Systematic analysis and experimental validation of the prognostic and immunological effects of SPP1 tumor-associated macrophage features in colorectal cancer

Yizhang Li  
Zhongnan Hospital of Wuhan University

Zhenwei Yang  
Zhongnan Hospital of Wuhan University

Jialong Liu  
Zhongnan Hospital of Wuhan University

Haizhou Wang  
Zhongnan Hospital of Wuhan University

Fan Wang  
Zhongnan Hospital of Wuhan University

Hailin Zhang  
Zhongnan Hospital of Wuhan University

Qiu Zhao  
Zhongnan Hospital of Wuhan University

Jing Liu (liujing_GI@whu.edu.cn)  
Zhongnan Hospital of Wuhan University

Research Article

Keywords: colorectal cancer, tumor-associated macrophages, EMT, angiogenesis, SPP1, immunotherapy

Posted Date: February 7th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2550089/v1

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

Purpose

Tumor associated macrophages (TAM) influence colorectal cancer (CRC) development, and their clinical significance has been widely established. We intend to depict a full macrophage landscape in order to increase our understanding of CRC heterogeneity and give improved precision medicine techniques.

Methods

Use Seurat and Cellchat to conduct single cell analysis on GSE178341 to determine the interaction between cells and understand the influence of core cell subsets on immune response. SsGSEA was used to quantify the immune related cells of TCGA patients and further cluster them into subtypes. The effectiveness of combined COX and LASSO, SPP1 TAM characteristics in predicting prognosis was validated in several GEO datasets.

Then, Cell line culture and Quantitative real-time PCR were used to validate the hub genes of SPP1 TAM features.

Results and Conclusion

To summarize, we built a more comprehensive macrophage atlas to highlight the wide range and heterogeneity of macrophages present in people at various MMR stages. SPP1 TAM is not only enriched in dMMR patients, but also shows two characteristics of immune response, which may explain the reason why some dMMR patients have poor response to immunotherapy. The prognosis model constructed by Hub DEG SPP1 related to it has different responses to immune response and chemotherapy drugs, which provides new clues to inhibit the potential efficacy of SPP1 TAM.

Introduction

Colorectal cancer is the most frequent gastrointestinal cancer and one of the top causes of cancer-related deaths worldwide. [1] This malignancy has a low survival rate and a terrible prognosis. As a result of advances in screening techniques and targeted cancer therapy, CRC incidence and mortality rates have decreased[2]. Despite this, the prognosis of patients with CRC remains an urgent concern owing to the heterogeneity of CRC.

Cancer immunotherapies have altered cancer therapy over the last decade by leveraging the body's immune system to combat the tumor. [3]. In spite of recent advances in immunotherapy for cancer, the effectiveness of these strategy remains limited. Immune checkpoint blockade(ICBs) medications have shown considerable therapeutic advantages for the PD-1/PD-L1 axis[4], but PD-1 blockade has not been proven to be as effective against colorectal cancer as previously thought[5]. Colorectal cancer shows different immune responsiveness between dMMR and pMMMR in cancer immunotherapies. According to
several studies, macrophage infiltration induces varied endpoints in CRC individuals with distinct MMR stages. [6].

Tumor-associated macrophages (TAMs) are regarded to be an important component of the body's immune response, as well are thought to be a prominent component of Tumor Microenvironment (TME), which is vital in tumor formation, invasion, and migration[7]. Among several different malignancies, TAM is classified as "traditional" M1 and M2, or, in other words, pro-inflammatory and anti-inflammatory. [8]. TAM in colorectal cancer, on the other hand, demonstrates significant dichotomy and has been classified as SPP1 TAM and C1QC TAM [9]. Furthermore, the particular roles of SPP1 TAM and C1QC TAM in colorectal cancer cannot be determined only by binary studies such as pro-inflammatory or anti-inflammatory phenotypes.

Notably, SPP1 TAM is specifically enriched for epithelial mesenchymal transition (EMT), angiogenesis, hypoxia and many other pathways associated with tumour progression and metastasis. And their changes in different MMR states still need to be further explored[9].

Macrophages have different phenotypes and functions in different stages of colorectal cancer and other types of cancer, showing dynamic changes[9–11]. It has a variety of impacts on cancerous cells' potential to multiply, infiltrate, and spread. To differentiate subgroups of CRC TAM, single-cell sequencing technologies must be used.

The macrophages heterogeneity in tumors obtained from individuals with colorectal cancer with varying MMR standings was examined at the single-cell level throughout this research. We found significant differences in SPP1 TAM in different MMR states and determined that its marker SPP1 has a significant impact in phenotyping. Based on the level of immune infiltration of SPP1 TAM, we classified all CRC patients into two subtypes and identified the differences between the two subtypes of expression genes. We created a methodology for the characteristic genes linked with SPP1 TAM in order to build a validated prognostic model for TAM. COX and LASSO examined these genes in order to establish and verify risk stratification signatures, which may play a significant influence on prognosis prediction and therapy efficacy monitoring.

Materials And Methods

Data Collection

To evaluate TAM function and alterations at the single-cell level, we had used single-cell sample GSE178341 [6]. In this data, we isolated all tumor tissue cells (containing 34 patients with dMMR status as well as 28 CRC patients with pMMR status). Cells with less than 100,000 transcripts or fewer than 200 genes were filtered out of all CRC single cells. Mitochondrial and ribosomal genes were eliminated. Furthermore, only genes expressed in at least 10 single cells at a level greater than one were included for downstream analysis.
Furthermore, using the TCGAbiolinks R package[12], we collected the bulk Rnas sequence dataset for colon and rectal cancers in TCGA. They were incorporated into TCGA-CRC. TPM format was utilized for future analysis. Because they had comprehensive overall survival (OS) information, CRC samples from these files (TCGA-CRC, GSE29621[13], GSE38832[14], GSE17537[15], GSE72970[16]) were utilized to develop and validate our signature. We evaluated individuals with an overall survival of more than one month to decrease the influence of unanticipated, non-disease variables on prognosis. In addition, for future reference, we retrieved somatic mutation data from CRC patients using the TCGAbiolinks R program.

**Clustering of colorectal cancer cells**

To analyze the single-cell data, we utilized the R program Seurat, as with most other options. [17]. While downloading single cell data from several centers and patients from various geographies, we also utilized the Harmony R package to eliminate various batch effects from the single cell data[18]. We next sorted all cells into six subgroups after performing an initial post-dimensional reduction clustering. Following that, we chose cell types in the myeloid lineage that could be recognized as macrophages and retrieved all of their macrophages for further clustering analysis. We discovered that adjusting the resolution parameter of the FindClusters function to 0.1 and 0.2 provided us the appropriate results in these two clustering phases.

**Differentially Expressed Genes Identification and Function Analysis for TAM**

We used the Seurat R tool as part of the standard analysis to identify marker genes in each macrophage subset. Following that, we used the same technique to identify differentially expressed genes in SPP1 TAM, a subset of macrophages in various MMR states. All absolute values of logfoldchange > 0.25, while concurrently corrected for a P value of 0.05, were our screening criterion for discovering differential genes in subpopulations in the preceding phases. In order to explore the diverse roles of distinct macrophage subsets, we utilized the GSVA approach to measure the activity of different macrophage subpopulations in different pathways[19]. In our GSVA research, we took the Hallmark gene set, a classic dataset available from the MSigDB database, and then we attempted to analyze the different pathway scores in different MMR states using the limma R tool. Simultaneously, we utilized ClusterProfiler enrichGO function to undertake GO pathway enrichment analysis of the differential genes acquired in different MMR states, selecting those pathways of relevance to us with adjusted P values less than 0.05 for further demonstration[20].

**Cell abundance changes**

We compared the variations in cell abundance across the two MMR states using the scRNAtoolVis R software package. In parallel, we examined differences in cell abundance across the groups (tumour stage and nodal stage).

**Trajectory of cell differentiation**
To deduce the key differentiation trajectories for SPP1 TAM, we applied the 'slingshot' technique in the R package dyno\[21\]. We inferred the direction of macrophage development trajectories based on genetic alterations that characterize The development of macrophages using marker genes from six macrophage subpopulations, including SPP1 TAM.

**Intercellular communication**

CellChat, a R tool, was used to analyze intercellular interactions\[22\]. We analyzed entire humans ligand-receptor results in CellChatDB and visualized ligand-receptors linked with EMT as well as angiogenesis.

**The pivotal genes in the pathway**

We used lasso regression using the R package glmnet to identify genes linked with the EMT or angiogenic pathways for SPP1 TAM in various MMR stages. Pathway scores and differential genes in distinct MMR states were employed as dependent and independent variables in the following study. The acquired genes were intersected to get crucial genes involved in various MMR states.

**Immune Response Prediction about SPP1 TAM**

We used the R package EaSleR to construct immunotherapy reaction ratings for TCGA-CRC individuals\[23\]. This R program calculates the probability of an immune reaction primarily by computing a weighted score for tumor mutation burden (TMB) and immunological micro-environment. The higher the relative score, the greater the likelihood of an immunotherapy response. We sought to characterize the various pMMR statuses based on the immunotherapy response score and compared the extent of SPP1 TAM infiltration to the different MMR status in TCGA data.

**Analysis of different TAM subtypes**

Using the ConsensusClusterPlus R package\[24\], we performed consensus clustering in TCGA-CRCs based on the differential genes between dMMR and pMMR in the TAM single-cell data (log2-fold change > 0.25 and FDR < 0.05). We then defined the most appropriate classification into two subtypes based on CDF graphs and heatmaps of the consensus cluster.

**TAM scoring in CRC patients**

From earlier single-cell investigations, we were able to identify marker genes for representative TAMs\[25\]. Based on these marker genes, we were able to use the 'ssGSEA' method in the GSVA R package to determine the activity of various TAMs in CRC samples. In order to confirm the stability and robustness of the results, we also selected a number of different algorithms summarized in the IOBR R package to evaluate (e.g. IPS, xCell, MCP, TIMER, ESTIMATE, EPIC, quanTIseq, etc.)\[26\].

**Creation of the riskScore model and signatures connected to TAM**

We obtained differentially expressed genes based on the two subgroups using the "limma" R package\[27\] and differential genes were filtered utilizing the following standards: log2 foldchange > 0.585 and False
discovery rate (FDR) < 0.05.

An individual tumor's risk score is computed to evaluate the prognostic significance of TAM in CRC. One-way Cox analysis of differentially expressed genes using follow-up data from the TCGA-CRC cohort to determine genetic signatures associated with TAM. Then, TCGA-CRC was used as the training group and other CRC cohorts (GSE17537, GSE29621, GSE38832, and GSE72970) were used as the test group. Following that, we utilized the Lasso algorithm (1000 random simulations to avoid overfitting) to further pick critical features for the genes in the training set for which prognostic correlations were established using the one-factor COX approach. Finally, risk models were built using the stepwise COX method findings, and risk scores were calculated using the "survminer" R package to analyze the influence of risk score prognostic prediction using KM survival analysis and the construction of receiver operating characteristic (ROC) curves.

**Role of TAM-related signatures in response to anti-PD-1/L1 immunotherapy**

To predict the effect of TAM-related signatures in immunotherapy response, we systematically collected immunotherapy-related cohorts and downloaded anti-PD-L1 cohorts (IMvigor210 study, GSE78220, GSE179351)[28–30].

We estimated semi-inhibitory concentration (IC50) ratings for chemotherapy medications routinely used to treat CRC utilizing "pRRophetic" R package to investigate variations in responsiveness to CRC chemotherapy[31].

**Cell line culture and Quantitative real-time PCR (RT-PCR)**

The following human colon cancer cell lines: SW620, HT29, SW480, LOVO, HCT116 and the normal cell line NCM460 bought from China Center for Type Culture Collection (Wuhan, CN). SW620 cells were culture in L15 (HyClone, USA), LOVO cells were culture in F-12K (HyClone, USA), SW480 cells were cultured in DMEM (HyClone, USA) and HT29, HCT116, NCM460 cells were maintained in RPMI 1640 (HyClone, USA), with 10% foetal bovine serum (FBS) from Gibco (USA) and 100 µg/ml streptomycin/penicillin (Hyclone, USA), in a 5% CO2 environment at 37°C

Samples of nine cancerous and matched adjacent normal tissues were collected from CRC patients of the Zhongnan Hospital of Wuhan university. The Ethics Committee of the Zhongnan Hospital of Wuhan university granted permission to perform the research and every participant availed consent. The total RNA from the frozen colorectal tissues and colorectal cancer cells were performed by TRIzol reagent (Invitrogen, USA) in accordance with the manufacturer's guidelines, and 1µg of total RNA was employed to produce complementary DNA (cDNA), utilizing a RevertAid First Strand cDNA synthesis kit (Thermo Scientific, USA) following the manufacturer's specifications. On a Roche LightCycler 96, RT-PCR was carried out utilizing an UltraSYBR mixture (CWBio, China) under normal PCR conditions. Primer sequences are described in the Table S3.
Statistical analysis

The R (ver 4.2.1) program was used to compute and visualize all of the data. Results with adjusted p-values of 0.05 or FDRs of 0.05 were deemed statistically meaningful.

Results

SPP1 TAM subsets in dMMR colorectal cancer tumor tissues have a higher rate of aggression compared to pMMR colorectal cancer

First, we confirmed macrophage diversity in the malignant immunology milieu by examining TAM subtypes, marker gene expression, and functional alterations, by dividing the cells into 6 different clusters via Seura packaging, we can see no significant change in the proportion of myeloid cells in dMMR and pMMR (Fig. 1A). After that, we separated all myeloid cells from six distinct clusters, conducted uniform flow approximation and projection (abbreviated UMAP, a downscaling approach) of macrophages in myeloid cells, and showed them based on representative markers.(Fig. 1B, 1C). Specifically, all cell clusters are essentially highly expressed with SPP1 and CIQC (Fig. 1D), while overexpression of classically labeled SPP1 is labeled SPP1 + TAM. Overexpression of the classically labeled C1Q gene is labeled C1QC + TAM. TAMs with substantial binary categorization, dubbed SPP1 TAM and C1QC TAM, were discovered in Lei’s work. SPP1 TAM and C1QC TAM are also distinct from "classically activated" M1 and "alternately activated" M2 macrophages. It has been demonstrated that the infiltration of macrophages in individuals is heavily impacted by the patient’s MMR status. TAM enrichment is reported to be linked to tumor growth and metastasis. We evaluated macrophage abundance in different MMR states and discovered that the presence of most TAM subgroups was significantly varied (Fig. 1E). We can see that the ratio of SPP1 TAM and C1QC TAM has altered in dMMR and pMMR, with SPP1 TAM being the most apparent. This might imply that SPP1 TAM and C1QC TAM play different roles in various colorectal cancer states, hence we chose SPP1 TAM and C1QC TAM for further investigation. Overall, SPP1 TAMs and C1QC TAMs are broadly dispersed in various MMR states with varying genetic activation levels (Fig. 1F), as well as their infiltration and function require further investigation.

We performed gene set variation analysis on published characteristic gene sets to determine the functional activity of distinct cell types (Fig. 2A). Most functional pathways' activity in SPP1 TAM changed significantly, with the epithelial-mesenchymal transition (EMT) and angiogenic pathways with increased activity in SPP1 TAM being more significant in dMMR than pMMR. The C1QC TAM did not significantly upregulate in most pathways. Previous research has shown that C1QC TAM has increased immunological activity and reduced tumor advancement activities, showing that they do not induce malignancies. SPP1 TAM has a high angiogenesis and EMT activity, revealing that it contains tumor-promoting capability. This might explain the considerable variability in partly immunopathological activity in various TAM tissues. Unsurprisingly, SPP1 TAM in a higher proportion of dMMR patients during cancer progression than pMMR (Fig. 2B), and the fundamental explanation of this phenomena has to be investigated further.
At the same time, we also explored the dynamics of SPP1 TAM and C1QC TAM cell type and expression characteristics, inferred the differentiation trajectories of six macrophage subsets (Fig. 3C), and mapped the maker genes in developmental order. The trajectory begins with the C1QC TAM and progresses via an intermediate stage of unclassified TAM to the SPP1 TAM (Fig. 3A, 3B). However, SPP1 TAM did not further differentiate into M2-like TAM, which may reflect that SPP1 TAM does not simply divide into M1 macrophages or M2 macrophages.

**SPP1 plays an important role in cellular communication in colorectal cancer**

We explored cell-to-cell communication of CRCs and analyzed receptor-ligand interactions via CellChat. The findings clearly reveal that macrophages employ SPP1 to communicate with all types of cells, particularly SPP1-CD44 interactions (Fig. 4A, 4B). The SPP1 signaling pathway was considerably elevated in both the dMMR and pMMR states, showing that it performed a crucial component in the evolution of colorectal cancer of various states. Recent research has revealed malignant cells may act as a bridge between HCC malignant cells and macrophages, promoting macrophage transformation, in which SPP1-CD44 plays an important role. Similar results were found in pancreatic cancer and lung squamous cell carcinoma[32–34], and may also be suitable for colorectal cancer.

**SPP1-induced hyperexpression of EMT and angiogenesis of SPP1 TAMs in dMMR**

To explore the pivot gene that causes SPP1 TAM in dMMR and pMMR, We conducted differential analysis and discovered 284 DEGs (Table S1). Among these, 129 DEGs were found to be considerably elevated in dMMR, whereas 216 DEGs were found to be significantly upregulated in pMMR (Fig. 4D). We found that they are also rich in EMT and angiogenesis pathways (Fig. 4E). Meantime, We present here in detail the varied pathway activities of SPP1 TAM in different MMR states based on the results of the previous GSVA research(Fig. 4C). EMT and angiogenesis pathways were significantly enriched in dMMR. Next, we further screen DEGs associated with EMT and angiogenesis using LASSO (Fig.S1A, S1B). A total of 13 genes are associated with EMT and angiogenesis of SPP1 TAM in dMMR and pMMR(Fig.S1C), of which SPP1 and VCAN are the two most important genes(Fig.S1A,S1B). Based on the above results, SPP1 and VCAN are significantly upregulated in dMMR, and these genes may play a significant influence in EMT and angiogenesis in dMMR.

**Relationship between SPP1 TAM infiltration and response to immunotherapy**

Based on the findings, SPP1 TAM is a cell with tumor-promoting capability, and its influence on the immunotherapy reaction has to be evaluated. In the TCGA-CRC data, we assessed the probability of individuals reacting to immunotherapy. We also looked at the manifestation of Immunological checkpoints genes on macrophage interfaces and discovered CTLA4 ligands CD80 and CD86 were
strongly expressed in a subset of SPP1 TAM cells, and CD80 was only highly expressed in SPP1 TAM, which represented differences in immune characteristics of different TAMs (Fig. 5A). To assess infiltration levels, marker genes associated with SPP1 TAM survival were used as the characteristic gene set. It was discovered to be substantial disparities in SPP1 TAM invasion and SPP1 overexpression in different MMR states (Fig. 5B, 5C), and the infiltration of SPP1 TAM and the expression of SPP1 were also positively correlated with reponse score and tmb (Fig. 5D). We discovered that patients with dMMR had higher response scores (Fig. 5E), revealing that they would be extra likely to receive benefit from immunotherapeutic, and that SPP1 TAM invasion rate in patients with dMMR colon cancer was positive and significantly associated with immunotherapeutic reaction, possibly implying that MMR patients with higher SPP1 TAM infiltration rates may respond to immunotherapy (Fig. 5F). However, our study shows that the level of infiltration of SPP1 TAM is poor in predicting the prognosis of CRC patients (Fig. 5G), so we need to build a prognostic model associated with SPP1 TAM to anticipate prognostic and therapeutic results in individuals with CRC.

**Molecular subtypes of colorectal cancer based on TAM**

In the TAM single cell data, we did an examination of consensus sorting depending on the various genes of dMMR and pMMR. Based on the CDF graphs and matrix heat maps of the consensus clusters, we divided the CRC samples into two categories. The immunological infiltration richness of the two consensus groupings (C1 and C2) differed substantially, with C2 having much greater total infiltration richness than C1 (Fig. 6A-D, Table S2). As a result, we classified C1 as 'non-immune' and C2 as 'immune'. In parallel, we employed various IOBR methods (such as IPS, xCell, MCP, TIMER, ESTIMATE, EPIC, quantIseq, and others) to validate the stability and robustness of molecular subtypes (Fig. 6E, G). KM survival analysis further showed that C2 patients had worse OS (Fig. 6F). However, the expression of ips and PD-L1 showed that C2 was insensitive to immunotherapy.

**Construction of TAM subtype risk scoring model**

We further evaluated the association between TCGA-CRC-based cohort prognosis and TAM. Based on the correlation with SPP1 and the correlation with uni-Cox analysis (p-value < 0.05), prognostic features associated with TAM were found. In order to avoid overfitting of prognostic features, the prognostic gene features with minimal λ were extracted by lasso regression (Fig. S2A, S2B), and the model was optimized by stepwise COX. The riskScore model was constructed using 17 prognostic traits genes (ATP6V1B2 + ADAM8 + P4HA1 + RAB3B + SCG2 + ASPHD2 + TIMP1 + DAPK1 + CPA3 + COMP + GADD45B + MMP1 + PCOLCE2 + CSRP2 + FECH + LXN + SPP1) using the predict function to calculate riskScore. The TCGA-CRC cohort was separated into high-risk and low-risk rating subgroups, with the median risk score generated from the survminer R tool serving as a cut-off value. Patients with high-risk scores had a shorter life expectancy than those with low-risk scores (Fig. 7A). Furthermore, the 1-, 3-, and 5-years riskScore rate of survival are noted by Area Under Curve (AUC) values of 0.79, 0.82, and 0.84, respectively (Fig. 7B). As to validate the riskscore model's stability, we tested the prognostic value of numerous external CRC cohorts. Our findings reveal that the risk model offers consistent and accurate prognostic
predictions across all CRC cohorts (Fig. 7C-L. The AUCs corresponding to 1-, 3-, and 5-years are: gse29621: 0.87, 0.92, 0.95; gse17537: 0.9, 0.9, 0.91; gse72970: 0.76, 0.75, 0.66; gse38832: 0.9, 0.94, 0.96; geo-cohort: 0.79, 0.79, 0.81.). We further evaluated each clinical feature in TCGA and GEO with the AUC values of the riskscore (Fig. 7M, 7N), establishing the riskscore as an independent risk factor for OS after correcting for accessible clinical variables (e.g., age, stage, risk score). Based on the median risk scores in TCGA, individuals in CRC cohort were split into two groups: high risk and low risk. Highrisk individuals lived longer than lowrisk individuals. Then, using the results of multiple Cox regression, nomograms were created to predict OS rates at 1-year, 3-years, and 5-years (Fig. 7O,7P). Based on such a risk model, nomogram plots predict that the patient will perform well in OS (Fig. 7Q,7R). In addition, we also show the risk curve and risk heat map for the model in the supplementary material (Fig.S2C, S2D).

Features of the riskScore

We looked into the riskscore characteristics to determine the clinical and biological functional validity of these 17 genetic markers. Clinically, we discovered that individuals having higher riskscores had higher ACJJ stage, T stage, N stage, and M stage (Fig. 8A-E). The GSEA method was then used to investigate the biology of the risk scores. The high risk group had higher levels of EMT, angiogenesis, and hypoxia, whereas in the low risk category, there were higher activities of fatty acid metabolism, MYC, and adipogenesis (Fig. 8F, G).

We as well looked into whether risk scores could be used as a decision-making tool in immunotherapy or chemotherapy. To begin, we used three immunotherapy datasets, IMvigor210, GSE78220, and GSE179351, to assess the use of 17 gene combinations in the prediction of immunotherapy response. We discovered that individuals having higher riskscore not only had a dismal outlook but also responded poorly to anti-PDL1 treatment in all three datasets (Fig. 9A-C). As a result, we propose the risk score as a reliable predictor of PD1/PDL1 blockade-based therapies. Furthermore, he riskscore, as we discovered, could predict the sensitivity of traditional colorectal cancer chemotherapeutic agents (5-fluorouracil, irinotecan, oxaliplatin, darifenib, and trametinib). Individuals with in high risk category were much more susceptible to these chemotherapeutic drugs, which was anticipated above for immunotherapy (Fig. 9D-H).

We discovered that the ranking among the most frequently mutated genes is mostly similar by the two subgroups after analyzing the mutation profile. However, the comparatively high TP53 mutation rate in the high risk vs low risk groups might explain some of the observed differences in treatment effects (Fig. 9I, 9J).

Based on WGCNA analysis, we selected turquoise modules containing SPP1 and intersected with 17 risk genes and finally selected three hub genes (SPP1, TIMP1 and COMP) that were significantly upregulated in CRC tissue (Fig.S3A-E). SPP1, TIMP1 and COMP found to be elevated among the 17 genes in the signature in the high-risk group (Fig. 9K, Fig.S3F, J). Furthermore, we discovered that SPP1 expression was associated with inferior ACJJ, T, and N staging (Fig. 9L-N), with TIMP1 and COMP demonstrating similar results (Fig.S3G-I, K-M). Following TIMER analysis, SPP1 was linked to infiltration of CD4 T cells,
CD8 T cells, macrophages, and so on, confirming its role in tumor immunity (Fig. 9O). Similar results were demonstrated in TIMP1 and COMP (Fig.S3N, O).

Three signature-related genes were found to be expressed in five CRC cell lines (SW620, HT29, SW480, LOVO, and HCT116) and one human normal intestinal epithelial cell line (NCM460). SPP1, TIMP1, and COMP expression levels were almost all considerably higher in CRC cells compared to human normal intestinal epithelial cells (Fig. 10A-C). The same results were seen in tumor tissue as well as neighboring normal tissue (Fig. 10D-F). These findings support the risk model's stability and dependability.

**Discussion**

TME is characterized by macrophage infiltration, which is regulated mostly by individual's MMR type. Previous research has found variations in the number associated with specific macrophage subgroups in individuals between dMMR and pMMR. Higher immunoreactive macrophage activation in dMMR as well as more TAM in pMMR may explain why individuals with pMMR have a worse clinical features[35]. According to our findings, SPP1 TAM is more numerous in dMMR, and several of its pro-tumor capabilities are elevated in dMMR, suggesting that SPP1 TAM possesses both pro-tumor and immunoreactive functions. TAMs are often classed as "conventional" pro-inflammatory (M1-like) or anti-inflammatory (M2-like) TAMs[36], TAMs in CRC, according to Lei et al., demonstrate strong dichotomy and are classified as SPP1 TAM and C1Q1 TAM[37]. Furthermore, C1QC TAM and SPP1 TAM in CRC cannot be characterized via transcriptional methods that rely on M1 and M2 TAM-related markers. According to our findings, SPP1 TAMs play a critical part in tumor angiogenesis and metastasis, and some researchers have found that SPP1 + TAMs do not exist in primary liver cancer, but are significantly enriched in the liver metastasis site of colorectal cancer. Analysis of lymph node sites also found that SPP1 TAM was only recognized in patients with lymph node metastasis, suggesting that SPP1 TAM promotes the metastasis of CRC cancer cells, and its functional inhibition may be beneficial in the treatment of CRC patients[38, 39].

Investigating its impact of TAM features on individual immunotherapeutic success is now a key focus of malignancy immunotherapy work. TAM has the capability to express several immunological checkpoint markers, block immune system response, and facilitate malignant flee. At present, the research on SPP1 TAM is still insufficient. In this study, SPP1 TAM specifically expresses CTLA4 ligand CD80[40], and their highly immunosuppressive properties indicate that they may be harmful factors of anti-tumor immunity. However, we discovered a positive association among both SPP1 TAM infiltration level as well as dMMR state clients' predictive reaction to immunotherapy, implying that SPP1 TAM has two phenotypes and that immunotherapy may be challenging to accurately predict the role of SPP1 TAM in patients who had high infiltration rate.

In addition, the high expression of SPP1 is associated with the a high amount of penetration of SPP1 TAM. In CRC with different MMR states, SPP1 and SPP1 TAM have similar results. At the same time, our research results show that SPP1 and VCAN are hub genes of SPP1 TAM in EMT and angiogenesis[41].
SPP1, being the main gene of SPP1 TAM, is important in malignant cell interaction. The above results are predictable. The above findings indicate that the penetration of SPP1 TAM has significant therapeutic significance. SPP1 mediated EMT and angiogenesis can increase the infiltration of SPP1 TAM. Therefore, it is necessary to construct two immune subtypes with SPP1 as hub gene.

We furthermore discovered those people mostly with greatest levels of TAM as well as other immune cell infiltration (immune group C2) suffered the lowest survival. At the same time, patients with immune subtype C2 do not respond as well to immunotherapy as patients with immune subtype C1. As a result, the immunological subgroups we discovered cannot easily be assigned to malignancies with immunological activation. Whereas the 2 subgroups discovered from all Colon cancer patients may assist us in better comprehending CRC diversity, they are not suitable for reproducing this complex subtype in clinical settings. Therefore, we use comprehensive bioinformatics tools to develop risk models. In this study, we modeled the prognosis of 17 genes through lasso and stepwise Cox. The prognostic model we establish can well distinguish patients with different prognosis in the test and validation datasets. Among 5 publicly available datasets, riskscore exhibits good reliability and consistent quality. Furthermore, it can anticipate various responses to immunotherapy and chemotherapeutic medication applications at immune checkpoints.

Riskscore can predict that the response of high-risk patients to immunotherapy is not as good as that of low-risk patients, which is similar to the prognosis of some patients with high and low SPP1 TAM infiltration levels. In view of the dual characteristics of SPP1 TAM, it may be related to the insufficient response of SPP1 TAM to immunotherapy, which indicates that immunosubtype C2 and high-risk scores are more inclined to suppress immune response and promote tumor escape. But at the same time, the patients in the high-risk group have higher sensitivity to the classic drugs of the classic colorectal cancer chemotherapy. The three drugs 5-Fluorouracil, Irinotecan and Oxaliplatin in folfox and forfiri, the classic chemotherapy protocols of CRC, are more sensitive to high-risk patients. As the third line therapy for patients with advanced colorectal cancer with BRAF mutation, dabrafenib and trametinib are also more sensitive to high-risk patients[42, 43]. This provides ideas for the therapeutic therapy of high-risk patients with insufficient response to immunotherapy.

However, this study has several limitations. Although we have discussed the role of SPP1 TAM in immunotherapy and prognosis of CRC patients, we need to further verify the SPP1 TAM population through in vivo and in vitro experiments. Whether SPP1 TAMs related genes can be used as immunotherapy targets needs further research.

To summarize, we constructed a more comprehensive macrophage atlas to highlight the great range and diversity of macrophages found in people with various MMR states. SPP1 TAM is not only enriched in dMMR patients, but also shows two characteristics of immune response, which may explain that patients with colorectal cancer with dMMR are not satisfactory in chemoradiotherapy and targeted therapy. Coincidentally, studies of adjuvant chemotherapy after gastric cancer surgery have also found that patients with dMMR who receive adjuvant chemotherapy have impaired survival outcomes[44].
prognosis model constructed by Hub DEG SPP1 related to it has different responses to immune response and chemotherapy drugs, which provides new clues to inhibit the potential efficacy of SPP1 TAM.

**Abbreviations**

CRC  
Colorectal cancer  
TME  
Tumor microenvironment  
SPP1  
Secreted Phosphoprotein 1  
TAM  
Tumor-associated macrophages  
MMR  
Mismatch Repair  
dMMR  
different Mismatch Repair  
pMMR  
proficient Mismatch Repair  
EMT  
epithelial mesenchymal transition  
ICBs  
Immune checkpoint blockade  
OS  
overall survival  
IC50  
semi-inhibitory concentration  
WGCNA  
Weighted correlation network  
GSEA  
Gene set enrichment analysis  
GSVA  
Gene set enrichment analysis  
FDR  
False discovery rate  
KM  
Kaplan–Meier  
PD-L1  
Programmed cell death ligand 1  
CTLA-4
Cytotoxic T lymphocyte antigen 4
UMAP
Uniform flow approximation and projection

Declarations

Data Availability Declaration


Competing Interests

The authors state that they have no conflicts of interest.

CONTRIBUTIONS FROM AUTHORS

The study was planned and designed by YZL and ZWY. The analytical processes were carried out by YZL. The findings were examined by YZL and HZW. The manuscript was written by YZL and ZWY. This manuscript was evaluated by QZ and JL. The final manuscript was co-written and approved by all writers.

Funding

The work was supported by the National Natural Science Foundation of China (Jing Liu, Grant No. 82072753); the Young Top-notch Talent Cultivation Program of Hubei Province (Jing Liu, Grant No. HBQB20220401); Wuhan University (Jing Liu, Grant No. 2042019kf0206); Zhongnan Hospital of Wuhan University (Jing Liu, Grant No. lcyf202202).

References


**Figures**
Figure 1

Atlas of tumour-associated macrophages: (A) the proportion of myeloid cells in dMMR and pMMR; (B) a UMAP clustering of macrophages; (C) top five marker genes in all subpopulations; (D) marker genes in all subpopulations; (E) the proportion of macrophages in dMMR and pMMR; (F) expression of marker genes in dMMR and pMMR.
Figure 2

Functions and cell ratios: (A) the functional activity of all macrophages subpopulations; orange symbol is significant in dMMR; black symbol is significant in pMMR; (B) cell ratios of macrophages subpopulations in dMMR and pMMR at different ACJJ stages and N stages.
Figure 3

Pseudotime trajectories analysis: (A) pseudotime differentiation trajectory of macrophages subpopulations; (B) pseudotime trajectory from C1QC TAM to SPP1 TAM; (C) heatmap of pseudotime trajectory.
Figure 4

Cellular communication and differential pathways and differential genes of SPP1 TAM in different MMR states: (A) SPP1 signaling pathway network in dMMR; (B) SPP1 signaling pathway network in pMMR; (C) differential pathways of SPP1 TAM in different MMR states; (D) differential genes of SPP1 TAM in different MMR states; (E) GO enrichment analysis for differential genes of SPP1 TAM in different MMR states.
Figure 5

SPP1 TAM infiltration and response to immunotherapy: (A) expression of immune checkpoints in macrophages subpopulations; (B) the infiltration of SPP1 TAM in different MMR states; (C) the expression of SPP1 in different MMR states; (D) heatmap of correlation between SPP1 TAM infiltration, SPP1 expression, TMB and responsescore; (E) responsescore in different MMR states; (F) correlation
between SPP1 TAM infiltration and different MMR states; (G) survival analysis of the infiltration of SPP1 TAM.

**Figure 6**

Molecular subtypes of TAM: (A) the consensus score matrix; (B) The CDF curves of consensus matrix; (C) the infiltration of different macrophages in molecular subtypes; (D) the distribution of the two molecular subtypes; (E) Immunesco, Stromalscore and Tumorpurity between the two subtypes; (F) survival analysis of molecular subtypes; (G) the infiltration of different immune cells in molecular subtypes; (H) IPS and the expression of PD-L1 between the two subtypes.
Figure 7

Construction of TAM subtype risk scoring model: (A-B) Kaplan–Meier curves of OS and ROC for predicting OS for riskscore in TCGA; (C-D) Kaplan–Meier curves of OS and ROC for predicting OS for riskscore in gse29621; (E-F) Kaplan–Meier curves of OS and ROC for predicting OS for riskscore in gse17537; (G-H) Kaplan–Meier curves of OS and ROC for predicting OS for riskscore in gse72970; (I-J) Kaplan–Meier curves of OS and ROC for predicting OS for riskscore in gse38832; (K-L) Kaplan–Meier
curves of OS and ROC for predicting OS for riskscore in geo-cohort; (M) AUC for clinical feature in TCGA; (N) AUC for clinical feature in GEO; (O) multiple Cox for TCGA; (P) multiple Cox for GEO; (Q) nomograms for TCGA; (R) nomograms for GEO.

Figure 8
Features of the riskscore and GSEA analysis: (A-E) riskscore in different ACJJ stages, T stages, N stages, M stages and MMR states; (F) GSEA volcano map for different risk groups; (G) GSEA for different risk groups.

Figure 9
Features of the riskscore: (A) Kaplan–Meier curves of OS, response to anti-PD-L1 treatment and ROC for predicting response for riskscore in IMvigor210; (B) Kaplan–Meier curves of OS, response to anti-PDL1 treatment and ROC for predicting response for riskscore in GSE78220; (C) Kaplan–Meier curves of OS, response to anti-PDL1 treatment and ROC for predicting response for riskscore in GSE179351; (D-H) the sensitivity of traditional colorectal cancer chemotherapeutic agents in different risk groups (5-fluorouracil, irinotecan, oxaliplatin, darifenib, and trametinib); (I-J) mutation landscape in different risk groups; (K-N) the expression of SPP1 in different risk groups, ACJJ stages, T stages and N stages; (O) correlations between SPP1 and the infiltration of different immune cells.

Figure 10
Validation of the expression patterns of 5 signature genes: (A) the expression of SPP1 in cell lines by RT-PCR; (B) the expression of TIMP1 in cell lines by RT-PCR; (C) the expression of COMP in cell lines by RT-PCR; (D) the expression of SPP1 in CRC tissue and normal tissue by RT-PCR; (E) the expression of TIMP1 in CRC tissue and normal tissue by RT-PCR; (F) the expression of COMP in CRC tissue and normal tissue by RT-PCR.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.zip