MUC4 mutation correlates with tumor mutation burden and the immune microenvironment in colorectal cancer

Ting Li  
Xinjiang Medical University  https://orcid.org/0000-0001-9470-0451

Wenjia Hui  
People's Hospital of Xinjiang Uygur Autonomous Region

Halina Halike  
People's Hospital of Xinjiang Uygur Autonomous Region

Feng Gao  (xjgf@sina.com)  
People's Hospital of Xinjiang Uygur Autonomous Region  https://orcid.org/0000-0001-7355-2672

**Primary research**

**Keywords:** colorectal cancer (CRC)  
tumor mutation burden  
MUC4  
tumor-infiltrating immune cells

**DOI:** https://doi.org/10.21203/rs.3.rs-101882/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.  
[Read Full License](#)
Abstract

Background: Immunotherapy is a new strategy for Colorectal cancer (CRC) treatment. Tumor mutation burden (TMB) may act as an emerging biomarker for predicting responses to immunotherapy. Nevertheless, no studies investigate if these gene mutations correlate to TMB and tumor-infiltrating immune cells.

Methods: Somatic mutation data for CRC samples were obtained from The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC) datasets. Then, we investigated the relationship between these mutant genes, TMB and overall survival outcomes. GSEA analysis was performed to explore the underlying mechanism of mutant gene. Finally, we further verified the connection between gene mutations and immune response.

Results: We identified 17 common mutant genes shared by both two cohorts. Further analysis found that MUC4 mutation was strongly related to higher TMB and predicted a poorer prognosis. Moreover, functional enrichment analysis of samples with MUC4 mutation revealed that they were involved in regulating the natural killer cell mediated cytotoxicity signaling pathway. Significant changes in the proportion of the immune cells of CD8 T cells, activated NK cells, M1 macrophages and resting memory CD4 T cells were observed using the CIBERSORT algorithm.

Conclusions: Our research revealed that MUC4 mutation significantly correlated with increased TMB, a worse prognosis and modulating the immune microenvironment, which may be considered a biomarker to predict the outcome of the immune response in colorectal cancer.

Background

Colorectal cancer (CRC) is one of the most common and aggressive malignancies. It is also a heterogeneous disease at the molecular level resulting from complex genetic and epigenetic alterations [1]. The incidence of CRC tends to increase year by year in China [2], Though the development at full speed of treatments, such as targeted therapies, chemotherapy, surgery and radiotherapy, high recurrence rate and poor prognosis remain a concern [3].

Breakthroughs have been made in the treatment of CRC by immunotherapy. It is considered to be a more precise treatment with high potential for the future, particularly for deficient mismatch repair (MMR) or high microsatellite instability (MSI-H). Researchers have recently found that the tumor mutation burden (TMB) is an emerging characteristic of cancer for assessing the effectiveness of immunotherapy [4, 5]. Tumor cells with mutations can be translated and might generate new antigens that can be recognized by the immune system [6, 7]. Patients with high TMB can benefit from immunotherapy according to several previous studies [4, 8]. At present, the predictive value of the therapeutic effect of TMB on CRC immunotherapy has been confirmed in several clinical studies [9–11]. The results indicated that the TMB test could help reclassify CRC patients. Increased mutation load is associated with microsatellite instability [12]. Another study result suggests that TMB has a certain correlation with chemotherapy
results in CRC patients. Therefore, TMB potential to act as sensitive biomarkers or predicting immunotherapy response. Nevertheless, the immune landscape between tumors with different TMB was unclear in CRC patients.

At present, we explored the somatic mutations of CRC patients in TCGA and ICGC data portal. The common mutant genes shared by these two cohorts were identified. Then, we investigated the relationship between these mutant genes, TMB and overall survival outcomes. Finally, we further verified the connection between gene mutations and immune response. Overall, our study aimed to provides additional prognostic information for immunotherapy in CRC.

**Methods**

**Data source**

The mutation data were obtained from the TCGA (https://tcga-data.nci.nih.gov/tcga/) and ICGC dataset (https://dcc.icgc.org/). The TCGA dataset including 510 American CRC samples, 434 patients had complete clinical information. The ICGC portal included 321 Chinese CRC samples.

**Calculation of TMB**

TMB was referred to the average number of genes mutations per megabase (Mb). TSV files, including data of somatic variants for Chinese CRC samples, were annotated and visualized by the GenVisR package. The project “TCGA-COAD” including data of somatic variants for American CRC samples, which downloaded as a MAF file that was detected using VarScan and visualized with the GenVisR package. The TMB estimate for each sample is equal to the total number of mutation frequency/38. The 38 Mb was routinely used to estimate the exome size of a human.

**Gene set enrichment analysis**

CRC patients from the TCGA dataset were divided into two groups according to the mutation status of MUC4. Annotated gene sets of c2.cp.kegg.v7.2.symbols.gmt was considered as the reference gene sets in GSEA software (4.0.1), p-value <0.05 was used as the cut-off values.

**Estimate of immune cells infiltration**

CIBERSORT is a deconvolution tool that can analyze immune cell landscapes according to gene expression data [13]. The CIBERSORT was performed using the reference of the leukocyte subtype with 100 permutations, at a threshold of P-value < 0.05. The result of immune cells distribution was visualized using the vioplot R package.

**Statistical analysis**

Survminer R package was used to draw the survival curve. Kaplan-Meier survival curves and log-rank test were employed to analyze the prognosis of CRC patients. Univariate and multivariate Cox regression was
utilized to investigated if the MUC4 mutation was an independent prognostic parameter. Mann-Whitney U test was performed to identify whether the mutant genes were associated with TMB. Statistical analyses were performed using R software (version 4.0.0). *P*-value < 0.05 was considered significant.

Results

The mutation spectrum in CRC

After analyzing the data, we discover the missense mutations are the most common. The waterfall diagram revealed the 30 frequently mutated genes from the TCGA cohort (Figure 1A). At the same time, we explored 30 frequently mutated genes from the ICGC cohort contained Chinese CRC samples (Figure 1B). The results showed that some genes frequently mutated in both two cohorts. Thus, the Venn diagram was used to identify the intersecting part and visualizes the 17 common genes shared by these two cohorts, including APC, TP53, TTN, KRAS, MUC16, MUC4, SYNE1, FLG, FAT4, OBSCN, FAT3, RYR2, PIK3CA, FBXW7, DNAH11, MUC5B and ZFHX4 (Figure 1C).

MUC4 mutation is tightly linked with high TMB and inferior prognosis in CRC patients

The result revealed that mutation in TTN, MUC16, MUC4, SYNE1, FLG, FAT4, OBSCN, FAT3, RYR2, PIK3CA, FBXW7, DNAH11, MUC5B and ZFHX4 presented relatively higher TMB in the CRC samples (Figure 2A). Kaplan-Meier survival analysis was conducted to investigate the potential relationship between these gene mutations and the prognosis of CRC patients (Figure 2B). Among the mutated genes only MUC4 mutation (P= 0.002) was related to an inferior prognosis. The univariate and multivariate Cox regression analyses were conformed to determine if MUC4 mutation was an independent prognostic parameter (Table 1). The results obtained from the univariate Cox model suggested that MUC4 mutation was correlated to a worsening of the outcome (HR=2.084 p=0.003, 95% CI [1.286-3.376]). Afterward, the entire variables were analyzed via multivariate Cox regression. Further statistical tests revealed that the MUC4 mutation leads to an inferior prognosis (HR=2.093, p =0.003 95% CI [1.290–3.396]), suggesting that MUC4 mutation could act as an independent prognostic parameter in CRC patients.

Table 1. Exploration of the independent prognostic parameters in colorectal cancer.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Univariate</th>
<th>Multivariate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>age(year) (&lt;65, ≥ 65)</td>
<td>1.57 (0.99-2.50)</td>
<td>0.06</td>
</tr>
<tr>
<td>gender (male, female)</td>
<td>1.37 (0.88-2.12)</td>
<td>0.16</td>
</tr>
<tr>
<td>stage (low, high)</td>
<td>2.82 (1.79-4.43)</td>
<td>0.00</td>
</tr>
<tr>
<td>TMB (low, high)</td>
<td>1.00 (0.99-1.01)</td>
<td>0.91</td>
</tr>
<tr>
<td>MUC4 (wide, mutant)</td>
<td>2.08 (1.29-3.38)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
CI: confidence interval; HR: hazard ratio.

**Enrichment pathway analysis of MUC4 mutation**

We performed GSEA enrichment analysis to explore the underlying mechanism of MUC4 mutation—the result revealed that natural killer cell mediated cytotoxicity were mainly enriched in patients with MUC4 mutation. Thus, we concluded that MUC4 mutation regulated pathways referred to the immune system (Figure 3).

**Correlation of MUC4 mutation with tumor-infiltrating immune cells in CRC**

We further explored the connection between MUC4 mutation and tumor-infiltrating immune cells in CRC patients. We compared the landscape of the immune cell in the MUC4 mutation group and wild group by CIBERSORT algorithm (Figure 4A). The difference analytical results showed that CD8 T cells, activated NK cells and M1 macrophages were comparatively higher infiltrating in tumors with MUC4 mutation, while resting memory CD4 T cells were higher infiltrating in the wild group (Figure 4B). The result of Correlation analysis revealed that CD8+T cells had a negative association with resting memory CD4+T cells (Figure 4C).

**Discussion**

The tumorigenesis of CRC is commonly known as a multi-stage process, which involves a series of genetic changes interacting with the tumor microenvironment (TME) [14, 15]. This study was conducted to analyze mutations in CRC samples from the TCGA dataset and ICGC dataset. We found a high mutation frequency of MUC4 in both cohorts. We were able to show that mutation in MUC4 was significantly correlated with increased TMB and indicated a worse prognosis. Moreover, samples with MUC4 mutations were enriched in the NK cell mediated cytotoxicity signaling pathway. The result of Immune cells infiltration demonstrated that MUC4 mutated with higher infiltration levels of memory CD8 T cells, activated NK cells and M1 macrophages, but less infiltrated in resting memory CD4 T cells.

MUC4 exhibited a relatively high mutation rate in CRC patients. Mucins are important glycoproteins, falling into two subgroups: the membrane-bound mucins and the secreted mucins [16, 17]. MUC4 is a membrane-bound mucin that can be seen as a useful biomarker of carcinogenic progression [18–21]. Its biological effects are mainly used in pancreatic cancer, lung cancer, ovarian cancer and oesophageal cancer [22–26]. MUC4 is expressed in epithelial cells of the digestive tract, as well as some goblet cells [27]. Abnormal MUC4 expression is correlated with the progression of esophageal cancer [23] and CRC [28, 29]. Besides, an increased expression in MUC4 demonstrates a poor survival outcome, specifically in the early stages of the CRC [30]. Due to the size of the MUC4 gene is relatively large, this may increase its mutation probability. Thus, MUC4 is the most common mutated gene in the state of stress exposure, for example, aging or nicotine treatment [31, 32]. MUC4 mutation was significant in Renal Carcinoma and was related to survival outcomes. The mutation of MUC4 could be tightly linked with smoking in kidney cancer [13] [33]. In our study, we found that CRC samples with a mutation in MUC4 had a significantly
worse survival outcome and were highly correlated with increased TMB. A high TMB referred to the production of mutation-related neoantigens, thereby increased tumor-infiltrating immune cells [34, 35]. It is inferred that the high TMB have affected response to immunotherapy [36]. Hence, we postulate that the immune response was depressed by the muc4 mutation.

We used GSEA analysis to explore the potential biological function of MUC4 mutation in CRC. The result showed that the NK cell-mediated cytotoxicity pathway was the major mechanism for MUC4 mutation influencing the outcome of the CRC patients. Together these functional analyses suggested that the MUC4 mutation had an important function in regulating the immune response in CRC.

Previous studies have indicated that the tumor immune microenvironment (TIM) has a significant meaning in predicting prognosis and evaluating therapeutic efficacy [8]. In our study, we found that MUC4 mutation was significantly related to increased CD8 T cells, NK cells and M1 macrophages. CD8 + T cells play essential roles in tumor immunity. We hypothesized that an obvious increase in the abundance of CD8 + T cells in samples with MUC4 mutated was due to an increased number of tumor neoantigens, leading to immune response activation. Interestingly, the result of the correlation analysis revealed that CD8 + T cells had a negative association with resting memory CD4 + T cells. We also discovered a higher proportion of activated NK cells infiltrated in patients with MUC4 mutated, indicating that MUC4 mutation could stimulate activation of NK cells. Studies show that NK cells were major players in the early detection, control of tumor growth and metastasis [37–39]. Previous studies reported that M2-tumor-associated macrophages (M2-TAMs) were common in malignant tumors, which creates a favorable microenvironment for tumor progression [40, 41]. M2-TAMs were found to be linked with poor survival outcome, while M1-TAMs show the opposite results [42–44]. M1 macrophages could be transformed into M2 with tumor progression. Interestingly, our study found that MUC4 mutated samples were more infiltrated in M1 macrophages. This is probably caused by an early staging of the majority cases in the TCGA datasets. Therefore, further studies are required to verify this hypothesis.

Conclusions

In conclusion, our data implicate that MUC4 had a relatively high mutation rate in colorectal cancer. MUC4 mutation was closely related to increased TMB and a worse prognosis. Finally, the NK cell mediated cytotoxicity signaling pathway were mainly enriched in samples with MUC4 mutations, and we also verified the association between TMB and immune cell infiltration. These findings identify that MUC4 mutation may be considered a biomarker to predict the outcome of the immune response in colorectal cancer.

Abbreviations

CRC: Colorectal cancer; GSEA: Gene set enrichment analysis; ICGC: International Cancer Genome Consortium; M2-TAMs: M2-tumor-associated macrophages; MSI-H: high microsatellite instability; MUC4:
Mucins 4; TCGA: The Cancer Genome Atlas; TIM: tumor immune microenvironment; TMB: tumor mutation burden; MMR: mismatch repair

Declarations

Acknowledgements

Not applicable.

Authors’ contributions

TL conceptualized and planned the study, wrote the manuscript. WJ.H and HH analyzed the interpreted results. FG designed the study, revised it critically for important intellectual information, and approved the final version for publication. The final manuscript was read and approved by all authors.

Funding

Not applicable.

Availability of data and materials

The datasets analysed in this study can be found in the TCGA (https://tcga-data.nci.nih.gov/tcga/) and ICGC dataset (https://dcc.icgc.org/).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References


**Figures**
Figure 1

Primary mutation of genes in colorectal cancer. (A) Waterfall of the top 30 mutated genes in the CRC patients from the TCGA cohort. (B) Waterfall of the top 30 mutated genes in the CRC patients from the ICGC cohort. (C) Venn diagram of the top 30 mutated genes shared by two cohorts.
Figure 2

Correlation of mutated Gene with TMB and prognosis of CRC patients. (A) The majority of mutated genes are highly associated with higher TMB. *p<0.05 ** p<0.01; *** p<0.001. (B) K–M curves of patients with gene mutations. WT, wild type; MT, mutant type.
Figure 3

MUC4 mutation correlated enrichment gene analysis with GSEA. NK cell-mediated cytotoxicity was mainly involved in samples with MUC4 mutation.
Figure 4

MUC4 mutation is closely related to tumor-infiltrating immune cells. (A) The landscape of tumor-infiltrating immune cells in TCGA samples. (B) Violin plot showing the fractions of infiltrated immune cells in MUC4 mutation groups and MUC4 wild group. (C) correlation analysis of the 22 immune cells.