Complement C1q stimulates the progression of hepatocellular tumor through the activation of discoidin domain receptor 1. *Sci Rep* 2018;8(1):4908. <https://doi.org/10.1038/s41598-018-23240-6>
- [29] Roumenina LI, Daugan MV, Noe R, Petitprez F, Vano YA, Sanchez-Salas R, et al. Tumor cells hijack macrophage-produced complement C1q to promote tumor growth. *Cancer Immunol Res* 2019;7(7):1091-105. <https://doi.org/10.1158/2326-6066.cir-18-0891>
- [30] Li X, Zhang Q, Chen G, Luo D. Multi-omics analysis showed the clinical value of gene signatures of C1QC+ and SPP1+ TAMs in cervical cancer. *Front Immunol* 2021;12(2663):694801. <https://doi.org/10.3389/fimmu.2021.694801>
- [31] Chen LH, Liu JF, Lu Y, He XY, Zhang C, Zhou HH. Complement C1q (C1qA, C1qB, and C1qC) May Be a potential prognostic factor and an index of tumor microenvironment remodeling in osteosarcoma. *Front Oncol* 2021;11:642144. <https://doi.org/10.3389/fonc.2021.642144>

Related Articles Published In BJBMS

1. Development and validation of a ferroptosis-related lncRNAs prognosis signature in colon cancer

Hua-jun Cai et al., BJBMS, 2021

2. MALAT1 inhibits the Wnt/ β -catenin signaling pathway in colon cancer cells and affects cell proliferation and apoptosis

Junjun Zhang et al., BJBMS, 2020

3. The role of the Hedgehog signaling pathway in cancer: A comprehensive review

Ana Marija Skoda et al., BJBMS, 2018