

# Expression signature, prognosis value, and immune characteristics of MUC1 identified by pan-cancer analysis

**Zi-Jian Zhang**

Second Xiangya Hospital Department of General Surgery

**YunPeng Huang**

Second Xiangya Hospital Department of General Surgery <https://orcid.org/0000-0002-0578-1235>

**Zhong-Tao Liu**

Second Xiangya Hospital Department of General Surgery

**Kai Liu**

Second Xiangya Hospital Department of General Surgery

**Yong-Xiang Wang**

Second Xiangya Hospital Department of General Surgery

**Yang Yang**

Second Xiangya Hospital Department of General Surgery

**Li Xiong**

Second Xiangya Hospital Department of General Surgery

**Heng Zou**

Second Xiangya Hospital Department of General Surgery

**Yu Wen** (✉ [wenyu2861@csu.edu.cn](mailto:wenyu2861@csu.edu.cn))

<https://orcid.org/0000-0001-7444-3021>

---

## Primary research

**Keywords:** MUC1, pan-cancer, prognosis, immunotherapy

**Posted Date:** April 12th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-390447/v1>

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

[Read Full License](#)

---

# Abstract

**Background:** Mucin 1 (MUC1) plays a major role in the occurrence and development of tumor by regulating the process of tumor cell proliferation, epithelial-mesenchymal transformation and epigenetics. However, the relationship between MUC1 expression and tumor prognosis and its role in tumor immunity is still worth exploring.

**Methods:** MUC1 expression was analyzed via GTEx database and TCGA database. We used the Kaplan-Meier survival estimation method to evaluate the influence of MUC1 on tumor prognosis through the survival information from TCGA database. The correlations between MUC1 and immune cell infiltration, tumor microenvironment were investigated through TIMER algorithm and ESTIMATE. In addition, we used Spearman correlation test to examine the correlation between MUC1 and TMB, MSI. Spearman correlation test was also designed to predict the correlation between MUC1 and immune checkpoint genes, four methyltransferases. Further, Gene Set Enrichment Analysis was used to explore the potential mechanism of MUC1 in Adrenocortical carcinoma (ACC) and Liver hepatocellular carcinoma (LIHC).

**Results:** Our study reported that MUC1 is highly expressed in most tumors, differing between cancer types. In most cancers, high expression of MUC1 means poor prognosis indicators, such as overall survival (OS), disease free survival (DSS), disease free survival (DFS) and progression free survival (PFS). MUC1 was positively correlated with infiltrating levels of B cells, CD4+ T cells and CD8+ T cells, dendritic cells, macrophages, neutrophils in LIHC, Prostate adenocarcinoma (PRAD) and Thyroid carcinoma (THCA). Besides, we reported that the expression of MUC1 correlated significantly with immune checkpoint gene, TMB and MSI and in most tumors. However, we found that MUC1 is negatively correlated with most immunotherapy-related indicators in BRCA and LUAD. The relationship between MUC1 and tumor neoantigens and DNA methylase is different in different tumors.

In ACC and LIHC, MUC1 can promote the metabolism of many substances. MUC1 can inhibit amyotrophic lateral sclerosis, DNA replication, base excision repair, proteasome in ACC. Similarly, MUC1 inhibits axon guidance, the interaction of cytokine and cytokine receptor, focal adhesion, and endocytosis in LIHC.

**Conclusions:** Our research demonstrates that MUC1 is correlated with prognosis and tumor-infiltrating immune cells, including those of B cells, CD8+ T cells, CD4+ T cells, dendritic cells, macrophages, neutrophils in different tumors. In addition, MUC1 is related to immune checkpoint gene, neoantigen, and some prognostic indicators of immunotherapy such as TMB and MSI, suggesting that MUC1 can also be invoked as a target and prognostic biomarker of immunotherapy. By analyzing the relationship between the expression of MUC1 and methyltransferase, we found that MUC1 regulates DNA methylation. Finally,

we used GSEA to study the function of MUC1 in ACC and LIHC. Briefly, our study highlights the significance of MUC1 in the study of tumor immunity from the perspective of pan-cancer.

## Introduction

Cancer is now being the main cause of death and an important obstacle to the extension of life expectancy. Depending on estimates from the World Health Organization in 2019, cancer is the first or second leading cause of death before the age of 70 in 112 countries, and the third or fourth cause of death in 23 other countries[1]. Among women, the top five cancers are breast cancer, lung cancer, colorectal cancer, prostate cancer and gastric cancer. Among men, the most common cancers are lung cancer, prostate cancer, colorectal cancer and liver cancer. Lung cancer and breast cancer are the most frequent causes of death for men and women, respectively[2]. After surgery, radiotherapy and chemotherapy, cancer immunotherapy is rapidly developing into the fourth most prominent cancer therapy[3]. As the first membrane mucin to be identified, MUC1 is up-regulated, aberrantly glycosylated and polarization disappears in many cancers[4]. Using MUC1 as a target for immunotherapy has made excellent progress. However, there are no studies to explore the significance of MUC1 and its role in immunotherapy from the perspective of pan-cancer.

In this study, we used bioinformatics techniques to predict the effect of MUC1 on the prognosis of pan-cancer and infiltration of immune cells in different tumors. Besides, we analyzed the relationship between MUC1 and forty common immune checkpoint genes as well as the relationship of MUC1 expression levels with the expression of four methyltransferases (DNMT1, DNMT2, DNMT3A, DNMT3B) in different tumors. Finally, we explain the regulatory mechanism of MUC1 in different tumors by GSEA enrichment analysis. Our results further provide evidence for MUC1 as a target for immunotherapy and suggest that MUC1 can be invoked as a prognostic marker for immunotherapy.

## Results

### Expression of MUC1 in pan-cancer

First, we collated tumor samples and adjacent nearby tissues from The Cancer Genome Atlas (TCGA), normal samples from Genotype-Tissue Expression (GTEx) database to evaluate the mRNA expression characteristics of MUC1 in humans. Through pan-cancer analysis, we found that the expression of MUC1 was quite different in different normal tissues (Figure 1A). In general, comparing tumor tissues with adjacent nearby tissues from TCGA, we found that the expression of MUC1 was upregulated, in such as Bladder Urothelial Carcinoma (BLCA), Breast invasive carcinoma (BRCA), Cholangiocarcinoma (CHOL), Glioblastoma multiforme (GBM), Brain Lower Grade Glioma (LGG), LIHC, Stomach adenocarcinoma (STAD), THCA, Uterine Corpus Endometrial Carcinoma (UCEC). However, MUC1 was found to be downregulated in Colon adenocarcinoma (COAD), Esophageal carcinoma (ESCA), Head and Neck squamous cell carcinoma (HNSCN), Kidney Chromophobe (KICH), Kidney renal clear cell carcinoma (KIRC), Kidney renal papillary cell carcinoma (KIRP), Lung squamous cell carcinoma (LUSC) (Figure 1B).

Then we analyzed the difference of MUC1 expression between tumor tissue from TCGA and normal samples from GTEx, MUC1 was upregulated in BLCA, BRCA, Cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), CHOL, GBM, LGG, Lung adenocarcinoma (LUAD), Ovarian serous cystadenocarcinoma (OV), Pancreatic adenocarcinoma (PAAD), THCA and UCEC. MUC1 was found to be downregulated in ACC, HNSC, KICH, KIRC, KIRP, Acute Myeloid Leukemia (LAML), Lung squamous cell carcinoma (LUSC), PRAD, Skin Cutaneous Melanoma (SKCM), STAD, and Testicular Germ Cell Tumors (TGCT) (Figure 1C).

### **Correlation analysis between MUC1 expression level and prognostic value**

After evaluating the characteristics of MUC1 expression at the mRNA levels, we used Kaplan–Meier plotter to investigate the correlation between MUC1 mRNA expression levels and OS, DSS, DFS and PFS in different tumors. Using TCGA data, our results showed that upregulated MUC1 expression was found to be associated with shorter OS in ACC, LAML, LGG, LUSC and PAAD, and with longer OS in Sarcoma (SARC) (Figure 2 A-F). Besides, upregulation of MUC1 expression was linked to shorter DSS in ACC, LGG and LUSC (Figure 2 G-I). High MUC1 expression was related to poor PFS in LGG, LUSC, Pheochromocytoma and Paraganglioma (PCPG), UCEC and Uveal Melanoma (UVM) (Figure 2 J-N). Furthermore, the high expression of MUC1 was also associated with poor DFS in ACC (Figure 2 O). These results confirmed that the expression of MUC1 is negatively correlated with the prognosis of most cancers.

### **Correlation analysis between MUC1 expression and immune cells**

After exploring the relationship between MUC1 expression and tumor prognosis, we studied the relationship between MUC1 and immune infiltration in different tumors, the analysis of tumor-infiltrating immune cells included B cell, CD4+T cell, CD8+T cell, dendritic cell, macrophage cell, neutrophil cell. With the cutoff of p value=0.05, MUC1 were positive with tumor-infiltrating immune cells from the aspect of correlation analysis. In LIHC (Figure 3 A), MUC1 is positive with B cell (correlation coefficient=0.33, p value=8.16e-11), CD4+T cell (correlation coefficient=0.374, p value=1.02e-13), CD8+T cell (correlation coefficient=0.353, p value=2.79e-12), dendritic cell (correlation coefficient=0.387, p value=8.13e-15), macrophage (correlation coefficient=0.396, p value=1.71e-15) and neutrophil (correlation coefficient=0.391, p value=3.85e-15). In PRAD (Figure 3 B), MUC1 is positive with B cell (correlation coefficient=0.29, p value=4.5e-11), CD4+T cell (correlation coefficient=0.41, p value=0), dendritic cell (correlation coefficient=0.344, p value=4.5e-15), macrophage (correlation coefficient=0.295, p value=1.97e-11) and neutrophil (correlation coefficient=0.449, p value=0). In THCA (Figure 3 C), MUC1 is positive with B cell (correlation coefficient=0.266, p value=1.24e-09), CD4+T cell (correlation coefficient=0.355, p value=1.75e-16), CD8+T cell (correlation coefficient=0.119, p value=0.00752), dendritic cell (correlation coefficient=0.501, p value=0), macrophage (correlation coefficient=0.263, p value=1.83e-09) and neutrophil (correlation coefficient=0.511, p value=0).

### **Relationship between MUC1 expression and tumor immune microenvironment**

Because of the interaction of immune cells, we also explored the relationship between MUC1 and immunity from the perspective of tumor immune microenvironment. The ESTIMATE method was used to analyze the correlation between MUC1 expression in TCGA tumor samples and the ratio of both stromal cells and immune cells (Figure 4). In GBM, KIRC, LGG, LIHC, LUSC, OV, PCPG, PRAD, THCA, SKCM and UVM, we found that MUC1 significantly positively correlated with stromal score, immune score, and estimated score. MUC1 had the highest correlation with stromal score in TGCT ( $r = 0.646$ ,  $P < 0.001$ ), while the highest correlation with immune score ( $r = 0.556$ ,  $P < 0.001$ ) and estimate score ( $r = 0.513$ ,  $P < 0.001$ ) was found in LGG.

### **Relationship between MUC1 expression and immune checkpoint gene expression in different tumors**

Immune checkpoint refers to a series of molecules expressed on immune cells and can regulate the degree of immune activation. They play an important role in preventing autoimmunity (such as inhibiting abnormal immune function and preventing overactivation of immune function). Tumors can use these molecules to escape recognition and destruction of the body. Therefore, blocking these molecules will cause anti-cancer response and promote the immune clearance of tumor cells, thus inhibit the occurrence and development of the tumor. We used the mRNA sequence database to assess whether there is a relationship between MUC1 expression and more than forty common immune checkpoint genes. The results showed that MUC1 was highly correlated with genes in many different tumors. In ACC, LGG, THCA, Thymoma (THYM) and UVM, we found that MUC1 and other immune checkpoint genes were significantly co-expressed. Besides, the expression of MUC1 was negatively correlated with most immune checkpoint genes in BRCA, LUAD and PAAD (Figure 5 A).

### **Relationship between MUC1 expression, and TMB, MSI in different tumors**

Immune checkpoint inhibitors can kill tumor cells by breaking immune tolerance and activating immune activity. At present, a variety of immune checkpoint inhibitors (ICI) drugs have been used in clinic. Tumor Mutational Burden (TMB) and Microsatellite instability (MSI) are important indicators to predict the prognosis of ICI-treated patients and have been explored extensively. Our results showed that the expression of MUC1 correlated significantly with TMB in THYM, THCA, PAAD, LGG, HNSC, ESCA, COAD and BLCA ( $P < 0.05$ ), and THYM had the highest coefficients, while LUAD and BRCA had the lowest score (Figure 5 B). The coefficient value refers that MUC1 was positively correlated with high mutation status in THYM, and positively correlated with low mutation status in BRCA and LUAD. In addition, we also analyzed the correlation between MUC1 expression and MSI in different tumors. Results showed that MUC1 was significantly correlated with MSI in TGCT and COAD ( $P < 0.05$ ), indicating a positive correlation between MUC1 expression and MSI in these cancers. In contrast, the expression of MUC1 had negative coefficients in UCS, UCEC, CESC, and BRCA, indicating that there is a significant negative correlation between MUC1 expression and MSI in these cancers (Figure 5 C).

### **Relationship between MUC1 expression and neoantigen in different tumors**

Neoantigen is a new antigen encoded by mutant genes of tumor cells, which is different from that expressed by normal cells. According to the specific immune activity of neoantigen, the new antigen vaccine can be designed and synthesized through the mutation of tumor cells to achieve the therapeutic effect. Our research evaluated the correlation between MUC1 expression and neoantigen in different tumors, and MUC1 was positively correlated with neoantigen in HNSC and STAD ( $P < 0.05$ ). However, in LUAD, BRCA, LIHC and SKCM, MUC1 was negative correlated with neoantigen (Figure 5 D).

### **Relationship between MUC1 expression and four methyltransferases expression in different tumors**

DNA methylation is a form of chemical modification of DNA, which can cause changes in chromatin structure, DNA conformation, DNA stability and the interaction between DNA and protein, thus regulating gene expression. It can modify the genetic expression without altering the DNA sequence through the action of DNA methyltransferase. In our research, we discussed the correlation between MUC1 expression and four methyltransferases expression (DNMT1, DNMT2, DNMT3A and DNMT3B) (Figure 6). In STAD, THCA, UCEC, UVM, KIRP, LAML, LGG, LIHC, PCPG, PARD and READ, MUC1 was positively correlated with DNA methylation ( $P < 0.05$ ). On the contrary, MUC1 was negatively correlated with DNA methylation in BRCA, COAD, ESCA, LUAD, LUSC, OV, and PAAD ( $P < 0.05$ ).

### **Gene set enrichment analysis (GSEA) of MUC1 in ACC and LIHC**

After preliminary exploration of the gene epigenetic mechanism of MUC1, we used GSEA to examine the function of MUC1 in ACC and LIHC. In ACC, MUC1 can promote the cycle of citrate and TCA, the metabolism of alanine, aspartate and glutamate, the metabolism of xenobiotics and drugs by cytochrome P450 (Figure 7 A). In addition, MUC1 can inhibit amyotrophic lateral sclerosis, DNA replication, base excision repair, proteasome in ACC. In LIHC, MUC1 advances the metabolism of many substances, such as drug, propanoate, fatty acid, glycine, serine and threonine (Figure 7 B). Conversely, MUC1 inhibits axon guidance, the interaction of cytokine and cytokine receptor, focal adhesion, and endocytosis.

## **Discussion**

MUC1 is a high molecular weight glycosylated transmembrane protein expressed mainly on the apical surface of glandular or luminal epithelial cells of almost all tissues, such as the mammary gland, gastrointestinal, respiratory, urinary and reproductive tracts[5]. Early studies have suggested that the function of MUC1 is to form physical barriers, block pathogens and protect and lubricate the epithelium[6, 7]. However, abnormal expression of MUC1 was also found in a variety of cancers, including breast cancer, gastric cancer, hepatobiliary cancer, bladder cancer, pancreatic cancer and other tumors[8-12]. The high expression of MUC1 is associated with poor prognosis in many cancers. In lung cancer, MUC1 can be regulated by various signal molecules and transcription factors, such as the extreme expression of STAT3 in NSCLCs [13] and the hypoxia-induced HIF1A[14]. In addition, the overexpression of MUC1 can enhance the stemness and Paclitaxel resistance of NSCLCs, affecting the prognosis[15]. MUC1 promoted the expression of platelet-derived growth factor- A, which causing the tumor growth, angiogenesis and

metastasis of pancreatic ductal adenocarcinoma[16]. Another study shows that MUC1 can directly regulate the expression of multidrug resistance genes in pancreatic cancers cells, increasing resistance to chemotherapeutic drugs (gemcitabine and etoposide)[17]. In endometrial cancer, MUC1 was hormonally upregulated and significantly associated with poor prognosis[18]. These researches are consistent with the results of our study. Therefore, MUC1 is supposed to be a diagnostic and prognostic marker of different tumors.

MUC1 changes both quality and quantity in different tumors, leading to the emergence of new antigenic epitopes. MUC1 is a cell surface molecule that comes into contact with the immune system, so it has been considered as an ideal target for immunotherapy in recent studies. Dendritic cells (DCs) are important parts of immune microenvironment, more and more studies have demonstrated that effectiveness of vaccination with DCs pulsed with MUC1 to initiate adaptive cytolytic immune responses via T cells. It has been noted that MUC1 can be presented to DC, to activate CD8+T cell and induce anti-cancer-related immune response[19]. A study evaluated the clinical effectiveness of a MUC1 peptide-loaded DC vaccine in 12 pancreatic and biliary cancer patients following surgical resection. The vaccine was tolerated and no toxicity was observed. Surprisingly, four vaccinated patients survived for up to five years[20]. Similarly, the efficacy of the vaccine in pancreatic cancer has been confirmed by many researches[21, 22]. Ge et al.[23] reported that DC vaccine pulsed with survivin and MUC1 was well tolerated without dose-limiting toxicity, and significantly improved the quality of NSCLS patients' life. Another clinical study showed that the DC-based vaccine targeting MUC1 induced an antitumor immune response that promoted prolonged survival of patients with refractory NSCLC. The median survival time (MST) after the initial vaccination was 7.4 months, and the 1-year survival rate was 25.0%. When the patients were vaccinated more than 6 times, the MST could be extended to 9.5 months, and the 1-year survival rate was 39.3%[24]. The vaccine can be also against a variety of tumors, such as prostate cancer, metastatic renal cancer and metastatic colorectal cancer[25-27]. Targeting MUC1 to chimeric antigen receptor (CAR)-T cells is also a promising immunotherapy. Zhang et al.[28] confirmed MUC1-CAR-T cells have long-lasting tumor killing and proliferative capabilities in esophageal cancer, and it had significant antitumor function and a prolonged half-life in subcutaneous transplantation models and PDX models. In another study, Mei et al. found[29] that human IL22 recombinant protein could increase the MUC1 expression and enhance the function of T cells. They constructed a fourth-generation CAR (CAR-MUC1-IL22 T cells) that secretes IL22, and confirmed that it has a stronger anti-tumor effect on MUC1+HNSCC cells. MUC1-CAR-T cells have good therapeutic effects in seminal vesicle cancer, NSCLC, triple-negative breast cancer, and pancreatic ductal adenocarcinoma[30-33].

Our study attempts to deeply analyze the relationship between immunity and MUC1, so we explored the correlation between the expression of MUC1 and 6 kinds of immune cells. Our results showed that in LIHC, PRAD and THCA, the expression of MUC1 was positively correlated with immune cells. MUC1 may be a potential target for immunotherapy of these cancers.

TMB is the total number of mutations per megabase in tumor, reflecting cancer mutation quantity. The more tumor mutations, the greater the difference between tumor and normal cells. Therefore, the higher

the value of TMB, the more likely the tumor is supposed to be recognized by immune cells, and the more effective it is for immunotherapy. In recent years, TMB has been taken into account as a biomarker of immunotherapy[34]. Similarly, the formation of MSI caused by DNA MMR proteins defect leads to the accumulation of mutations and the production of neoantigens. MSI is another important marker that determines the effectiveness of immunotherapy[35]. Neoantigens are also given the same meaning. Neoantigens are not expressed in normal tissues but only expressed in tumor tissues. Because of its high specificity and can be recognized by immune cells, it is the target of personalized therapeutic cancer vaccines[36]. Previous studies have confirmed that MUC16 is associated with TMB in hepatocellular carcinoma and gastric cancer[37, 38]. MUC family proteins are highly related to MSI in colon cancer[39]. MUC2 expressed in mucinous carcinomas, a distinct subtype of colon cancer associated with MSI[40]. In our research, we found that MUC1 is associated with TMB, MSI and neoantigen in a variety of tumors. Especially in BRCA, these three indexes are all negatively correlated with MUC1. Our results suggest that the expression of MUC1 in BRCA may be used to predict the effectiveness of immunotherapy. In general, our study revealed a relationship between the expression of MUC1 and the effectiveness of immunotherapy. At present, in specific cancers, a lot of researches are still needed to examine the possibility of MUC1 as a target of immunotherapy and predictive.

In recent years, immune checkpoint inhibitors have made a significant breakthrough as a novel type of immunotherapy. Antibodies blocking the cytotoxic T lymphocyte-associated protein 4 (CTLA-4) or the programmed cell death 1 (PD-1) pathway are the two most common inhibitors with good clinical effects[41]. At present, there are many other target antibodies under study, such as LAG-3, TIM-3, TIGIT. On the basis of exploring the relationship between MUC1 and immunotherapy, we also studied the relationship between MUC1 and immune checkpoint genes. Our result showed that the expression of MUC1 was negatively correlated with most immune checkpoint genes in BRCA, LUAD and PAAD. Interestingly, the expression of MUC1 in BRCA and LUAD is negatively correlated with most immunotherapy-related indicators in this study.

More and more studies have demonstrated that DNA methylation is highly related to diagnosis, prognosis and prediction of response to therapies of tumors, such as NSCL, colorectal cancer, hepatocellular carcinoma and metastatic breast cancer[42-45]. Yamada et al. found that the DNA methylation level of MUC1-negative cancer cell lines was greater than the MUC1-positive cell lines[46]. This research indicated that DNA methylation play a critical role in MUC1 gene expression. Another research showed that MUC1 and NF- $\kappa$ B p65 form a complex that binds to the promoter of DNMT1 and DNMT3b, promoting transcription of DNMT1 and DNMT3b[47]. DNMT1 and DNMT3b are two crucial DNA methyltransferases, so MUC1 can promote DNA methylation. In pancreatic ductal adenocarcinomas, a study found that low methylation of MUC1 promoters correlates with decreased overall survival[48]. This study also revealed that the expression of several DNA methylation/demethylation factors have a significant correlation with MUC1 methylation status. Subsequent studies also concluded that aberrant methylation of MUC1 promoters are potential prognostic biomarkers for pancreatic ductal adenocarcinomas[49]. DNA methylation refers to the chemical modification that transfers methyl to specific bases in the DNA chain under the action of DNA methylase. In our research, we found that the correlation between expression

levels of DNA methyltransferase and MUC1 varied in different tumors. In addition, although we have explored the function of MUC1 in ACC and LIHC, a lot of research is still needed to explore the epigenetic changes of MUC1 and its potential functions, which may contribute to discover new cancer treatments and predict the prognosis of cancer patients.

## **Materials And Methods**

### **Obtain MUC1 expression data from public databases**

We analyzed the expression of MUC1 in various normal tissues through the GTEx database. We have introduced 6678 healthy tissues and organs from the GTEx (<https://www.gtexportal.org/home/>) database sample. Sequencing data of 33 tumors were collected from 11,057 samples (10,327 tumor samples and 730 matched paraneoplastic samples) in the TCGA database. Some tumors with missing matches or too few paraneoplastic samples were excluded from the calculation of the difference in expression of MUC1 in tumor and paraneoplastic tissue. For the remaining tumor samples, the tumor samples were normalized and then subjected to the Wilcoxon test to analyze whether there were differences in MUC1 expression between these tumors and normal tissues.  $P < 0.05$  is considered statistically significant.

### **Correlation between the expression of MUC1 and the prognosis of tumor patients**

We downloaded the survival information (including overall survival, disease-specific survival, disease-free survival, and progression-free survival) of 10,327 tumor samples from 33 tumors in TCGA. The Kaplan-Meier survival estimation method was used to calculate the prognostic value of MUC1, and the log-rank test was used to determine the significant difference index. Comprehensive analysis of the results obtained by these two methods, when the overall survival test  $P < 0.05$ , we believe that the expression of MUC1 significantly affects the prognosis of the tumor.

### **Correlation analysis between immune cell infiltration and MUC1 expression**

Through the TIMER algorithm, we use gene expression profiles of 10897 samples from 32 tumors in TCGA to infer the number of tumor infiltrating immune cells (TIICs). TIMER is a statistical method of deconvolution, which can be used for approximate calculation of immune infiltration. We analyzed the relationship between the expression level of the MUC1 gene and the abundance of infiltrating immune cells (including CD4+ T cells, CD8+ T cells, B cells, neutrophils, dendritic cells, and macrophages) through the above methods.

### **Correlation analysis between tumor microenvironment and MUC1 expression level**

Immune cells and stromal cells are the two main non-tumor components in the tumor microenvironment. They are of great value to the diagnosis and prognosis of cancer. ESTIMATE (<https://bioinformatics.mdanderson.org/public-software/estimate/>) is a tool that uses gene expression data to predict the abundance of stromal and immune cell infiltration and tumor purity in tumor tissues.

Based on the expression profile model files of 9664 samples of 33 tumors, after removing the normal samples, according to the ratio of stromal cells and immune cells in each tumor sample, the estimate package is used to estimate the tumor purity. The stromal score, immune score, and estimated score are used to calculate tumor purity. Then we combine the MUC1 expression data and the score of estimate algorithm, use the spearman correlation test to calculate their correlation. After that, the correlations were plotted using the ggplot2 (<https://CRAN.R-project.org/package=ggplot2>), ggpubr, and ggExtra (<https://CRAN.R-project.org/package=ggExtra>) package.

### **The correlation of TMB, MSI and neoantigens with the expression of MUC1**

Tumor mutation burden (TMB) and microsatellite instability (MSI) are important evaluation indicators directly related to the effect of immunotherapy such as PD-1. We downloaded the mutation data of 10114 samples of these 33 tumors in TCGA, calculated the mutation score of each sample, and obtained the TMB information of each tumor. Further, we used Spearman correlation test to analyze the correlation between MUC1 expression and TMB, and used the fmsb package (<https://CRAN.R-project.org/package=fmsb>) to construct the relevant radar chart. We also downloaded and analyzed the MSI scores of 10,415 tumor samples, and used the same processing method to draw related radar charts of MUC1 and MSI. Predicted neoantigens of each

samples were retrieved from previous publication[50]. Correlation among TMB, MSI, neoantigens and MUC1 were calculated using the spearman correlation test.

### **The correlation between MUC1 and known essential marker genes**

After exploring the relationship between MUC1 and immune cell infiltration, it is necessary to further explore the correlation between MUC1 and specific immune checkpoint genes. We screened 47 immune checkpoint genes from the literature, and tested the correlation between MUC1 expression in tumors and the expression of these immune checkpoint genes through Spearman correlation test. Then, use the reshape2 package (<http://www.jstatsoft.org/v21/i12/>) to create related heat maps. DNA methylation changes chromatin structure, DNA conformation, DNA stability, and the way DNA interact with proteins. Therefore, we analyzed the correlation between MUC1 expression and four methyltransferases (DNMT1, DNMT2, DNMT3A and DNMT3B). The analysis method is basically the same as described above.

### **Gene Set Enrichment Analysis (GSEA) analysis**

Since MUC1 is closely related to patient prognosis and immune regulation responsiveness in LIHC and ACC, we further want to explore the potential mechanism of MUC1 in these tumors. To determine the most significant pathways of MUC1 in these tumors, the tumors were grouped into high and low expression groups based on the expression level of MUC1. The expression level of MUC1 was then analyzed by limma, org.Hs.eg.db, clusterProfiler (<http://bioconductor.org/packages/release/bioc/html/clusterProfiler.html>), and enrichplot package based on c2.cp.kegg.v7.2.symbols background file to perform Gene Set Enrichment Analysis on the tumor expression matrix file and the

most significant KEGG pathways were selected to plot enrichment curves.  $P < 0.05$  was considered a significant difference criterion.

## Statistical analysis

The R version 3.6.2 software (<https://www.r-project.org/>) and its ancillary packages were used for data analysis. Limma package and Student's t-test were used to analyze MUC1 expression, and  $P < 0.05$  was considered statistically significant. Kaplan-Meier curves were used for survival analysis using the log-rank test, and  $P < 0.05$  was considered statistically significant survival difference. Spearman or partial Spearman method was used to analyze correlations between genes and correlations between genes and immune cells.  $P < 0.05$  was considered statistically significant. All figures in this study were performed using R version 3.6.2.

## Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: All data generated or analyzed during this study are included in **Materials and Method** and **Figures** of this published article.

Competing interests: The authors declare no competing conflict for this article.

Funding: This work was supported by National Natural Science Foundation of China, No. 81970569 and No. 81773293; and Natural Science Foundation of Hunan Province, No. 2017SK50121.

Authors' contributions: Zi-Jian Zhang and Yun-Peng Huang are co-first authors of this article, we completed the data processing and writing of the article together. Zhong-Tao Liu, Kai Liu, Yong-Xiang Wang, Yang Yang and Li Xiong proposed amendments to this article. Heng Zou and Yu Wen are co-corresponding authors of this article, they provided guidance and valuable comments on this article. All authors read and approved the final manuscript

Acknowledgements: The authors gratefully acknowledge TCGA, GTEx and TIMER for open access to their database.

## References

[1] (WHO). WHO. Global Health Estimates 2020: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019. WHO;. Accessed December 11, 2020.; [who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death](http://who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death):

- [2] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021;
- [3] Song W, Musetti S and Huang LJB. *Nanomaterials for cancer immunotherapy*. 2017; 148: 16-30.
- [4] Taylor-Papadimitriou J, Burchell J, Graham R and Beatson RJB. Latest developments in MUC1 immunotherapy. 2018; 46: 659-668.
- [5] Gendler SJJ. Mucins and neoplasia. MUC1, the renaissance molecule. 2001; 6: 339-353.
- [6] Yolken R, Peterson J, Vonderfecht S, Fouts E, Midthun K and Newburg DJT. Human milk mucin inhibits rotavirus replication and prevents experimental gastroenteritis. 1992; 90: 1984-1991.
- [7] Belley A, Keller K, Göttke M and Chadee K. Intestinal mucins in colonization and host defense against pathogens. *Am J Trop Med Hyg* 1999; 60: 10-15.
- [8] Fremd C, Stefanovic S, Beckhove P, Pritsch M, Lim H, Wallwiener M, Heil J, Golatta M, Rom J, Sohn C, Schneeweiss A, Schuetz F and Domschke CJO. Mucin 1-specific B cell immune responses and their impact on overall survival in breast cancer patients. 2016; 5: e1057387.
- [9] Ando H, Aihara R, Ohno T, Ogata K, Mochiki E and Kuwano HJT. Prognostic significance of the expression of MUC1 and collagen type IV in advanced gastric carcinoma. 2009; 96: 901-909.
- [10] Kasprzak A and Adamek AJ. Mucins: the Old, the New and the Promising Factors in Hepatobiliary Carcinogenesis. 2019; 20:
- [11] Sangoi A, Higgins J, Rouse R, Schneider A, McKenney JJ. Immunohistochemical comparison of MUC1, CA125, and Her2Neu in invasive micropapillary carcinoma of the urinary tract and typical invasive urothelial carcinoma with retraction artifact. 2009; 22: 660-667.
- [12] Wang S, You L, Dai M, Zhao Y. Mucins in pancreatic cancer: A well-established but promising family for diagnosis, prognosis and therapy. 2020; 24: 10279-10289.
- [13] Gao J, McConnell M, Yu B, Li J, Balko J, Black E, Johnson J, Lloyd M, Altiock S and Haura E. MUC1 is a downstream target of STAT3 and regulates lung cancer cell survival and invasion. 2009; 35: 337-345.
- [14] Mikami Y, Hisatsune A, Tashiro T, Isohama Y, Katsuki H. Hypoxia enhances MUC1 expression in a lung adenocarcinoma cell line. 2009; 379: 1060-1065.
- [15] Ham S, Kwon T, Bak Y, Yu J, Hong J, Lee S, Yu D and Yoon D. Mucin 1-mediated chemo-resistance in lung cancer cells. 2016; 5: e185.

- [16] Sahraei M, Roy L, Curry J, Teresa T, Nath S, Besmer D, Kidiyoor A, Dalia R, Gendler S and Mukherjee PJO. MUC1 regulates PDGFA expression during pancreatic cancer progression. 2012; 31: 4935-4945.
- [17] Nath S, Daneshvar K, Roy L, Grover P, Kidiyoor A, Mosley L, Sahraei M and Mukherjee PJO. MUC1 induces drug resistance in pancreatic cancer cells via upregulation of multidrug resistance genes. 2013; 2: e51.
- [18] Sivridis E, Giatromanolaki A, Koukourakis M, Georgiou L and Anastasiadis PJH. Patterns of episialin/MUC1 expression in endometrial carcinomas and prognostic relevance. 2002; 40: 92-100.
- [19] Lee H, Park H, Yu H, Na K, Oh K and Lee EJP. Dendritic Cell-Targeted pH-Responsive Extracellular Vesicles for Anticancer Vaccination. 2019; 11:
- [20] Lepisto A, Moser A, Zeh H, Lee K, Bartlett D, McKolanis J, Geller B, Schmotzer A, Potter D, Whiteside T, Finn O and Ramanathan RJCt. A phase I/II study of a MUC1 peptide pulsed autologous dendritic cell vaccine as adjuvant therapy in patients with resected pancreatic and biliary tumors. 2008; 6: 955-964.
- [21] Kondo H, Hazama S, Kawaoka T, Yoshino S, Yoshida S, Tokuno K, Takashima M, Ueno T, Hinoda Y and Oka MJAr. Adoptive immunotherapy for pancreatic cancer using MUC1 peptide-pulsed dendritic cells and activated T lymphocytes. 2008; 28: 379-387.
- [22] Shindo Y, Hazama S, Maeda Y, Matsui H, Iida M, Suzuki N, Yoshimura K, Ueno T, Yoshino S, Sakai K, Suehiro Y, Yamasaki T, Hinoda Y and Oka MJJotm. Adoptive immunotherapy with MUC1-mRNA transfected dendritic cells and cytotoxic lymphocytes plus gemcitabine for unresectable pancreatic cancer. 2014; 12: 175.
- [23] Ge C, Li R, Song H, Geng T, Yang J, Tan Q, Song L, Wang Y, Xue Y, Li Z, Dong S, Zhang Z, Zhang N, Guo J, Hua L, Chen S and Song XJBc. Phase I clinical trial of a novel autologous modified-DC vaccine in patients with resected NSCLC. 2017; 17: 884.
- [24] Teramoto K, Ozaki Y, Hanaoka J, Sawai S, Tezuka N, Fujino S, Daigo Y and Kontani KJTaimo. Predictive biomarkers and effectiveness of MUC1-targeted dendritic-cell-based vaccine in patients with refractory non-small cell lung cancer. 2017; 9: 147-157.
- [25] Scheid E, Major P, Bergeron A, Finn O, Salter R, Eady R, Yassine-Diab B, Favre D, Peretz Y, Landry C, Hotte S, Mukherjee S, Dekaban G, Fink C, Foster P, Gaudet J, Garipey J, Sekaly R, Lacombe L, Fradet Y and Foley RJCir. Tn-MUC1 DC Vaccination of Rhesus Macaques and a Phase I/II Trial in Patients with Nonmetastatic Castrate-Resistant Prostate Cancer. 2016; 4: 881-892.
- [26] Wierecky J, Müller M, Wirths S, Halder-Oehler E, Dörfel D, Schmidt S, Häntschel M, Brugger W, Schröder S, Horger M, Kanz L and Brossart PJCr. Immunologic and clinical responses after vaccinations with peptide-pulsed dendritic cells in metastatic renal cancer patients. 2006; 66: 5910-5918.

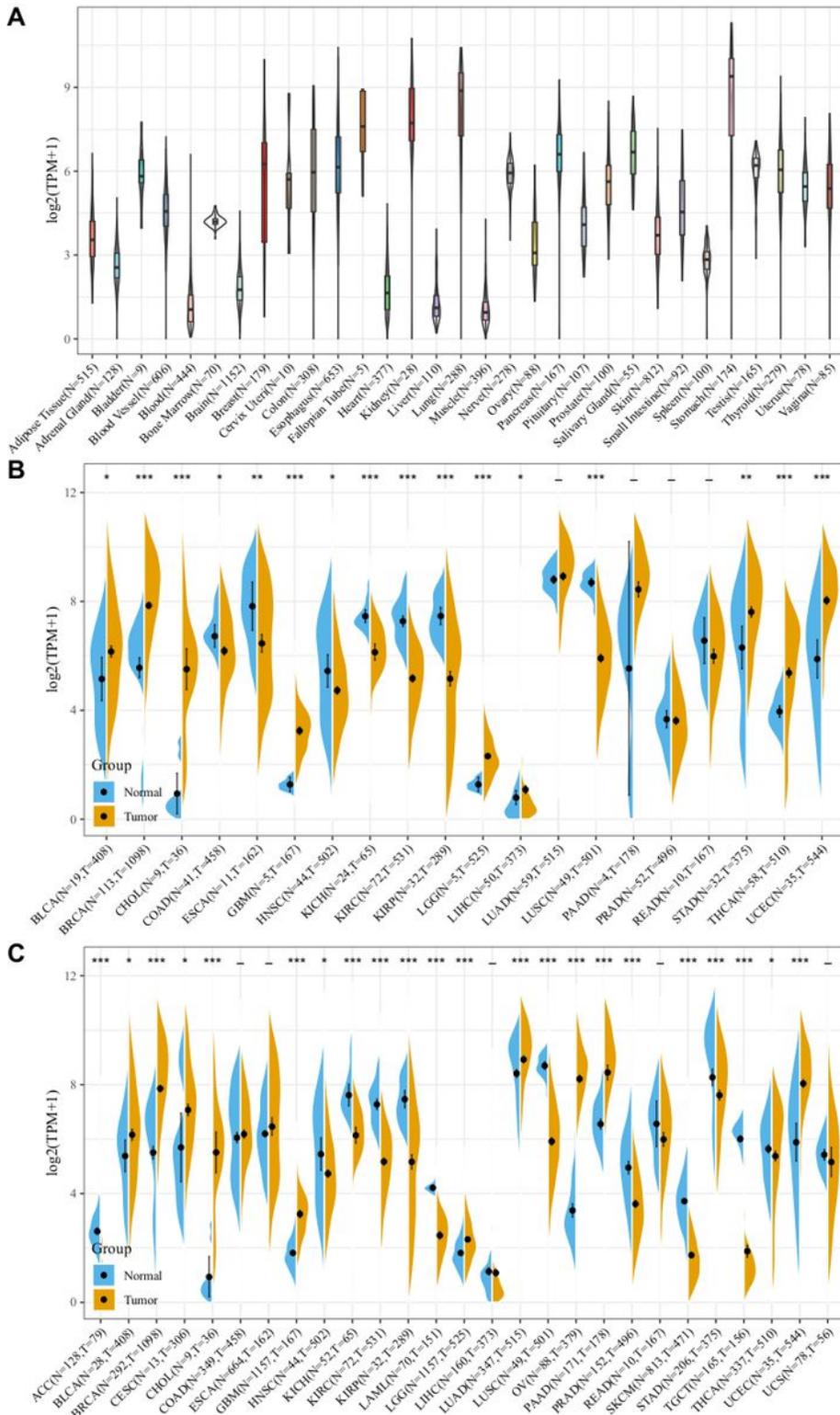
- [27] Morse M, Niedzwiecki D, Marshall J, Garrett C, Chang D, Aklilu M, Crocenzi T, Cole D, Dessureault S, Hobeika A, Osada T, Onaitis M, Clary B, Hsu D, Devi G, Bulusu A, Annechiarico R, Chadaram V, Clay T and Lyerly HJAos. A randomized phase II study of immunization with dendritic cells modified with poxvectors encoding CEA and MUC1 compared with the same poxvectors plus GM-CSF for resected metastatic colorectal cancer. 2013; 258: 879-886.
- [28] Zhang H, Zhao H, He X, Xi F, Liu JJ. JAK-STAT Domain Enhanced MUC1-CAR-T Cells Induced Esophageal Cancer Elimination. 2020; 12: 9813-9824.
- [29] Mei Z, Zhang K, Lam A, Huang J, Qiu F, Qiao B and Zhang YJ. MUC1 as a target for CAR-T therapy in head and neck squamous cell carcinoma. 2020; 9: 640-652.
- [30] Wei X, Lai Y, Li J, Qin L, Xu Y, Zhao R, Li B, Lin S, Wang S, Wu Q, Liang Q, Peng M, Yu F, Li Y, Zhang X, Wu Y, Liu P, Pei D, Yao Y and Li PJ. PSCA and MUC1 in non-small-cell lung cancer as targets of chimeric antigen receptor T cells. 2017; 6: e1284722.
- [31] Zhou R, Yazdanifar M, Roy L, Whilding L, Gavrill A, Maher J and Mukherjee PJ. CAR T Cells Targeting the Tumor MUC1 Glycoprotein Reduce Triple-Negative Breast Cancer Growth. 2019; 10: 1149.
- [32] Yazdanifar M, Zhou R, Grover P, Williams C, Bose M, Moore L, Wu S, Maher J, Dreau D and Mukherjee AJC. Overcoming Immunological Resistance Enhances the Efficacy of A Novel Anti-tMUC1-CAR T Cell Treatment against Pancreatic Ductal Adenocarcinoma. 2019; 8:
- [33] You F, Jiang L, Zhang B, Lu Q, Zhou Q, Liao X, Wu H, Du K, Zhu Y, Meng H, Gong Z, Zong Y, Huang L, Lu M, Tang J, Li Y, Zhai X, Wang X, Ye S, Chen D, Yuan L, Qi L and Yang LJ. Phase 1 clinical trial demonstrated that MUC1 positive metastatic seminal vesicle cancer can be effectively eradicated by modified Anti-MUC1 chimeric antigen receptor transduced T cells. 2016; 59: 386-397.
- [34] Jardim D, Goodman A, de Melo Gagliato D and Kurzrock R. The Challenges of Tumor Mutational Burden as an Immunotherapy Biomarker. 2021; 39: 154-173.
- [35] Petrelli F, Ghidini M, Ghidini A and Tomasello G. Outcomes Following Immune Checkpoint Inhibitor Treatment of Patients With Microsatellite Instability-High Cancers: A Systematic Review and Meta-analysis. 2020; 6: 1068-1071.
- [36] Blass E and Ott P. Advances in the development of personalized neoantigen-based therapeutic cancer vaccines. 2021;
- [37] Yin L, Zhou L and Xu R. Identification of Tumor Mutation Burden and Immune Infiltrates in Hepatocellular Carcinoma Based on Multi-Omics Analysis. 2020; 7: 599142.
- [38] Li X, Pasche B, Zhang W and Chen K. Association of MUC16 Mutation With Tumor Mutation Load and Outcomes in Patients With Gastric Cancer. 2018; 4: 1691-1698.

- [39] Slattery M, Pellatt D, Mullany L, Wolff R, Herrick JJG, chromosomes and cancer. Gene expression in colon cancer: A focus on tumor site and molecular phenotype. 2015; 54: 527-541.
- [40] Byrd J and Bresalier RJCmr. Mucins and mucin binding proteins in colorectal cancer. 2004; 23: 77-99.
- [41] Ribas A and Wolchok JJS. Cancer immunotherapy using checkpoint blockade. 2018; 359: 1350-1355.
- [42] Sandoval J, Mendez-Gonzalez J, Nadal E, Chen G, Carmona F, Sayols S, Moran S, Heyn H, Vizoso M, Gomez A, Sanchez-Cespedes M, Assenov Y, Müller F, Bock C, Taron M, Mora J, Muscarella L, Liloglou T, Davies M, Pollan M, Pajares M, Torre W, Montuenga L, Brambilla E, Field J, Roz L, Lo Iacono M, Scagliotti G, Rosell R, Beer D and Esteller MJJocoojotASoCO. A prognostic DNA methylation signature for stage I non-small-cell lung cancer. 2013; 31: 4140-4147.
- [43] Jung G, Hernández-Illán E, Moreira L, Balaguer F, Goel AJNrG and hepatology. Epigenetics of colorectal cancer: biomarker and therapeutic potential. 2020; 17: 111-130.
- [44] Xu R, Wei W, Krawczyk M, Wang W, Luo H, Flagg K, Yi S, Shi W, Quan Q, Li K, Zheng L, Zhang H, Caughey B, Zhao Q, Hou J, Zhang R, Xu Y, Cai H, Li G, Hou R, Zhong Z, Lin D, Fu X, Zhu J, Duan Y, Yu M, Ying B, Zhang W, Wang J, Zhang E, Zhang C, Li O, Guo R, Carter H, Zhu J, Hao X and Zhang KJNm. Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma. 2017; 16: 1155-1161.
- [45] Visvanathan K, Fackler M, Zhang Z, Lopez-Bujanda Z, Jeter S, Sokoll L, Garrett-Mayer E, Cope L, Umbricht C, Euhus D, Forero A, Storniolo A, Nanda R, Lin N, Carey L, Ingle J, Sukumar S and Wolff AJJocoojotASoCO. Monitoring of Serum DNA Methylation as an Early Independent Marker of Response and Survival in Metastatic Breast Cancer: TBCRC 005 Prospective Biomarker Study. 2017; 35: 751-758.
- [46] Yamada N, Nishida Y, Tsutsumida H, Hamada T, Goto M, Higashi M, Nomoto M and Yonezawa SJCr. MUC1 expression is regulated by DNA methylation and histone H3 lysine 9 modification in cancer cells. 2008; 68: 2708-2716.
- [47] Rajabi H, Tagde A, Alam M, Bouillez A, Pitroda S, Suzuki Y and Kufe DJO. DNA methylation by DNMT1 and DNMT3b methyltransferases is driven by the MUC1-C oncoprotein in human carcinoma cells. 2016; 35: 6439-6445.
- [48] Yokoyama S, Higashi M, Kitamoto S, Oeldorf M, Knippschild U, Kornmann M, Maemura K, Kurahara H, Wiest E, Hamada T, Kitazono I, Goto Y, Tasaki T, Hiraki T, Hatanaka K, Mataka Y, Taguchi H, Hashimoto S, Batra S, Tanimoto A, Yonezawa S and Hollingsworth MJO. Aberrant methylation of MUC1 and MUC4 promoters are potential prognostic biomarkers for pancreatic ductal adenocarcinomas. 2016; 7: 42553-42565.

[49] Yokoyama S, Hamada T, Higashi M, Matsuo K, Maemura K, Kurahara H, Horinouchi M, Hiraki T, Sugimoto T, Akahane T, Yonezawa S, Kornmann M, Batra S, Hollingsworth M and Tanimoto AJC. Predicted Prognosis of Patients with Pancreatic Cancer by Machine Learning. 2020; 26: 2411-2421.

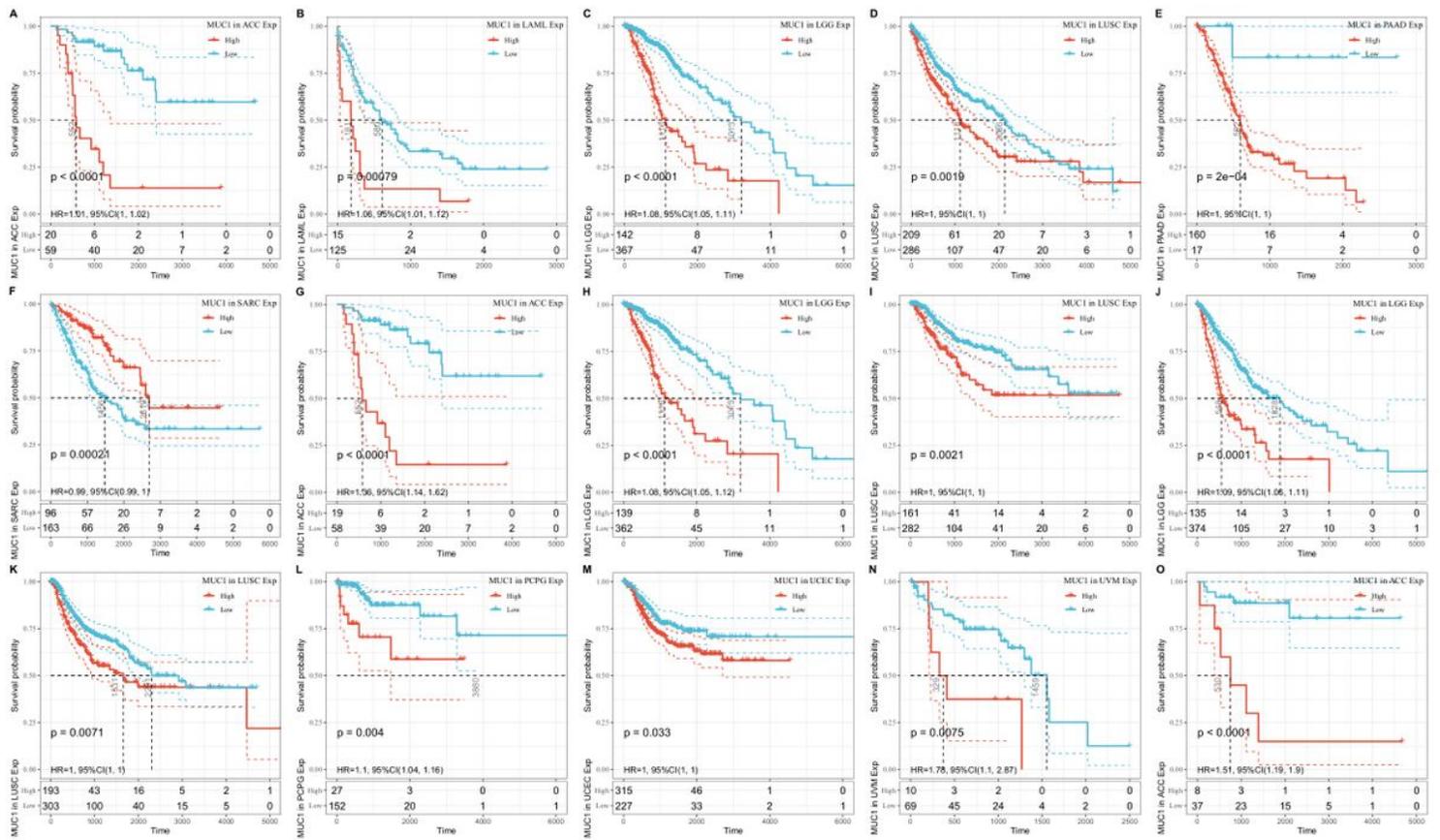
[50] Rooney M, Shukla S, Wu C, Getz G and Hacohen NJC. Molecular and genetic properties of tumors associated with local immune cytolytic activity. 2015; 160: 48-61.

## Figures



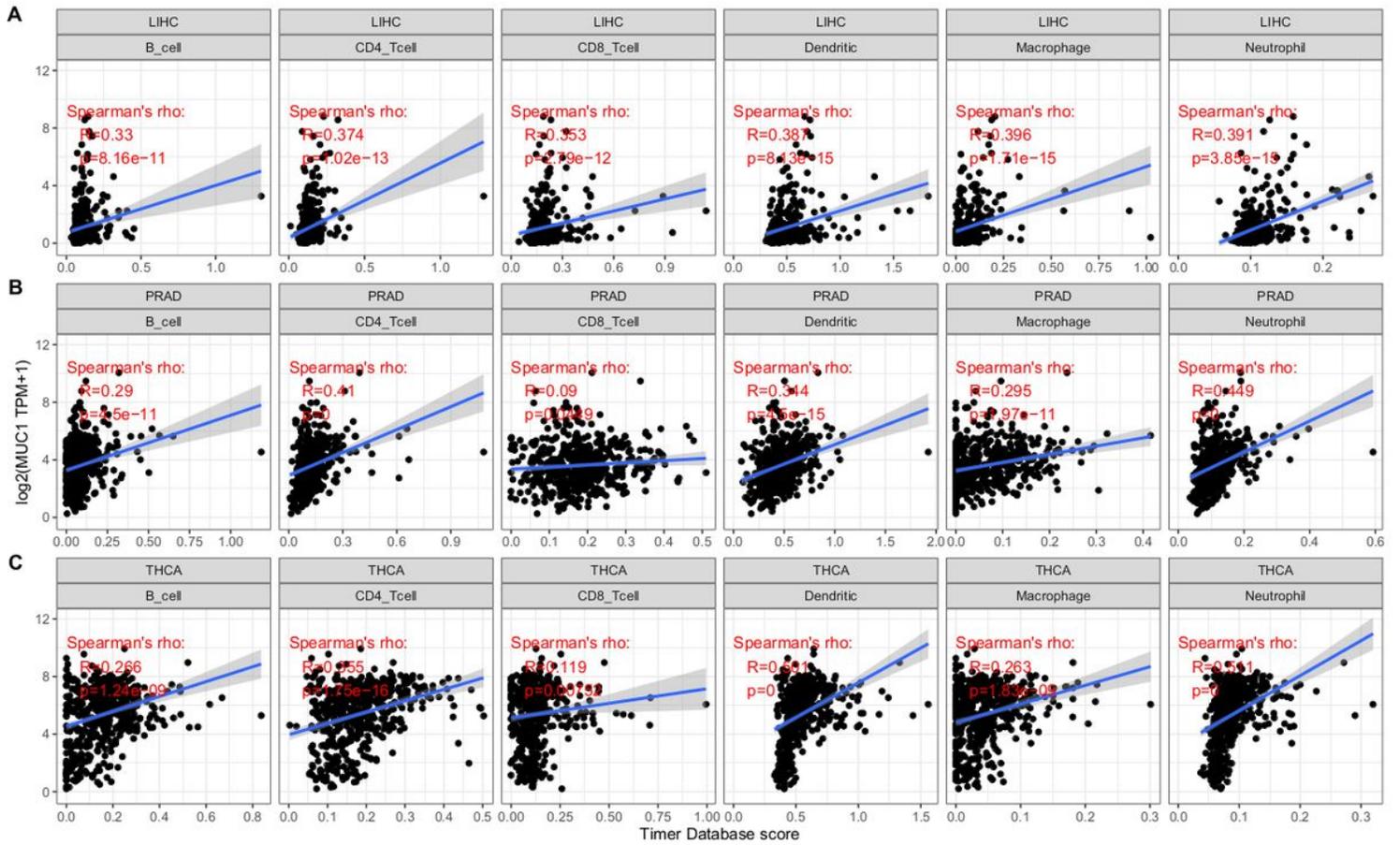
**Figure 1**

The expression level of MUC1 in cancer patients. (A) The MUC1 expression of 31 kinds of normal tissues in GTEx database. (B) The difference of MUC1 expression between cancer tissues and normal tissues of patients with 20 kinds of cancers in TCGA database. (C) The difference of MUC1 expression between cancer tissue and normal tissue of 27 kinds of cancer patients was analyzed by TCGA and GTEx database. \*P < 0.05. \*\*P < 0.01. \*\*\*P < 0.001.



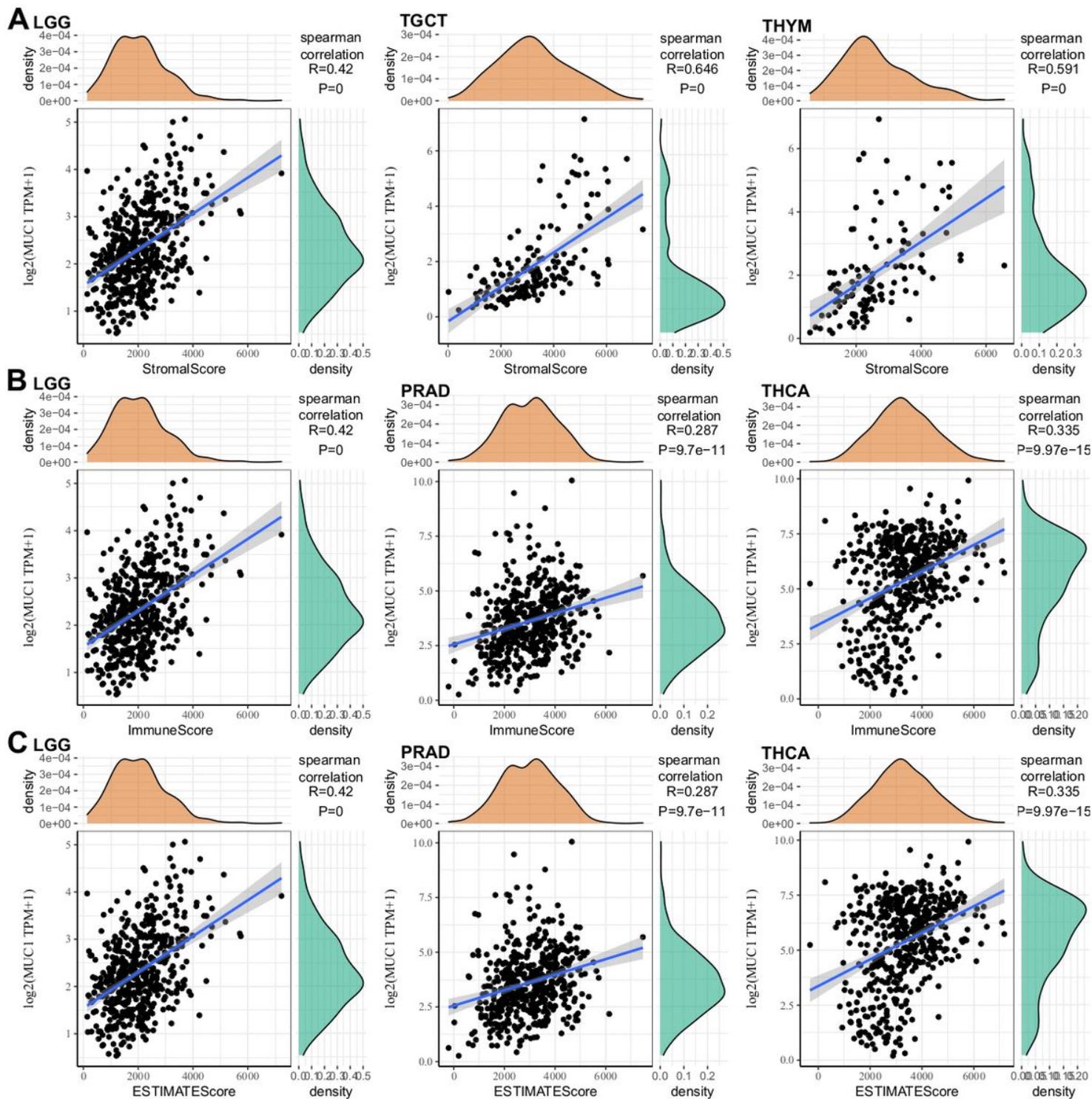
**Figure 2**

The relationship between the expression of MUC1 and the prognosis of patients with 33 kinds of cancers in TCGA database. (A-F) Overall survival. (G-I) Disease-specific survival. (J-N) Disease-free survival. (O) Progression-free survival.



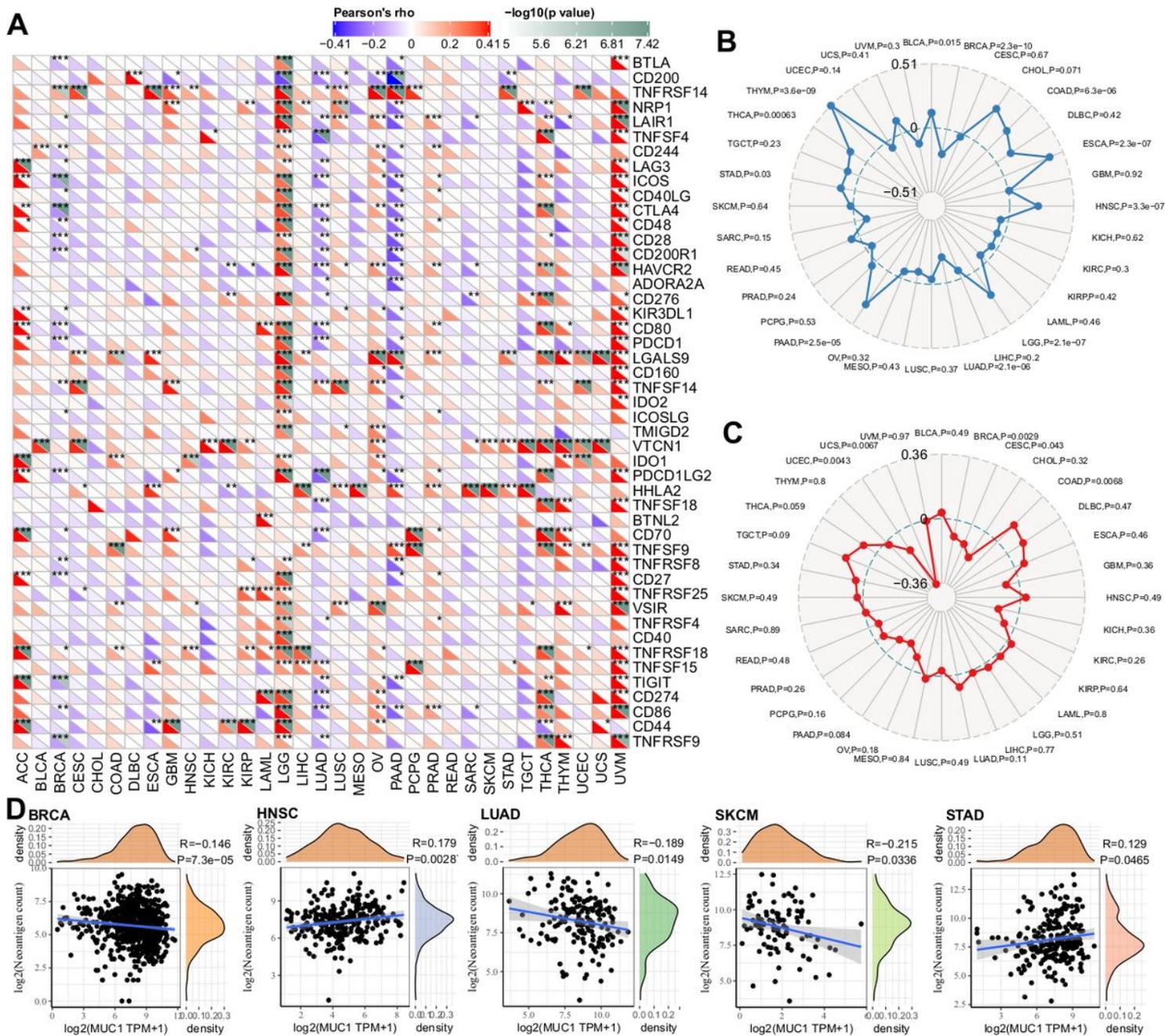
**Figure 3**

The relationship between the expression level of MUC1 and the infiltration level of different kinds of immune cells in cancer patients. (A, B, C) are the three tumors, which are LIHC, PRAD and THCA, with the strongest correlation with immune cell infiltration.



**Figure 4**

The relationship between the expression level of MUC1 and the score of ESTIMATE algorithm in cancer patients. (A, B, C) are StromalScore, ImmuneScore and ESTIMATEScore



**Figure 5**

The relationship between MUC1 expression and Immunoreactivity. (A) The correlation between MUC1 expression level and immune checkpoint gene expression. \* $P < 0.05$ . \*\* $P < 0.01$ . \*\*\* $P < 0.001$ . (B) The relationship between MUC1 expression level and tumor mutation burden (TMB). (C) MUC1 expression level and microsatellite instability (MSI) in cancer patients. (D) Correlation between tumor neoantigen and MUC1 expression level.

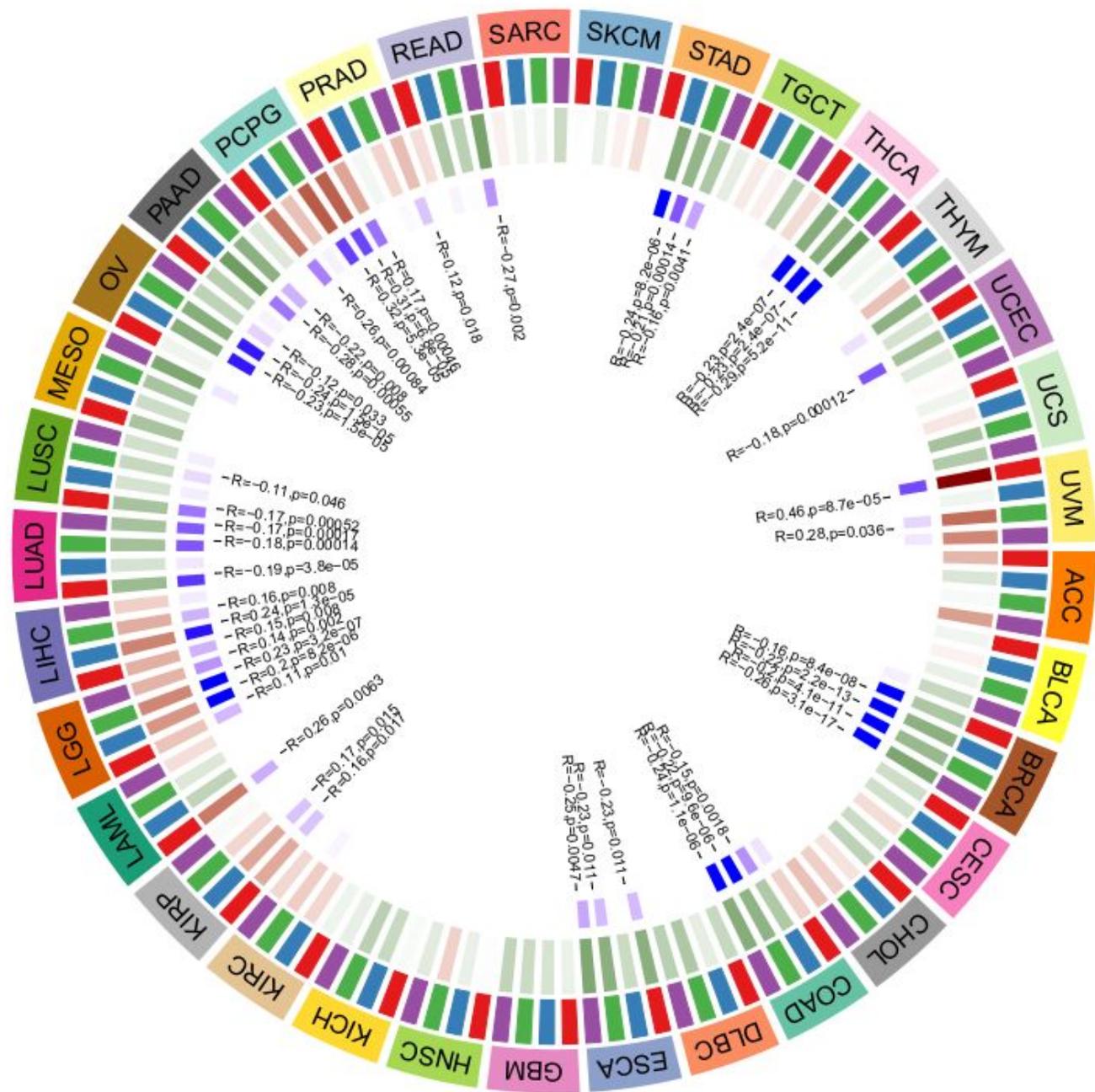
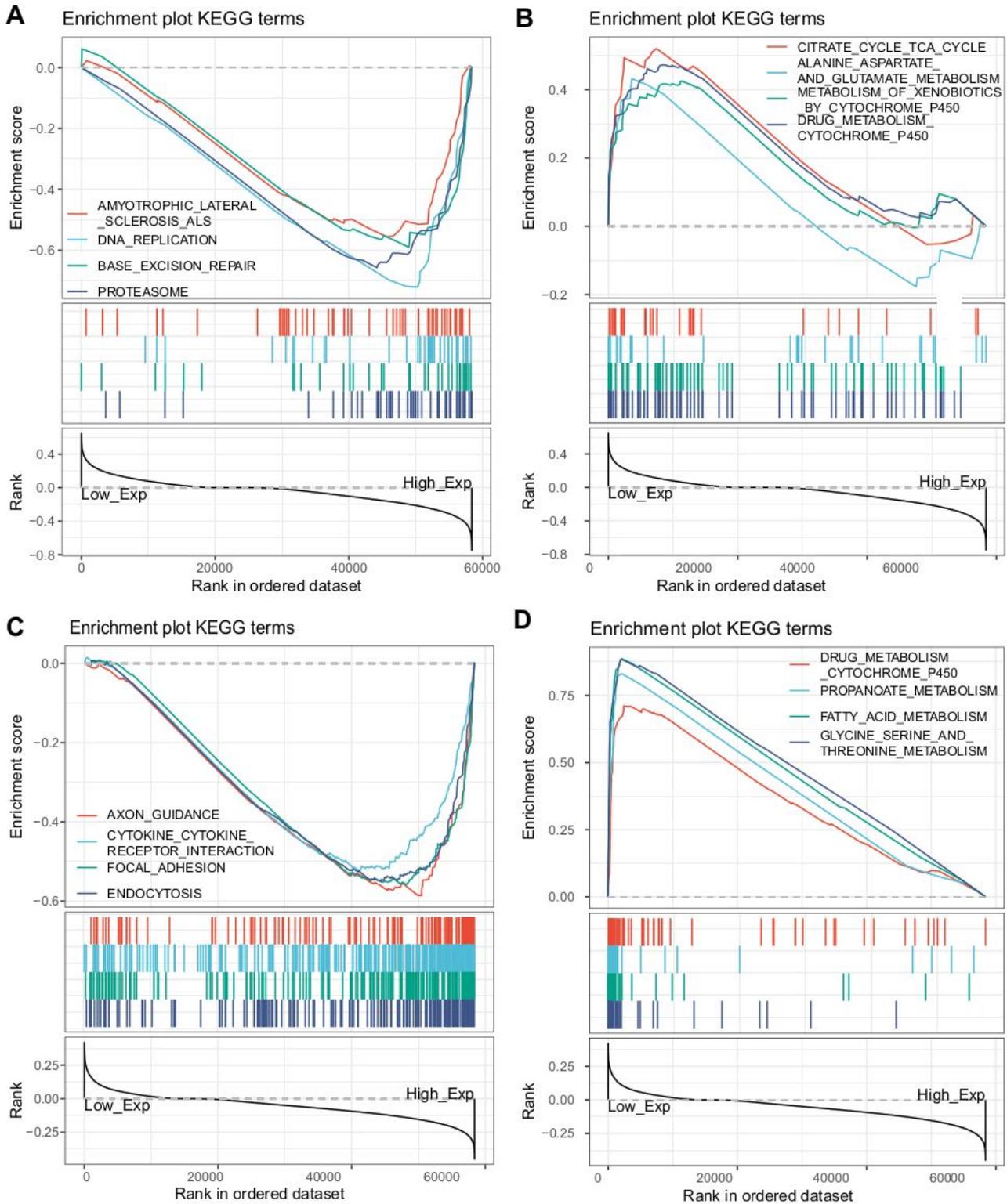


Figure 6

The relationship between the expression level of MUC1 and 4 methyltransferases.



**Figure 7**

KEGG pathway analysis of MUC1 by GSEA in ACC and LIHC whose prognosis is closely related to the expression of MUC1. (A, C) are the pathways negatively correlated with MUC1 expression in ACC and LIHC, respectively. (B, D) are the pathways that are positively correlated with MUC1 expression in ACC and LIHC, respectively.

## Supplementary Files

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

- [coverletters.docx](#)