Original Article

Transmembrane serine protease 2, a SARS-CoV-2 internalization protease, correlates with clinical outcome, molecular features, and immunotherapy response in colorectal cancer

Chenjing Zhu[1#], Jianfeng Wu[1#], Vishal G. Shelat[2], José María Sayagués[3], Seiichiro Yamamoto[4], Li Yang[5], Xia He[1]

[1]Department of Radiation Oncology, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & Affiliated Cancer Hospital of Nanjing Medical University, Nanjing Medical University Affiliated Cancer Hospital, Nanjing, China; [2]Department of General Surgery, Tan Tock Seng Hospital, Singapore, Singapore; [3]Pathology Service, Hospital Clínico Universitario de Salamanca and Biomedical Research Institute of Salamanca (IBSAL), University of Salamanca, Salamanca, Spain; [4]Department of Gastroenterological Surgery, Tokai University School of Medicine, Isehara, Japan; [5]Endoscopy Center, Shanghai Pulmonary Hospital, Tongji University School of Medicine, Shanghai, China

Contributions: (I) Conception and design: X He; (II) Administrative support: X He; (III) Provision of study materials or patients: C Zhu, VG Shelat, JM Sayagués, S Yamamoto; (IV) Collection and assembly of data: C Zhu, J Wu; (V) Data analysis and interpretation: C Zhu, J Wu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Xia He, MD. Department of Radiation Oncology, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & Affiliated Cancer Hospital of Nanjing Medical University, Nanjing Medical University Affiliated Cancer Hospital, 42 Baiziting, Nanjing 210009, China. Email: hexiabm@163.com.

Background: Transmembrane serine protease 2 (TMPRSS2) mediates the entry of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) into host cells. The relevant research indicates the intestine to be a target of SARS-CoV-2 infection, and thus we aimed to investigate the correlation between *TMPRSS2* expression and the prognosis, molecular features, and immunotherapy response in patients with colorectal cancer (CRC).

Methods: The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases were used in this study and a total of 1,385 patients were identified. The CIBERSORT algorithms were used to evaluate the relative infiltration levels of immune cell types in the tumor microenvironment (TME). The correlation between *TMPRSS2* expression and immunotherapy response rate was assessed in another 2 independent cohorts.

Results: *TMPRSS2* expression was significantly downregulated in cancer tissue compared to the adjacent normal tissue, and patients with CRC with lower *TMPRSS2* expression showed notably poorer prognosis. Functional enrichment analysis found that low *TMPRSS2* expression was significantly associated with cancer metastasis-related pathways. Further analysis based on the miRWalk tool and JASPAR database identified a list of microRNAs (miRNAs) and transcriptional factors targeting *TMPRSS2*. Distinct differences in immune cell infiltration and tumor purity reflected by estimate and mutant-allele tumor heterogeneity score were observed between patients with low and high *TMPRSS2* expression levels. Interestingly, patients with a low *TMPRSS2* expression level showed a higher response rate to immunotherapy.

Conclusions: CRC cells may be more resistant to SARS-CoV-2 infection due to the decreased expression of *TMPRSS2*, which could be a newly identified biomarker for prognosis and immunotherapy response prediction in patients with CRC.

Keywords: Colorectal cancer (CRC); immunotherapy; severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2); transmembrane serine protease 2 (TMPRSS2); tumor microenvironment (TME)

Submitted Aug 01, 2023. Accepted for publication Sep 24, 2023. Published online Oct 24, 2023.
doi: 10.21037/jgo-23-641
View this article at: https://dx.doi.org/10.21037/jgo-23-641

Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the cause of coronavirus disease 2019 (COVID-19) which has led to a worldwide healthcare emergency (1,2). Since December 2019, 475,391,321 cases and 6,287,850 deaths related to COVID-19 have been reported globally (https://www.worldometers.info/coronavirus/). Transmembrane serine protease 2 (TMPRSS2) has been identified as a key host cell factor for viral entry and pathogenesis of SARS-CoV-2, through a proteolytic process of the SARS-CoV-2 Spike (S) protein (3-5).

Colorectal cancer (CRC) is one of the most common types of malignant disease and the leading cause of cancer-related death globally (6). Although the past decades have witnessed advances in surgical procedures and adjuvant treatment schemes, the prognosis of patients with CRC with advanced stage remains poor (7,8). Surgical treatment followed by adjuvant chemotherapy is currently is the standard treatment scheme (9). Previous reports have suggested that patients with cancer are more susceptible to infections due to the immunosuppression caused by tumor burden and postoperative chemotherapy (10,11). In several bioinformatics analysis studies, lung and stomach cancer tissues expressed higher *TMPRSS2* expression levels than did matched healthy tissues (12-14). However, recent studies have revealed that *TMPRSS2* expression is downregulated in kidney and head and neck cancer tissues, indicating that these patients may be more resistant to SARS-CoV-2 infection (15,16). Research indicated that *TMPRSS2* is involved in the signal transduction between cancer cells and the extracellular environment (17). Ko *et al.* revealed that *TMPRSS2* expression was regulated by androgen and can promote prostate cancer cell invasion and metastasis (18). Of note, studies concerning the relationship between *TMPRSS2* expression and CRC are conflicting (14,19,20), and the potential functions of *TMPRSS2* gene and its correlation with molecular features and immunotherapy response have not been investigated in relation to CRC.

In this present study, The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases were used to investigate the association between *TMPRSS2* expression and oncogenic outcomes, molecular features, and immunotherapy response in patients with CRC. We present this article in accordance with the REMARK reporting checklist (available at https://jgo.amegroups.com/article/view/10.21037/jgo-23-641/rc).

Highlight box

Key findings

- Transmembrane serine protease 2 (TMPRSS2) was identified as a biomarker for prognosis and immunotherapy response prediction in colorectal cancer (CRC).

What is known and what is new?

- TMPRSS2 mediates the entry of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) into host cells, with research indicating the intestine to be a target of SARS-CoV-2 infection.
- We found a close correlation between *TMPRSS2* expression and the prognosis of patients with CRC and discovered that patients with low expression level had a higher response rate to immunotherapy.

What is the implication, and what should change now?

- CRC tumor cells may be more resistant to SARS-CoV-2 infection due to the decreased expression of *TMPRSS2*. *TMPRSS2* expression should be considered in the individualization of immunotherapy for patients with CRC.

Methods

Data availability and materials

Messenger RNA (mRNA) expression data of patients with CRC from TCGA were downloaded from the Genomic Data Commons (GDC) Data Portal (https://portal.gdc.cancer.gov/). Patients with RNA-sequencing (RNA-seq) data and complete follow-up information were included in this study. Three GEO datasets, GSE39582 (N=557), GSE33113 (N=90) and GSE17538 (N=200), produced by the Affymetrix HG-U133 plus 2.0 platform were included. CEL files were downloaded and normalized for GEO microarray data using robust multichip average method. The identified TCGA and GEO datasets were combined into a single dataset and the ComBat (combating batch effects when combining batches of gene expression

microarray data) method was used to remove the batch effects among the TCGA and GEO cohorts. This method was implemented in the SVA R package. Other three datasets [GSE32323 (17 paired normal and cancer tissues), GSE110225 (17 paired normal and cancer tissues), GSE44076 (98 paired normal and cancer tissues)] from the GEO database were identified to compare the *TMPRSS2* expression between normal and tumor tissues. The basic clinicopathological features of the identified patients from the GEO and TCGA databases are presented in the supplementary table (available at https://cdn.amegroups.cn/static/public/jgo-23-641-1.docx). This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Functional annotation of TMPRSS2-related genes

To identify *TMPRSS2*-associated genes, the linear models for microarray data (LIMMA) method was used to generate the differentially expressed genes (DEGs) between low and high *TMPRSS2* expression groups. The threshold for significance was set as an adjusted P value <0.05 and fold change >1.5. Functional annotation of Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) biological process enrichment was further performed in the R package "clusterProfiler" (R version 3.4.2, The R Foundation for Statistical Computing; https://www.r-project.org/).

Identification of microRNA (miRNA) and transcriptional factors targeting TMRPSS2

miRWalk (http://mirwalk.umm.uni-heidelberg.de/) was used to predict the potential miRNAs targeting *TMPRSS2*. Moreover, the University of California, Santa Cruz (USCS) Genome browser (https://genome.ucsc.edu/) integrating the JASPAR database was used to determine the potential transcriptional factors regulating *TMPRSS2*, identify the modulating networks for *TMPRSS2*, and provide intervention alternatives.

Characterization of immune cell infiltration and evaluation of therapeutic response to immune checkpoint blockade

To quantify the relative abundance of immune cell infiltration and characterize the features of the CRC tumor microenvironment (TME) based on different *TMPRSS2* expression levels, the CIBERSORT algorithm was implemented, and the calculated enrichment score of each cell was used to represent their abundance level in the TME. CIBERSORT is a deconvolution method that characterizes the cell composition of complex tissue from their gene expression profiles. It employs linear support vector regression, a machine learning approach, to deconvolute a mixture of gene expression. To evaluate the therapeutic response to immune checkpoint blockade based on *TMPRSS2* expression, two cohorts of anti-programmed cell death protein 1 (*PD-1*) (21) and anti-cytotoxic T-lymphocyte-associated antigen 4 (*CTLA-4*) (22) were selected to analyze and compare the response rate between patients with low and high *TMPRSS2* expression.

Statistical analyses

All statistical analyses were performed using R software. Data are presented as the mean ± standard deviation. To compare differences between groups, the Wilcoxon rank-sum test was used for data with a skewed distribution, while the Student's *t*-test was used for data with a normal distribution. The X-tile program determined the best cutoff value to separate patients into low and high *TMPRSS2* expression groups. The Kaplan-Meier method was used to plot survival curves, and the log-rank test was used to analyze significance. A P value <0.05 was considered significant.

Results

TMPRSS2 was downregulated in CRC tissues and was correlated with superior clinical outcomes

We examined TCGA and three datasets (GSE32323, GSE110225, GSE44076) from GEO database to compare the expression difference between cancer tissue and adjacent normal tissue. The results showed that *TMPRSS2* was significantly downregulated in cancer tissues among the four cohorts (*Figure 1A-1D*). The comparison of *TMPRSS2* expression between normal and cancer tissues among other solid tumors is shown in Figure S1. To further test the prognostic value of *TMPRSS2* in patients with CRC, a total of 1,387 patients from TCGA and GEO databases were included. Based on the best cutoff value, patients with a low expression of *TMPRSS2* showed notably worse survival than did patients with a high expression in TCGA (P=0.015; *Figure 1E*), GSE39582 (P=0.016; *Figure 1F*), GSE33113

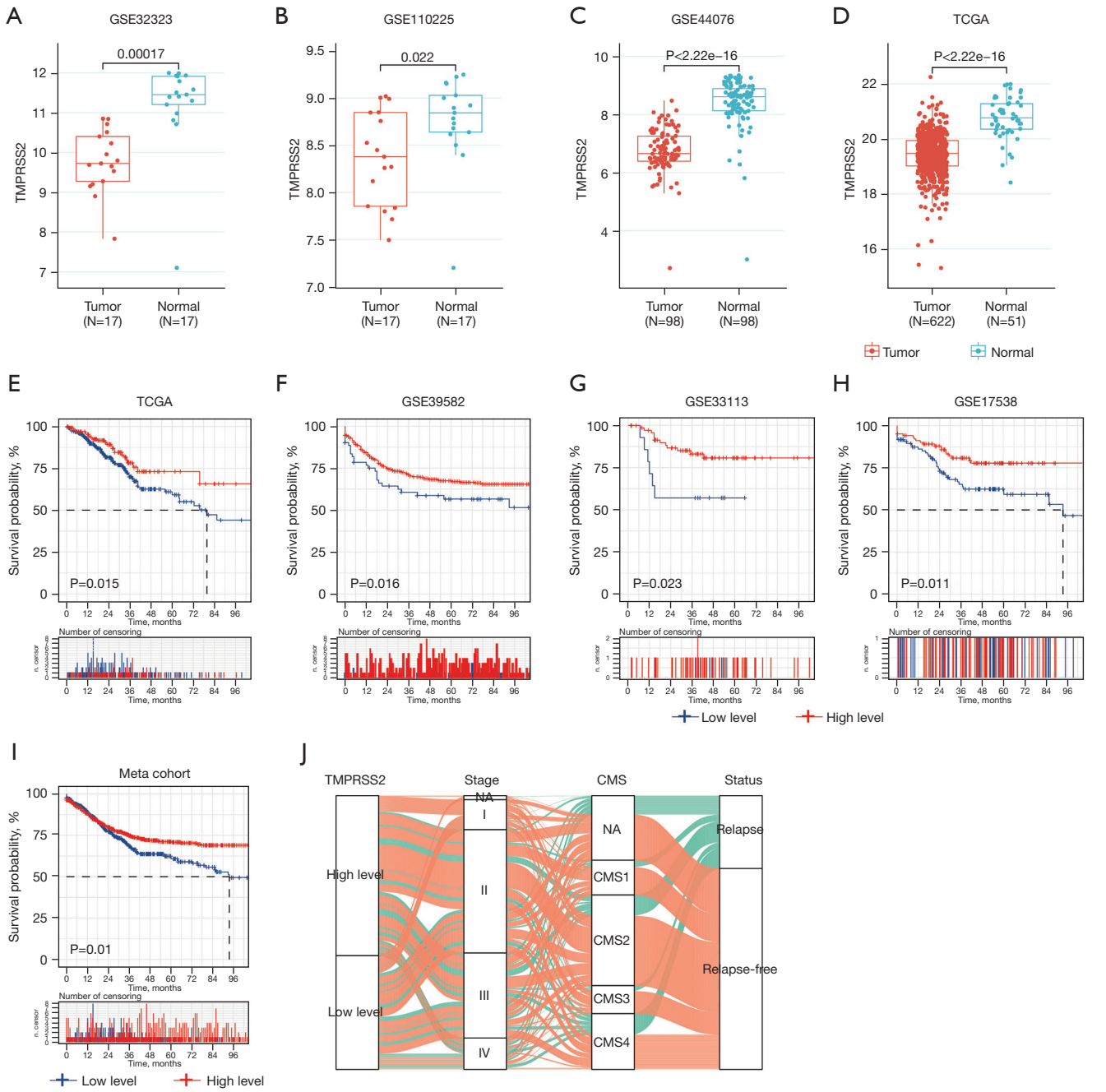

Figure 1 Expression difference of *TMPRSS2* between normal and cancerous tissues in the GSE32323 (A), GSE110225 (B), GSE44076 (C), and TCGA (D) datasets. The Kaplan-Meier survival curves comparing low and high expression of *TMPRSS2* in TCGA (E), GSE39582 (F), GSE33113 (G), GSE17538 (H), and the combined cohort (I). Correlation plot between *TMPRSS2* expression and clinical factors (J). TMPRSS2, transmembrane serine protease 2; TCGA, The Cancer Genome Atlas; NA, not applicable; CMS, consensus molecular subtypes.

(P=0.023; *Figure 1G*), GSE17538 (P=0.011; *Figure 1H*), and the combined (*Figure 1I*, P=0.01) cohorts. Further correlation analysis showed that low *TMPRSS2* expression was significantly associated with advanced stage (*Figure 1J*). These results revealed that *TMPRSS2* is a promising prognostic factor in CRC.

Identification of TMPRSS2-associated biological signaling pathways

To identify *TMPRSS2* expression-associated genes, we combined TCGA and GEO cohorts and performed differential expression analysis between *TMPRSS2* low and high expression groups. A total of 336 genes with fold change >1.5 and false-discovery rate (FDR) <0.05 were identified (the supplementary table is available at https://cdn.amegroups.cn/static/public/jgo-23-641-2.docx), with the top 10 differentially expressed genes being listed in *Figure 2A*. Next, we conducted function enrichment analysis based on the upregulated and downregulated genes in the *TMPRSS2* low-expression group. The GO and KEGG enrichment analysis indicated that a low expression of *TMPRSS2* was significantly associated with tumor metastasis-related biological processes or pathways, including extracellular matrix (ECM) organization, extracellular structure organization, focal adhesion, and ECM-receptor interaction (*Figure 2B-2E*); meanwhile, a high expression of *TMPRSS2* was associated with energy metabolism-related biological processes or pathways, including generation of precursor metabolites and energy, small-molecule catabolic process, fatty acid catabolic process, and citrate cycle (*Figure 2F-2I*). In summary, low *TMPRSS2* expression was notably associated with tumor progression-related pathways.

Prediction of miRNAs and transcriptional factors targeting TMPRSS2

To identify the potential regulators of *TMPRSS2* expression, we attempted to explore the potential miRNAs and transcriptional factors targeting *TMPRSS2*. First, we used the miRWalk tool to search for miRNAs that could putatively target *TMPRSS2* (the supplementary table is available at https://cdn.amegroups.cn/static/public/jgo-23-641-3.docx). Based on TCGA database, we examined six of the identified miRNAs (miR_379_5p, let_7i_5p, miR_513b_3p, miR_140_5p, miR_371a_5p, miR_708_5p) whose expression levels were inversely correlated with *TMPRSS2* expression (*Figure 3A*) and were significantly upregulated in cancer tissues (*Figure 3B*). Next, we used the UCSC genome database integrating JASPAR to predict the possible transcriptional factors targeting *TMPRSS2*. A list of transcriptional factors targeting *TMPRSS2* with the setting the track score set to >600 is provided in *Figure 3C*.

Characterization of the infiltrating immune cells in the TME among the different TMPRSS2 expression groups

To more precisely characterize the degree of immune infiltration between the *TMPRSS2* low and high expression groups, we performed CIBERSORT analysis; the resultant proportions of 22 kinds of immune cell components in CRC tissues are visualized in *Figure 4A*. A heatmap was then used to plot the differences in the immune cell infiltrations between the *TMPRSS2* low and high-expression groups (*Figure 4B*). Compared with those in the high *TMPRSS2* expression group, the infiltration levels of activated natural killer (NK) cells, resting NK cells, $CD8^+$ T cells, and M0, M1, and M2 macrophages were significantly higher in the low *TMPRSS2* expression group (*Figure 4C*). In summary, low *TMPRSS2* expression was significantly associated with immune-activating cell infiltration.

Potential of TMPRSS2 expression as an indicator of immunotherapy response

Based on the differences in immune cell infiltration between patients with low and high *TMPRSS2* expression, we hypothesized that *TMPRSS2* might be indicative of immunotherapy response. Tumor purity, heterogeneity, and immune checkpoint inhibitor (ICI) gene expression are established markers for immunotherapy response. We then explored the differences in tumor purity, heterogeneity, and ICI gene expression between the low and high *TMPRSS2* expression groups. We found that immune (*Figure 5A*), stromal (*Figure 5B*), and estimate scores (*Figure 5C*) were significantly higher in the low *TMPRSS2* expression group than in the high *TMPRSS2* expression group. Subsequently, the mutant-allele tumor heterogeneity (MATH) score was used to evaluate tumor heterogeneity between the low and high expression groups. The results showed that patients with low *TMPRSS2* expression had significantly higher MATH scores than those with high *TMPRSS2* expression (*Figure 5D*). We next compared the expression level of ICI genes [programmed death-ligand 1 (*PD-L1*) and *CTLA-4*] between patients with low and high *TMPRSS2* expression. It was found that patients with low *TMPRSS2* expression tended to show higher *PD-L1* and *CTLA-4* expression (*Figure 5E,5F*). Consensus molecular subtypes (CMS) are important molecular classification in CRC and is also an indicator for prognosis and immunotherapy response evaluation (23). We thus compared *TMPRSS2* expression differences among CMS types and found that the CMS4

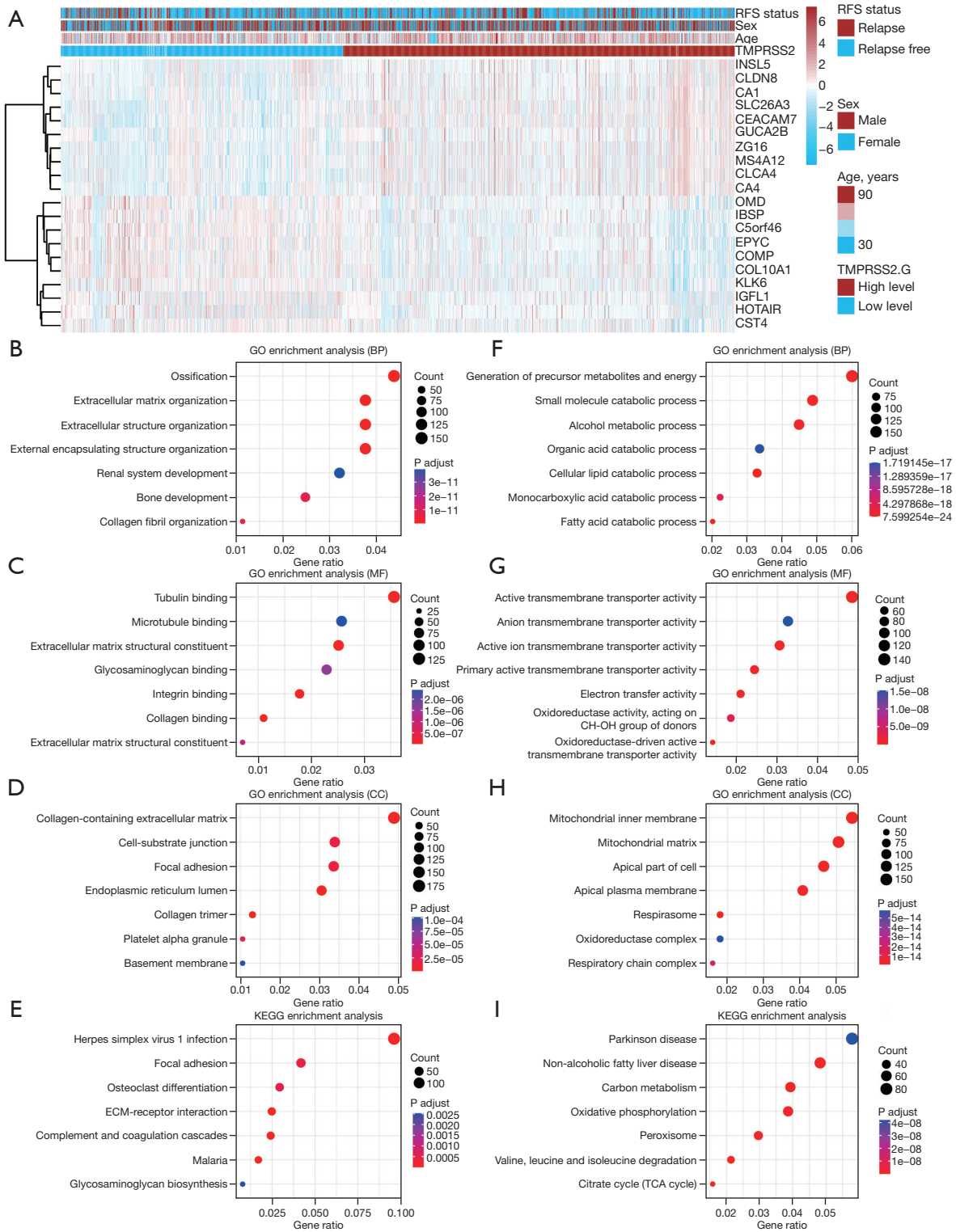

Figure 2 Heatmap of differentially expressed genes between the low and high *TMPRSS2* expression groups (A). Functional annotation of GO BP (B), GO MF (C), GO CC (D), and KEGG (E) based on the upregulated genes in the low *TMPRSS2* expression group. Functional annotation of GO biological process (F), GO molecular function (G), GO cellular component (H), and KEGG (I) based on the upregulated genes in the high *TMPRSS2* expression group. TMPRSS2, transmembrane serine protease 2; RFS, relapse free survival; KEGG, Kyoto Encyclopedia of Genes and Genomes; GO, Gene Ontology; BP, biological process; MF, molecular function; CC, cellular component.

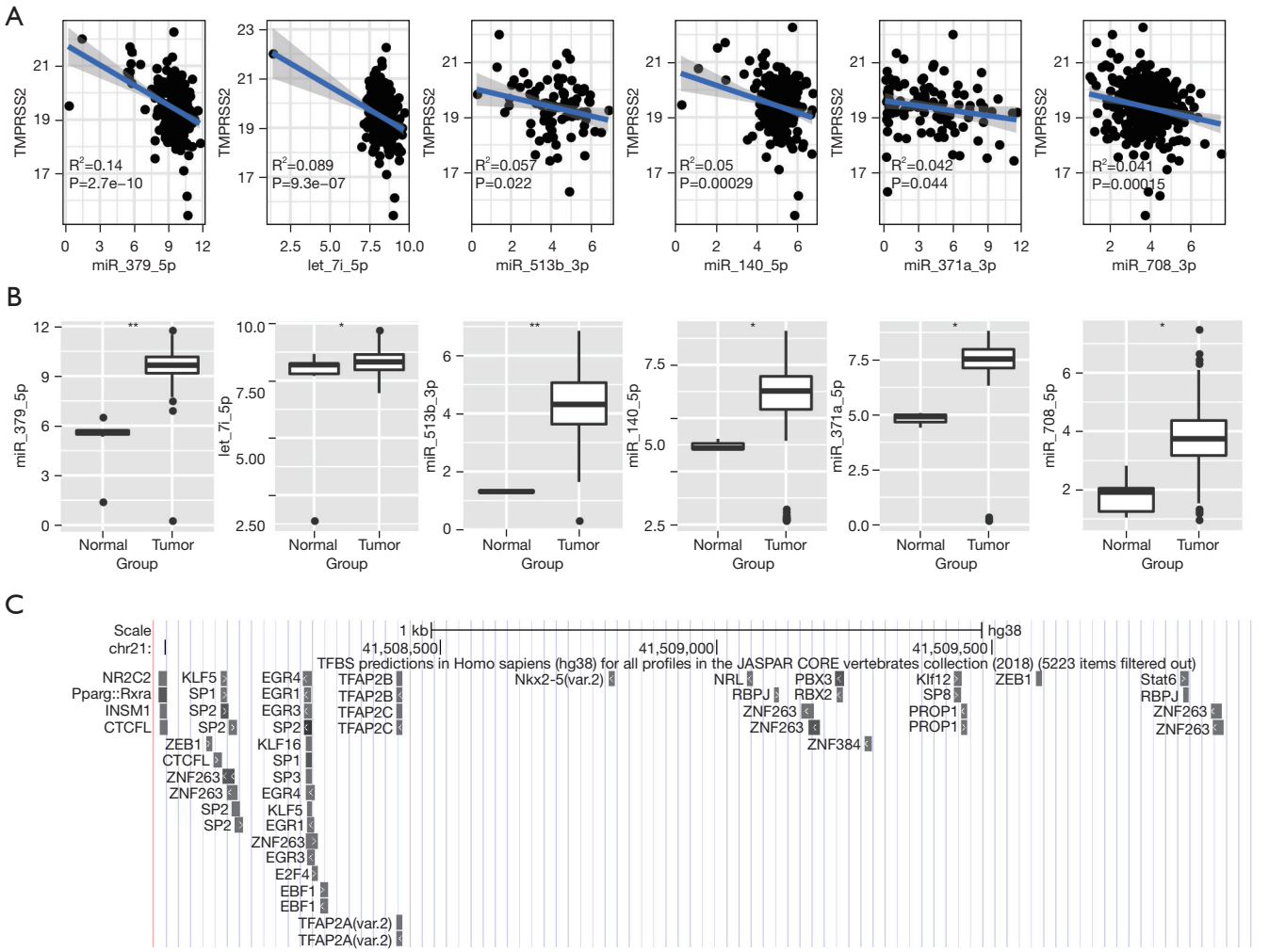

Figure 3 Correlation plot between indicated miRNAs and *TMPRSS2* in TCGA CRC dataset (A). Box-plot analysis representing the expression levels of miRNAs in normal and cancer tissues from patients with CRC in TCGA (B). Screenshot of the identified transcriptional factors with a track core >600 in the UCSC Genome Browser integrating the JASPAR database (C). *, P<0.05; **, P<0.01. TMPRSS2, transmembrane serine protease 2; miRNAs, microRNAs; TCGA, The Cancer Genome Atlas; CRC, colorectal cancer; UCSC, University of California Santa Cruz.

type was associated with the worst prognosis, while the immune-activated CMS1 type was associated with relatively low *TMPRSS2* expression (*Figure 5G*).

PD-1 and *CTLA-4* inhibitors are two representative immunotherapies. We next investigated whether *TMPRSS2* expression could be used to predict effective response to these immune checkpoint blockade therapies. In both the anti-PD-1 cohort (21) and anti-CTLA-4 cohort (22), patients with low *TMPRSS2* expression displayed more effective clinical response to anti-PD-1or CTLA-4 immunotherapies as compared to those patients with high *TMPRSS2* expression (response rate of anti-PD-1 cohort: 38% *vs.* 22%, *Figure 5H*; response rate of anti-CTLA-4 cohort: 57% *vs.* 32%, *Figure 5I*). Overall, these findings suggest that *TMPRSS2* could be a reliable biomarker for predicting immunotherapy response.

Discussion

SARS-CoV-2 is a newly identified coronavirus capable of infecting humans and rapidly inducing acute lung failure and multi-organ damage. *TMPRSS2* is one of the most

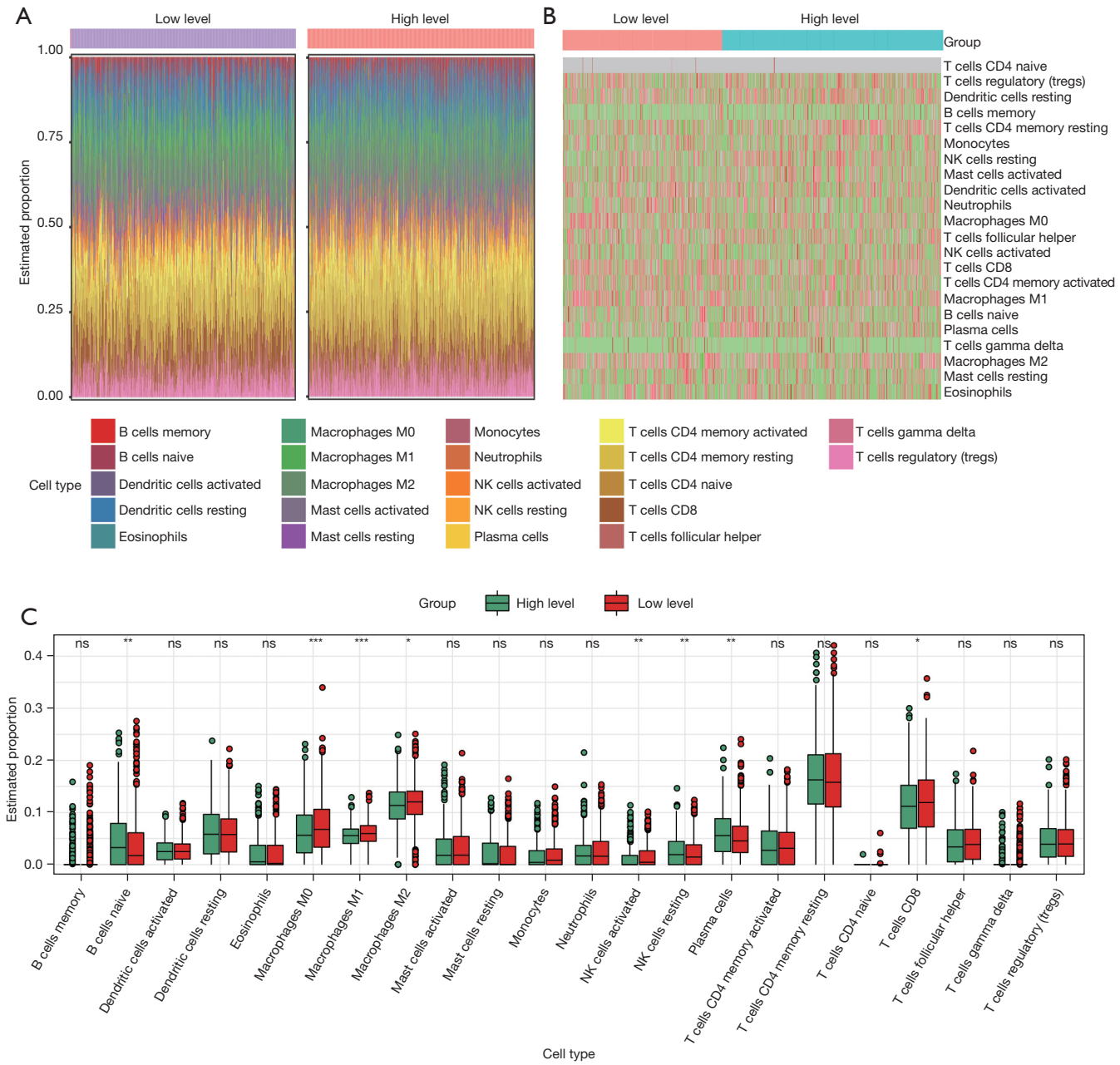

Figure 4 The proportions of different immune cell components in CRC samples with low and high *TMPRSS2* expression (A). Heatmap for the infiltration levels of immune cells in the low and high *TMPRSS2* expression groups (B). Box plots for the differences in infiltration levels of immune cells between the low and high *TMPRSS2* expression groups (C). ns, no significant; *, P<0.05; **, P<0.01; ***, P<0.001. NK, natural killer; TMPRSS2, transmembrane serine protease 2; CRC, colorectal cancer.

important factors mediating the susceptibility to SARS-CoV-2 and infection enhancement. Thus, exploring the susceptibility to coronavirus and the physiological functions of *TMPRSS2* in humans is essential to the ongoing global drive to manage this pandemic. Previous reports have suggested that patients with cancer are more susceptible to infections, and several bioinformatics analyses have shown that cancer tissues have higher *TMPRSS2* expression levels than matched healthy tissues in lung and stomach cancer (24,25). As previously

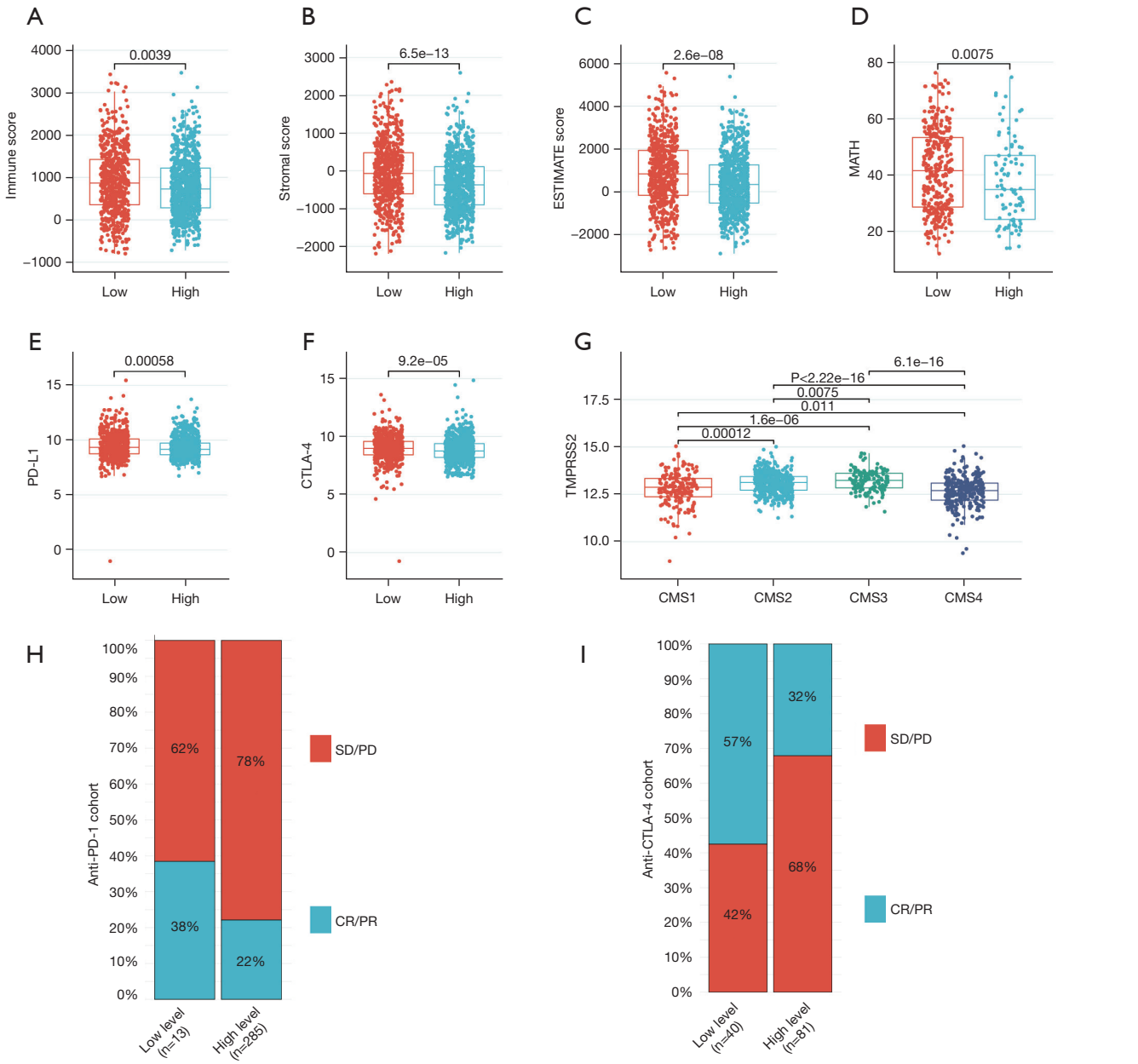

Figure 5 Box plots for the differences of immune score (A), stromal score (B), ESTIMATE score (C), MATH score (D), PD-L1 expression (E), and CTLA-4 expression (F) between the low and high *TMPRSS2* expression groups. Comparison of *TMPRSS2* expression among patients with different CMS types (G). Distribution of response rate to anti-PD-1 (H) and anti-CTLA-4 (I) in patients with low and high *TMPRSS2* expression. TMPRSS2, transmembrane serine protease 2; MATH, mutant-allele tumor heterogeneity; PD-L1, programmed death-ligand 1; CTLA-4, cytotoxic T-lymphocyte-associated protein 4; CMS, consensus molecular subtypes; PD-1, programmed death 1; CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease.

reported, patients with COVID-19 usually exhibit multi-organ damage, especially aerodigestive disease (26). We can speculate that a high expression of *TMPRSS2* in these tissues facilitates the susceptibility to SARS-CoV-2. However, studies concerning the implications of *TMPRSS2* expression in patients with CRC remain conflicting. Wang *et al.* showed that *TMPRSS2* was highly expressed in CRC which suggests severe symptoms of SARS-

CoV-2 Infection (27). We thus applied multiomics tools to explore the role of *TMPRSS2* expression in CRC and evaluated the correlation between *TMPRSS2* expression and immunotherapy response to glean new insights into the impact of COVID-19 infection on CRC.

To our knowledge, this is the first study to analyze the difference in *TMPRSS2* expression between normal and cancer tissues and its prognostic value based on multiple datasets. We found that *TMPRSS2* was highly expressed in normal colorectal tissues. In addition, a list of miRNAs and transcriptional factors that could putatively target *TMPRSS2* were identified. Notably, *TMPRSS2* expression was associated with immune cell infiltration, and a low expression of *TMPRSS2* was significantly correlated with higher immunotherapy response.

TMPRSS2 mediates the entry of SARS-CoV-2 into host cells. The low level of *TMPRSS2* in CRC cancerous tissue was confirmed in this study, which indicates that patients with CRC may be more resistant to SARS-CoV-2 infection. CRC is a type of immunogenic cancer. It was found that the spontaneous anticancer immune response may help to increase survival duration, while the immune escape may decrease survival (28). First lymphocytes will eliminate tumor cells while some cells are resilient and they survive and reach in equilibrium phase and thus escaping immune surveillance (29). Increased *PD-1* expression was found to be able to cause tumor infiltrating lymphocytes to be terminally exhausted (29). Recently, the application of several ICIs including anti-PD-1 and anti-CTLA-4 has been widely adopted in CRC (28). The aim of immunotherapy is to activate the immunosuppressive networks in the TME. Several immune-related markers have been used for predicting the response to immunotherapy (30). Our study revealed that a low expression of *TMPRSS2* was related to the infiltration of resting NK cells, activated NK cells, plasma cells, $CD8^+$ T cells, and M0, M1, and M2 macrophages, indicating the possible relation between *TMPRSS2* expression and immunotherapy response. Tumor purity, heterogeneity, and ICI gene expression level are established indicators for immunotherapy response. We found that patients with a low expression of *TMPRSS2* tended to have higher tumor heterogeneity and ICI gene expression level, which correlated with a higher immunotherapy response rate. As expected, a low expression of *TMPRSS2* was found to be associated with a higher response rate in both the anti-PD-1 and anti-CTLA-4 cohorts.

Although this is the first study to demonstrate a correlation between *TMPRSS2* expression and the prognosis, molecular features, and immunotherapy response in patients with CRC, several limitations should be noted. First, all the data were obtained from public databases, and further prospective validation is needed. Second, samples of patients with CRC with SARS-CoV-2 infection should be tested to confirm the conclusions of this study. Finally, the mechanisms underlying the correlations reported in this study remain unclear and should be further investigated.

Conclusions

In conclusion, CRC tumor cells may be more resistant to SARS-CoV-2 infection due to decreased *TMPRSS2* expression. In addition, the expression levels of this gene may represent a novel biomarker for the prognosis and immunotherapy response prediction of patients with CRC.

Acknowledgments

We would like to acknowledge the editors and reviewers for their helpful comments on this paper.
Funding: This work was supported by the National Natural Science Foundation of China (No. 82102850; No. 82002416), the China Postdoctoral Science Foundation (No. 2020M671399), the Natural Science Foundation of Jiangsu Province (No. BK20201089), and the Jiangsu Province Postdoctoral Fund (No. 247755).

Footnote

Reporting Checklist: The authors have completed the REMARK reporting checklist. Available at https://jgo.amegroups.com/article/view/10.21037/jgo-23-641/rc

Peer Review File: Available at https://jgo.amegroups.com/article/view/10.21037/jgo-23-641/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jgo.amegroups.com/article/view/10.21037/jgo-23-641/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. This study was

conducted in accordance with the Declaration of Helsinki (as revised in 2013).

References

1. Tabibzadeh A, Esghaei M, Soltani S, et al. Evolutionary study of COVID-19, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as an emerging coronavirus: Phylogenetic analysis and literature review. Vet Med Sci 2021;7:559-71.
2. Zhao J, Zhao S, Ou J, et al. COVID-19: Coronavirus Vaccine Development Updates. Front Immunol 2020;11:602256.
3. Senapati S, Banerjee P, Bhagavatula S, et al. Contributions of human ACE2 and TMPRSS2 in determining host-pathogen interaction of COVID-19. J Genet 2021;100:12.
4. Fricke-Galindo I, Falfán-Valencia R. Genetics Insight for COVID-19 Susceptibility and Severity: A Review. Front Immunol 2021;12:622176.
5. Dong M, Zhang J, Ma X, et al. ACE2, TMPRSS2 distribution and extrapulmonary organ injury in patients with COVID-19. Biomed Pharmacother 2020;131:110678.
6. Siegel RL, Miller KD, Fuchs HE, et al. Cancer Statistics, 2021. CA Cancer J Clin 2021;71:7-33.
7. Wrobel P, Ahmed S. Current status of immunotherapy in metastatic colorectal cancer. Int J Colorectal Dis 2019;34:13-25.
8. Biller LH, Schrag D. Diagnosis and Treatment of Metastatic Colorectal Cancer: A Review. JAMA 2021;325:669-85.
9. Benson AB, Venook AP, Al-Hawary MM, et al. Colon Cancer, Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 2021;19:329-59.
10. Veglia F, Sanseviero E, Gabrilovich DI. Myeloid-derived suppressor cells in the era of increasing myeloid cell diversity. Nat Rev Immunol 2021;21:485-98.
11. Sica A, Massarotti M. Myeloid suppressor cells in cancer and autoimmunity. J Autoimmun 2017;85:117-25.
12. Ilikci Sagkan R, Akin-Bali DF. Structural variations and expression profiles of the SARS-CoV-2 host invasion genes in lung cancer. J Med Virol 2020;92:2637-47.
13. Katopodis P, Anikin V, Randeva HS, et al. Pan-cancer analysis of transmembrane protease serine 2 and cathepsin L that mediate cellular SARS-CoV-2 infection leading to COVID-19. Int J Oncol 2020;57:533-9.
14. Liang W, Guan W, Chen R, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. Lancet Oncol 2020;21:335-7.
15. Sacconi A, Donzelli S, Pulito C, et al. TMPRSS2, a SARS-CoV-2 internalization protease is downregulated in head and neck cancer patients. J Exp Clin Cancer Res 2020;39:200.
16. Mihalopoulos M, Dogra N, Mohamed N, et al. COVID-19 and Kidney Disease: Molecular Determinants and Clinical Implications in Renal Cancer. Eur Urol Focus 2020;6:1086-96.
17. Lam DK, Dang D, Flynn AN, et al. TMPRSS2, a novel membrane-anchored mediator in cancer pain. Pain 2015;156:923-30.
18. Ko CJ, Huang CC, Lin HY, et al. Androgen-Induced TMPRSS2 Activates Matriptase and Promotes Extracellular Matrix Degradation, Prostate Cancer Cell Invasion, Tumor Growth, and Metastasis. Cancer Res 2015;75:2949-60.
19. Huang X, He C, Hua X, et al. Bioinformatic Analysis of Correlation between Immune Infiltration and COVID-19 in Cancer Patients. Int J Biol Sci 2020;16:2464-76.
20. He C, Hua X, Sun S, et al. Integrated Bioinformatic Analysis of SARS-CoV-2 Infection Related Genes ACE2, BSG and TMPRSS2 in Aerodigestive Cancers. J Inflamm Res 2021;14:791-802.
21. Liu D, Schilling B, Liu D, et al. Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. Nat Med 2019;25:1916-27.
22. Miao D, Margolis CA, Gao W, et al. Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. Science 2018;359:801-6.
23. López-Camacho E, Prado-Vázquez G, Martínez-Pérez D, et al. A Novel Molecular Analysis Approach in Colorectal Cancer Suggests New Treatment Opportunities. Cancers (Basel) 2023;15:1104.
24. Xiao X, Shan H, Niu Y, et al. TMPRSS2 Serves as a Prognostic Biomarker and Correlated With Immune Infiltrates in Breast Invasive Cancer and Lung

Adenocarcinoma. Front Mol Biosci 2022;9:647826.
25. Hoang T, Nguyen TQ, Tran TTA. Genetic Susceptibility of ACE2 and TMPRSS2 in Six Common Cancers and Possible Impacts on COVID-19. Cancer Res Treat 2021;53:650-6.
26. Walczak P, Janowski M. The COVID-19 Menace. Glob Chall 2021;5:2100004.
27. Wang H, Yang J. Colorectal Cancer that Highly Express Both ACE2 and TMPRSS2, Suggesting Severe Symptoms to SARS-CoV-2 Infection. Pathol Oncol Res 2021;27:612969.
28. Camidge DR, Doebele RC, Kerr KM. Comparing and contrasting predictive biomarkers for immunotherapy and targeted therapy of NSCLC. Nat Rev Clin Oncol 2019;16:341-55.
29. Verma NK, Wong BHS, Poh ZS, et al. Obstacles for T-lymphocytes in the tumour microenvironment: Therapeutic challenges, advances and opportunities beyond immune checkpoint. EBioMedicine 2022;83:104216.
30. Doroshow DB, Sanmamed MF, Hastings K, et al. Immunotherapy in Non-Small Cell Lung Cancer: Facts and Hopes. Clin Cancer Res 2019;25:4592-602.

Cite this article as: Zhu C, Wu J, Shelat VG, Sayagués JM, Yamamoto S, Yang L, He X. Transmembrane serine protease 2, a SARS-CoV-2 internalization protease, correlates with clinical outcome, molecular features, and immunotherapy response in colorectal cancer. J Gastrointest Oncol 2023;14(5):2146-2157. doi: 10.21037/jgo-23-641

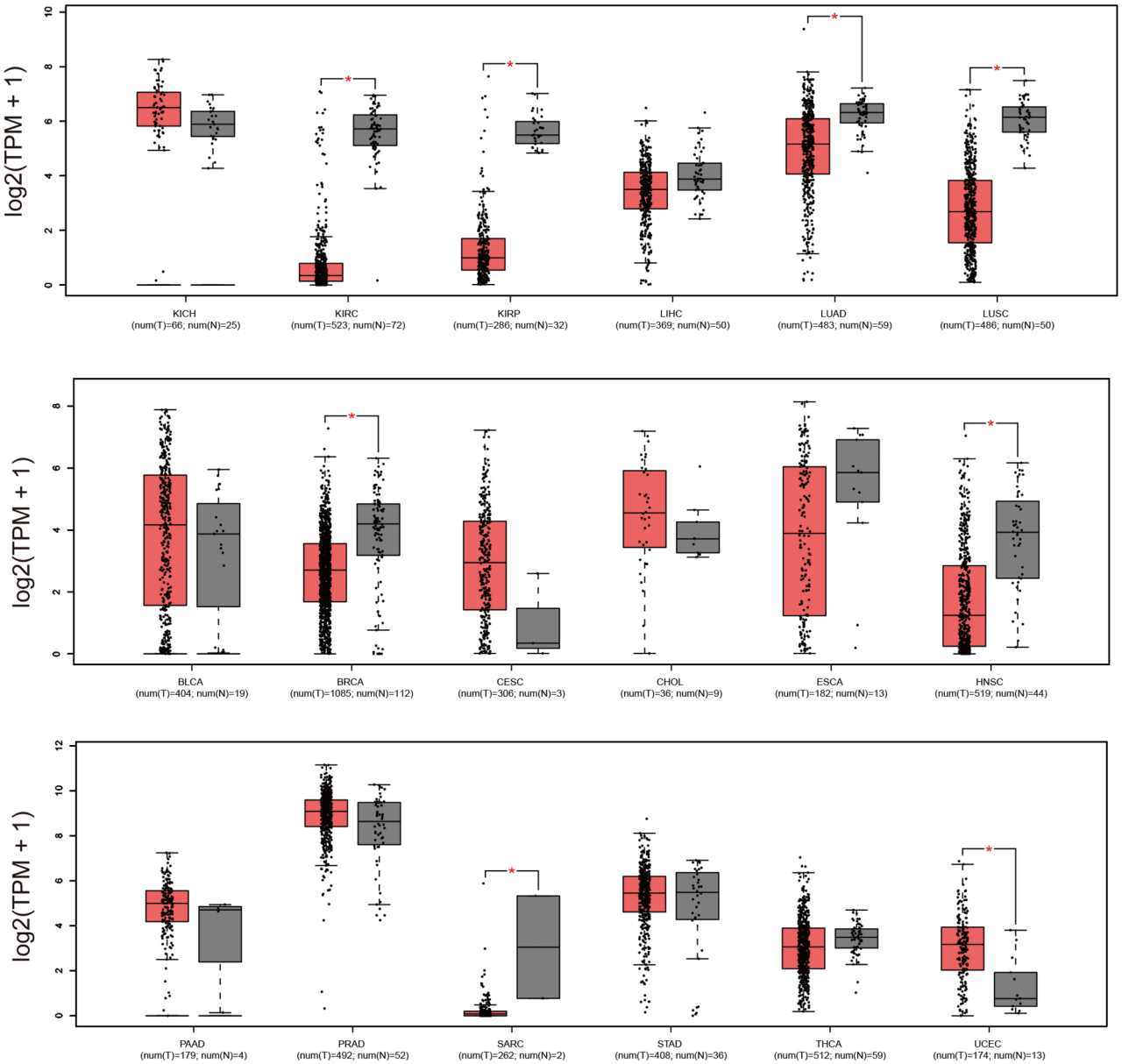

Figure S1 Box plot of TMPRSS2 expression between normal and cancerous tissues among solid tumors in TCGA database. *, P <0.05.

https://dx.doi.org/10.21037/jgo-23-641