

RESEARCH

Open Access



# Characterizing and forecasting neoantigens-resulting from MUC mutations in COAD

Min Chen<sup>1\*</sup>, Xin Zhang<sup>2</sup>, Zihe Ming<sup>3</sup>, Lingyu<sup>4†</sup>, Xiaorong Feng<sup>5</sup>, Zhenguo Han<sup>6</sup> and Han-Xiang An<sup>1,7\*</sup> 

## Abstract

**Background** The treatment for colon adenocarcinoma (COAD) faces challenges in terms of immunotherapy effectiveness due to multiple factors. Because of the high tumor specificity and immunogenicity, neoantigen has been considered a pivotal target for cancer immunotherapy. Therefore, this study aims to identify and predict the potential tumor antigens of MUC somatic mutations (MUCmut) in COAD.

**Methods** Three databases of TCGA, TIMER2.0, and cBioPortal were used for a detailed evaluation of the association between MUCmut and multi-factors like tumor mutation burden (TMB), microsatellite instability (MSI), prognosis, and the tumor microenvironment within the context of total 2242 COAD patients. Next, TSNAdb and the differential agretopicity index (DAI) were utilized to predict high-confidence neopeptides for MUCmut based on 531 COAD patients' genomic information. DAI was calculated by subtraction of its predicted HLA binding affinity of the MUCmut peptide from the corresponding wild-type peptide.

**Results** The top six mutation frequencies (14 to 2.9%) were from MUC16, MUC17, MUC5B, MUC2, MUC4 and MUC6. COAD patients with MUC16 and MUC4 mutations had longer DFS and PFS. However, patients with MUC13 and MUC20 mutations had shorter OS. Patients with the mutation of MUC16, MUC5B, MUC2, MUC4, and MUC6 exhibited higher TMB and MSI. Moreover, these mutations from the MUC family were associated with the infiltration of diverse lymphocyte cells and the expression of immune checkpoint genes. Through TSNAdb 1.0/NetMHCpan v2.8, 452 single nucleotide variants (SNVs) of MUCmut peptides were identified. Moreover, through TSNAdb2.0/NetMHCpan v4.0, 57 SNVs, 1 Q-frame shift (TS), and 157 short insertions/deletions (INDELs) of MUCmut were identified. Finally, 10 high-confidence neopeptides of MUCmut were predicted by DAI.

**Conclusions** Together, our findings establish the immunogenicity and therapeutic potential of mutant MUC family-derived neoantigens. Through combining the tools of TSNAdb and DAI, a group of novel MUCmut neoantigens were identified as potential targets for immunotherapy.

**Keywords** MUC, Tumor neoantigens, Colon adenocarcinoma, Bioinformatics, Immunotherapy

<sup>†</sup>Min Chen, Xin Zhang and Zihe Ming have contributed equally to this work.

\*Correspondence:

Min Chen

mchenst@outlook.com

Han-Xiang An

anhaxiang@hotmail.com

Full list of author information is available at the end of the article



## Background

Colon adenocarcinoma (COAD), the most prevalent form of colorectal cancer (CRC), ranks as the third most commonly diagnosed malignant tumor with a high mortality rate worldwide [1, 2]. Immunotherapy for cancer, like immune-checkpoint inhibitors (ICIs) that boost the body's natural anti-tumor immune responses, has demonstrated potential in treating recurrent or metastatic cancer [3–5]. The metastatic CRC patients with deficient mismatch repair (dMMR) or high microsatellite instability (MSI-H) receive a clear clinical response [6, 7], while those with proficient mismatch repair (pMMR)/microsatellite instability-low (MSI-L) CRC did not benefit from immunotherapy [8]. During the past 10 years, neoantigens, a class of antigens derived from tumor-specific mutations, have been widely regarded as optimal targets for activating the patient's immune system to recognize and eliminate cancer cells such as the mutations of KRAS [9], HER2 [10], and PIK3CA [11].

It has emerged as the most prevalent approach for the precise elimination of tumors through immunotherapy. Prof. Rosenberg's research group reported the identification rate of neoantigen reactive T lymphocytes in gastrointestinal cancer was high (83%–90%) [12, 13] and also found that most epithelial cancers, generally considered to be non-immunogenic, do elicit *in vivo* immune reactions. Their studies offer a rationale for developing new immunotherapeutic treatments by targeting the unique tumor-associated mutations expressed in epithelial cancer. With the promising outcome from the neoantigen-based immunotherapies [14–19], researchers and clinicians are exploring the possibility of using tumor neoantigens as a potentially effective treatment to improve the treatment options for CRC patients.

In pan-cancer studies, a peptide-based neoantigen vaccine, iNeo-Vac-P01, was conducted on Chinese patients with solid tumors (NCT03662815) and showed a promising outcome in 30 patients, and 80% of peptides enhanced immune response [20, 21]. In a recent study (NCT03639714), the integration of neoantigen cancer vaccines GRT-C901 and GRT-R902 (produced by Gritstone Bio) with nivolumab and ipilimumab drugs was shown to significantly improve overall survival rates among patients with NSCLC, metastatic urothelial, gastroesophageal, and microsatellite stable colorectal cancers [21]. In a phase I/IIa clinical trial (NCT01461148) including 22 dMMR CRC patients treated with a united neoantigen-peptide containing three frameshift mutant genes (AIM2, HT001, TAF1B), two of the three assessable patients hold stable disease as the best overall response, and one heavily pretreated patient with bulky metastasis also presented as stable disease over 7 months [22]. Overall, the promising results of these results provided clinical

prospects in terms of neoantigens-based immunotherapy in solid cancer.

The identification of immunogenic neoantigens from numerous sources is a crucial step in the development of effective immunotherapies [23]. Neoantigens may now be thoroughly screened across the entire cancer spectrum thanks to the convergence of whole-exome sequencing (WES), RNA-seq, and proteomic data from TCGA [24]. Several studies showed a remarkable number of somatic mutations and tens of putative mutational neoantigens in COAD [25, 26], resulting in a vast total number of potential multiple neopeptides. To verify whether an antigen can be presented by the major histocompatibility complex (MHC), two general methods are used: (1) *in silico* computational predictions and (2) mass spectrum analyses. Several software programs are available for *in silico* computational predictions, including TSNAdb [27], NetMHCpan [28, 29], and IEDB [30]. The main principle of *in silico* computational predictions is to establish a predictive model through a database acquired by proteasome splicing, transporter associated with antigen processing (TAP) channel selectivity, and epitopes through which the MHC molecules recognize the peptide. Thus, ongoing attempts or strategies to improve neo-peptide immunogenicity prediction are critical for the selection of neoantigen targets for cancer.

The MUC family characterized by its heavily glycosylated transmembrane mucins has the same protein sequence in normal and tumor cells but different glycosylation profiles, thus the glycopeptide epitopes are expected to be tumor-specific. Studies showed that MUC1 neoantigens were exclusively recognized by glycoform-specific T cell receptors (TCR) [31] or presented on MHC class II molecules [32, 33]. Especially in recent years, MUC1 neoantigens has been proven as a tumor-associated antigen highly expressed on many adenocarcinomas (breast, colon, lung, kidney, ovary, etc.), and MUC1 neoantigens has been successfully utilized in immunotherapeutic approaches for the development of peptide-, carbohydrate-, DNA-, and dendritic cell (DC)-based vaccines [33], and clinical trials (NCT00415818 [34] and NCT00409188 [35]) in lung cancer. Currently, several Phase II clinical studies (NCT02134925 and NCT00773097) examined the efficacy of immunotherapies targeting tumor neoantigens against MUC1 in COAD. Thus, tumor-associated glycopeptide antigens of the MUC family may serve as ideal targets for the development of immunotherapeutic vaccines in COAD.

The human MUC family consists of at least 22 members: MUC1-MUC22 [36]. The aberrant expression of the MUC family has been reported to be a common feature of CRC. MUC1, MUC2, MUC4, and MUC6 are up-regulated in CRC, and their overexpression

has been associated with disease progression [37–39]. However, the demethylation and down-expression of MUC5AC are predictive biomarkers for MSI and prognosis in CRC [38, 40–42]. A recent report presented a finding from multi-omics clinical data, including transcriptomics RNA-sequencing (mRNA, lncRNA, miRNA), DNA methylation, and gene mutations in the TCGA-STAD cohort demonstrated that MUC16 could be used as a potential biomarker to predict the response of immunotherapy and chemotherapy in gastric cancer [43], implying that MUC may be a broad-spectrum therapeutic target. However, an integrated bioinformatics analysis of MUC mutations in COAD across multiple databases has not been performed to date. Herein, in this study, we obtained neoantigen candidates of MUC mutant genes for COAD by the combination of a comprehensive analysis of public data and set up a prediction model for initiating the selection of high-confidence neoantigens derived from MUC mutations. We provided a series of potential candidates for tumor neoantigen-based immunotherapy in COAD.

## Methods

### Data collection and genetic mutation analysis

We procured somatic mutation data from 172 COAD samples within The Cancer Genome Atlas (TCGA), accessible through the official website (<https://portal.gdc.cancer.gov/>) and from 2070 COAD samples hosted on the cBioPortal, available at (<https://www.cbioportal.org/>). Additionally, the prognostic analysis results of overall survival (OS), disease-free survival (DFS), progression-free survival (PFS), and disease-specific survival (DSS) between the 19 MUC genes mutation group and the non-mutation group in COAD cancer patients were also conducted using data from these websites.

### Tumor-associated antigens prediction for MUC family

The Tumor Immune Estimation Resource 2.0 (TIMER2.0, <http://timer.cistrome.org/>) web server, a comprehensive resource integrating multiple immune infiltration estimation algorithms [44] was used for the analysis of tumor immunogenicity. The TISIDB web server, a web portal integrated with multiple heterogeneous data types, was used for tumor and immune system interaction analysis. A two-sided Wilcoxon rank-sum was employed to compare the expression levels of genes related to tumor-infiltrating immune cells, immunoinhibitors, immunostimulators, MHC molecules, chemokines, receptors and antitumor immunity between MUC family-mutated and wild-type groups.

### Neoantigen prediction

TSNAdb [27] collects 7748 tumor samples from TCGA and the Cancer Immunome Atlas (TCIA). We obtained the tumor-specific neoantigen data of 531 TCGA COAD samples from TSNAdb (<http://biopharm.zju.edu.cn/tsnadb>). Compared with NetMHCpan v2.8 [28, 30], NetMHCpan v4.0 [45] is trained based on both binding affinity data and mass spectrometry data, thus adopting stricter criteria for binding prediction. The binding affinities of wild-type and mutant peptides of the MUC family with the HLA alleles of patients were first predicted by the NetMHCpan (v2.8 and v4.0), algorithm in COAD using TSNAdb v1.0. Next, TSNAdb v2.0 combines the predicted results of DeepHLApan, MHCflurry, and NetMHCpan v4.0 for the identification of higher confidence neoantigens derived from SNV, INDEL, and Fusion [28]. To calculate DAI, MHC-I affinity was predicted for wildtype (WTA) and mutant (MA) peptide pairs arising from the same mutation and differing by a single amino acid. DAI was defined as WTA-MA for each peptide pair.

### Statistical analysis

Enumeration data including gene mutations were compared using the chi-square test or Fisher's exact test, while measurement data like age, TMB, abundance of immune cells, and immune-associated gene expression were compared by Mann–Whitney U-test. Clinical outcomes were assessed using the Kaplan–Meier method and compared by the log-rank test. The TISIDB database was used to explore the correlations between MUC expression and immune- or molecular subtypes in COAD, in which a  $p$ -value  $< 0.05$  was set as the satisfying criteria. To evaluate the predictive impact of MUC mutations on immunotherapy, we employed the 'IMvigor210CoreBiologies' package (version 1.0.0) within the R software, which encompasses clinical data of ICIs in patients with COAD. The TIMER database was employed for a comprehensive examination of tumor-infiltrating immune cells. In this analysis, the somatic copy number alterations (SCNAs) and mutation modules were evaluated using a two-sided Wilcoxon rank-sum test. The SCNAs module, as defined by GISTIC 2.0, includes the following categories: deep deletion (−2), arm-level deletion (−1), diploid/normal (0), arm-level gain (1), and high amplification (2). The value of  $p \leq 0.05$  was considered statistically significant.

## Results

### Exploring MUC genes mutations in colon adenocarcinoma datasets

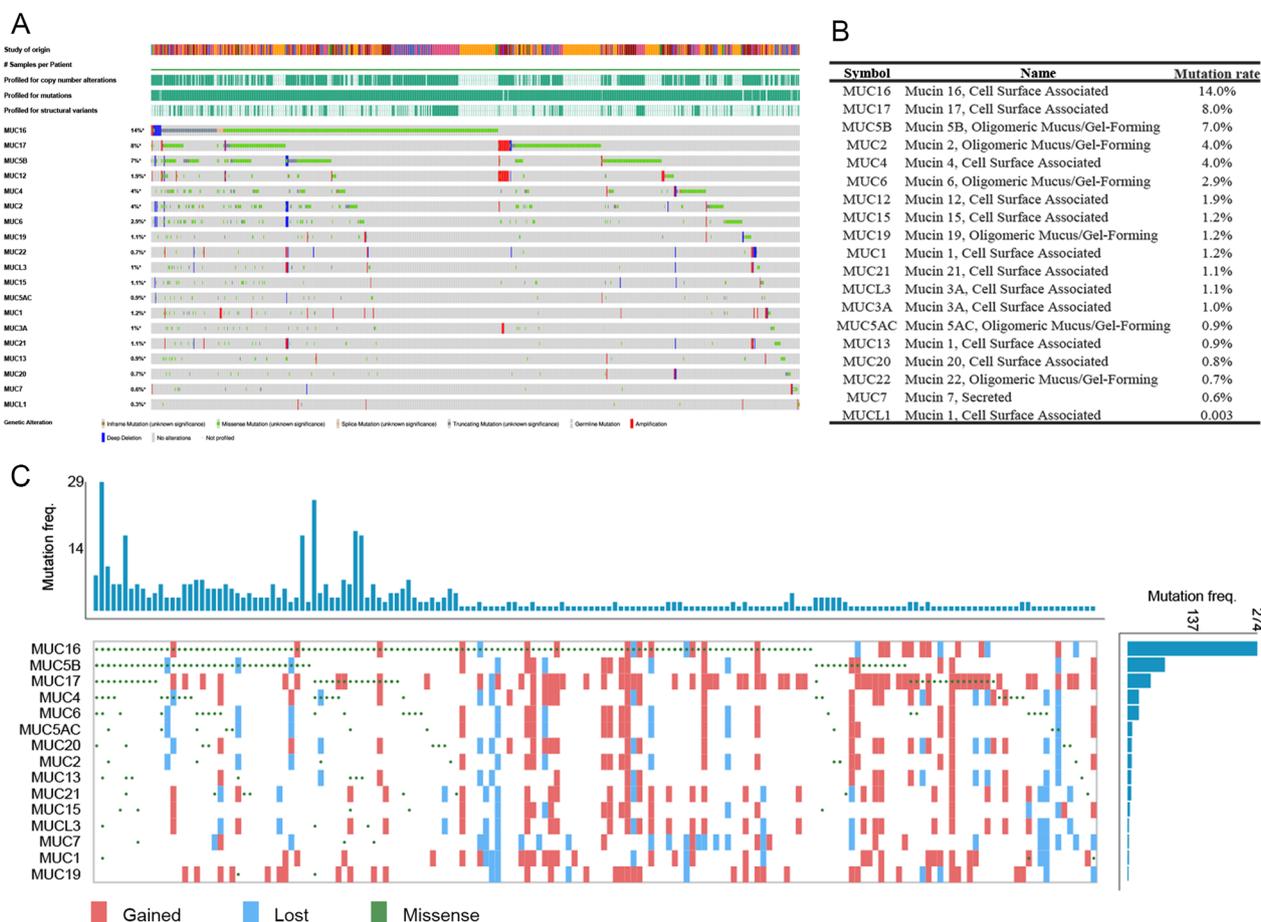
To identify potential tumor neoantigens in COAD, we determined the mutational frequencies of MUC genes

within the context of a total of 2242 COAD across two distinct datasets (n1=2070, n2=172). Firstly, we present the mutational frequencies of nineteen MUC genes as cataloged on the renowned platform, cBioPortal (Fig. 1A, n1=2070). The genes in which recurrent mutations were frequently observed were MUC16 (14%), MUC17 (8%), MUC5B (7%), MUC2 (4%), MUC4 (4%), and MUC6 (2.9%), respectively. Within this framework, we present an exhaustive tabular representation detailing the mutation rates of these aforementioned genes (Fig. 1B). Continuing our analysis, we also explored the mutation frequencies of the same MUC genes using data from the TCGA datasets (Fig. 1C, n2=172). This search yielded similar results, further validating our findings from the cBioPortal platform. Further delving into the results, we observed distinct mutation patterns across the MUC genes. While MUC16 showed the highest mutation frequency at 75.4%, this notably high rate could be influenced by the small sample size and the randomness

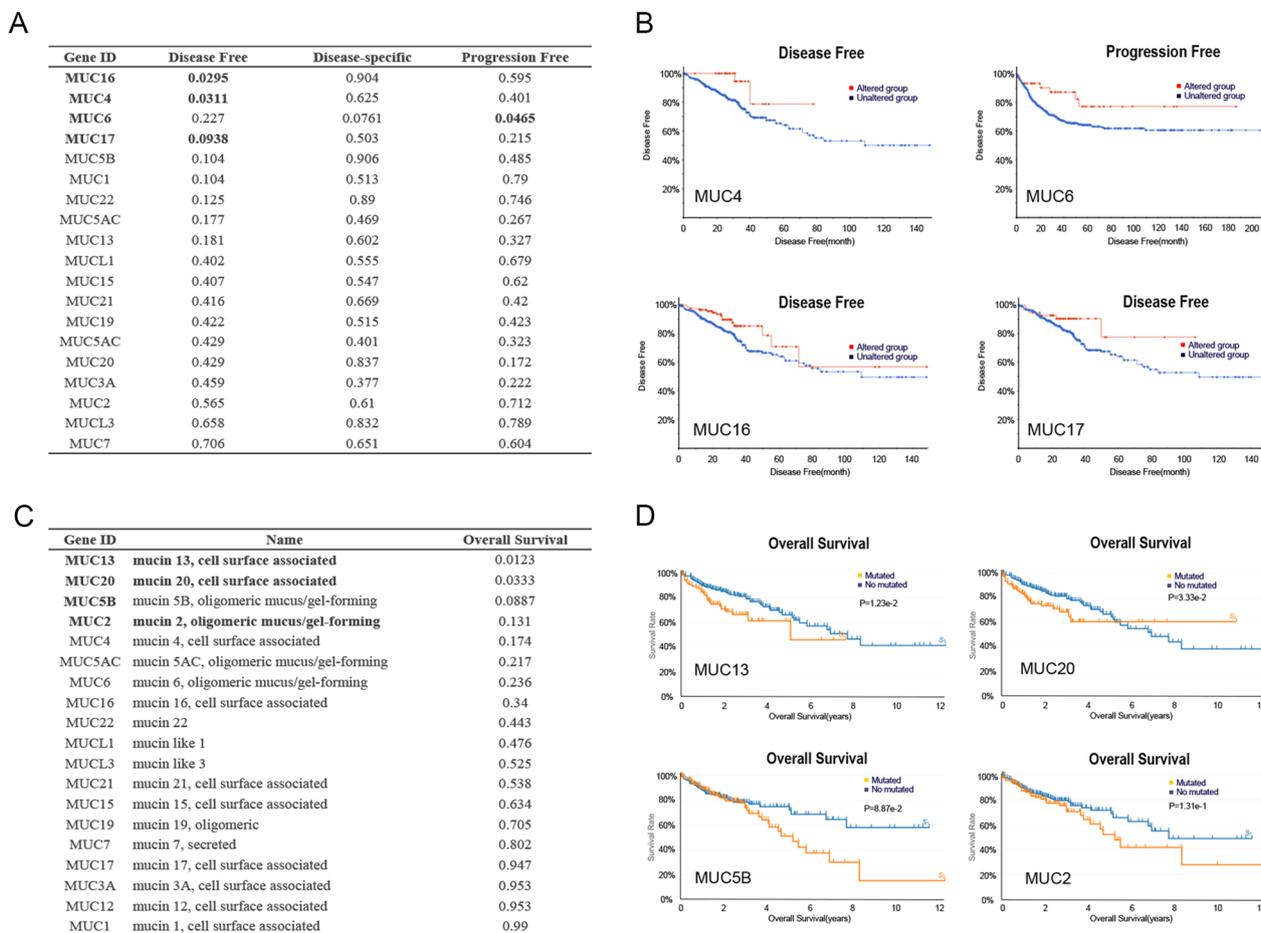
inherent in the data. This is followed by MUC5B at 22.8%, MUC17 at 13.5%, and MUC2 and MUC6 both at 8.7% (Additional file 6: Table S1). Through the analyses of these two databases, it becomes apparent that among the most prominent genes in terms of mutation frequency are MUC16, MUC17, MUC5B, MUC2, MUC4, and MUC6. These patterns highlight the varied roles and potential significance of these genes in cancer biology. It underscores the necessity for further research with a larger dataset to validate these findings, particularly for MUC16.

**Prognostic significance of MUC genes in COAD**

To investigate the MUC family that is functional as a potential candidate for mRNA vaccine targets, the prognostic significance of mutations in all MUC genes was analyzed in 2070 samples of COAD with survival data. We identified 3 genes associated with disease-free survival (DFS) (Fig. 2A), and 1 genes with progression-free



**Fig. 1** Interrogating the Mutational Landscape of MUC Genes in the Context of COAD Across Two Distinct Datasets. **A** Illuminating the Mutational Frequencies of Nineteen MUC Genes on cBioPortal website; **B** Providing a Tabular Representation of the Mutation Rates for these Genes; **C** The mutation frequencies of the 19 MUCs genes within the TCGA datasets



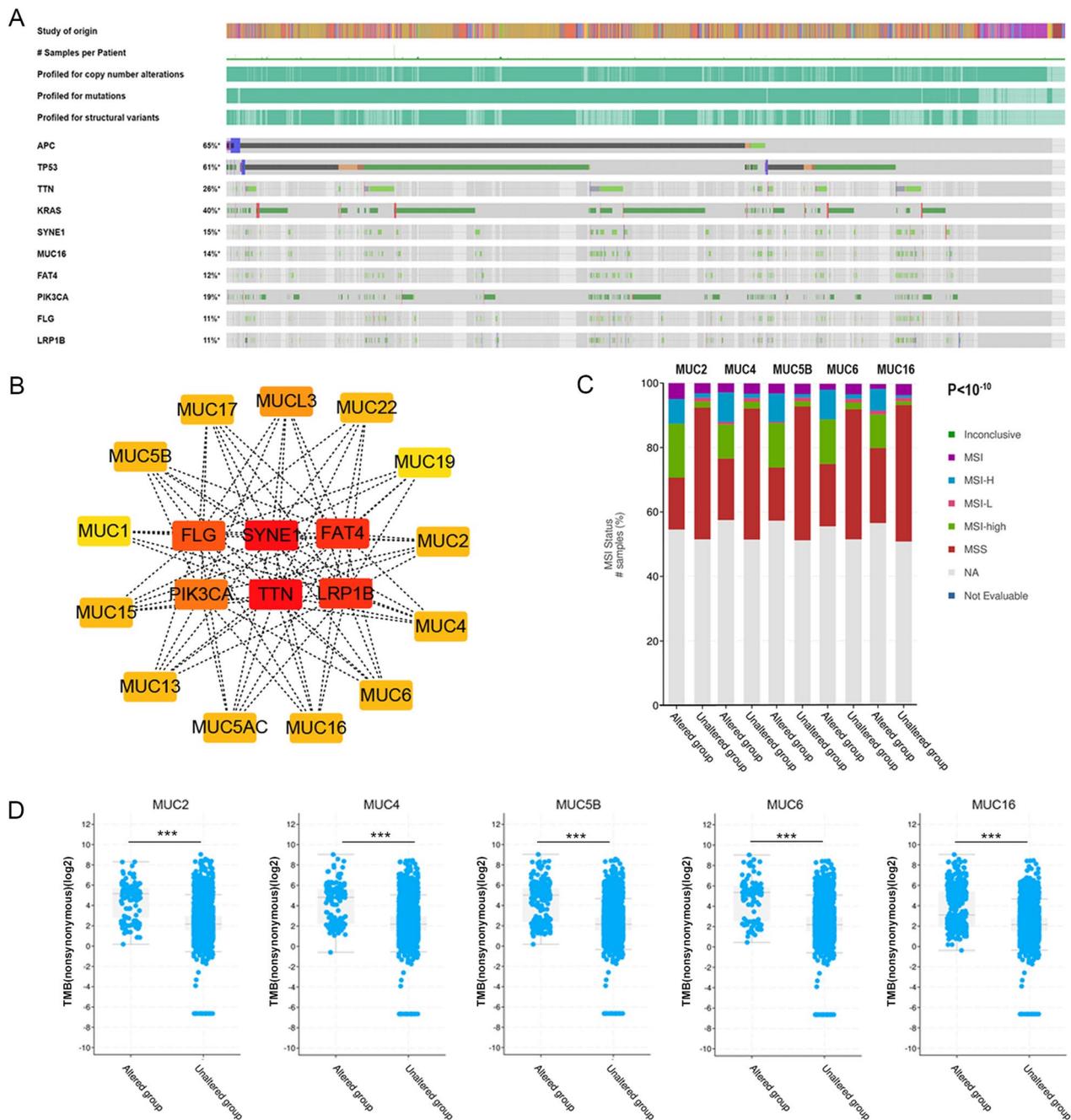
**Fig. 2** Revealing the Prognostic Significance of MUC Genes in two databases. **A** Exploring DFS, PFS, and DSS among COAD Patients, stratified between those with mutations in the 19 MUC genes and those without, as documented on the cBioPortal platform; **B** The Kaplan–Meier curve presents a portrayal of the four preeminent genes in COAD, characterized by their remarkable significance, as elucidated on the cBioPortal website; **C** The OS rate of the 19 MUCs genes mutation group and the non-mutation group in COAD cancer patients on the TCGA website; **D** The KM curve captures the top 3 genes with significant P value in the context of COAD within the TCGA website. DFS, Disease-Free Survival; PFS, Progression-Free Survival; DSS, Disease-Specific Survival; OS, Overall Survival

survival (PFS) (Fig. 2A, B), respectively. Only MUC13, MUC20, and MUC5B were revealed to be closely associated with OS (Fig. 2C, D). In summary, our findings emphasize the crucial role of MUC genes associated with the prognoses of COAD patients, shedding light on the potential implications for clinical practice and therapeutic interventions.

**Exploring the correlation of MUC genes in the context of COAD**

We analyzed the correlation of MUC genes with clinical factors within the context of COAD. Initially, we carefully identified the top ten genes with the highest mutation frequencies in COAD using the cBioPortal website, as depicted in Fig. 3A. Subsequently, we conducted a detailed analysis to explore the intricate interaction between mutated MUC genes and these top ten

genes. In COAD, the identified top ten most frequently mutated genes were APC, TP53, TTN, KRAS, SYNE1, MUC16, FAT4, PIK3CA, FLG, and LRP1B (Fig. 3A). Among these, a subset comprising SYNE1, FAT4, PIK3CA, FLG, and LRP1B demonstrated a significant correlation with mutations in the MUC gene family, which includes MUC1, 2, 3, 4, 5AC, 5B, 6, 13, 15, 16, 17, 19, and 22. Furthermore, we analyzed the relationship between the mutation status of MUC genes and genetic MSI and TMB (Fig. 3C). Our investigation found a strong correlation between the mutation status of five specific MUC genes and TMB (MUC2, MUC4, MUC5B, MUC6 and MUC16) (Fig. 3D). The relevance between MUC mutation and TMB/TSI implied the potential effect of MUC mutation on tumor progression and immunotherapy.



**Fig. 3** The correlation of MUCs with highly mutated genes and genetic instability. **A** The top ten genes with the highest mutation frequencies in COAD on the cBioPortal website; **B** The correlation of mutated MUC genes and the top ten highest mutation genes were analyzed by the cytoscope; **C** The relationship between the mutation status of MUC genes and MSI (MUC16, MUC2, MUC4, MUC5B, and MUC6); **D** The plots showing the association between TMB and mutation status of five MUC genes (MUC2, MUC4, MUC5B, MUC6 and MUC16) frequently mutated in the COAD cohort. MSI, Microsatellite Instability; TMB, Tumor Mutational Burden

### Immunological correlation of MUC family mutation

To depict the immunological role of MUC mutation comprehensively that targeting MUC mutation might benefit from immunotherapy in COAD, we investigated the association between MUC family mutation

and tumor microenvironment (TME) and immune subtype. The cancer-associated immune cells were related to MUC family mutation in COAD (Fig. 4A–E). We assessed the enrichment scores of MUC family mutated- and wildtype- cohorts in immune cells i.e.,

CD8<sup>+</sup>T cells, CD4<sup>+</sup>T cells, dendritic cells, myeloid-derived suppressor cells (MDSCs), macrophages, B cells, neutrophil cells, NK, and Treg cells in COAD. Furthermore, the enrichment scores within the different cohorts of MUC copy number variation (CNV) showed significant differences (Fig. 5A–E, Additional file 1: Fig. S1, Additional file 2: Fig. S2, Additional file 3: Fig. S3 and Additional file 4: Fig. S4) in immune cells, including CD8<sup>+</sup>T cells, CD4<sup>+</sup>T cells, dendritic cells, neutrophil cells, Th2 cells, B cells, and NKT. Comprehensive analyses of above data about the associations of MUCmut in the top six mutated genes with different immune parameters, we found the common points, that the mutations of the MUC family i.e., 16, 17, 5B, 4, 2, and 6 were immunogenic to patients, including higher immune scores and more immune cells, especially CD8<sup>+</sup>T cells.

The immune checkpoint (ICP) genes were found to play a role in immune cell infiltration and immunotherapy outcomes [43, 46]. Our result indicates that MUC family gene mutation was positively related to the expression of these ICP genes in COAD: i.e. immunoinhibitors including CD244, CD274, CD96, CTLA4, HAVCR2, IDO1, LAG3, PDCD1, PDCD1 LG2, TGFβ1 and TIGIT etc. in COAD (Fig. 6A); immunostimulators including C10orf54, CXCR4, KLRC1, KLRK1, MICB, TNFSF13, and TNFRSF25 (Fig. 6B); MHC molecules including HLA-DMA, HLA-G, Transporter associated with antigen processing (TAP)1, TAP2, TAPBP etc. (Fig. 6C); chemokine including CCL4, CCL5, CCL24, CXCL9, CXCL13, CXCL14 and CXCL16 (Fig. 6D); receptor including CCR3 and CXCR4 (Fig. 6E). Additionally, high MUC16 expression was also promoted in the C1 (wound healing) immune subtypes and MUC5B in CIN molecular subtypes (Fig. 6F, G), which indicated that MUC mutated genes may be involved in TME remodeling. Since TMB, MSI, and NEO are considered predictors for response to tumor immunotherapy within the TME [47, 48], our above-analyzed data showed a positive association of TMB and MSI with MUCmut in COAD (Fig. 3C, D). Collectively, MUC family mutation may affect antitumor immunity through its association with immune infiltrating cells, ICPs, MSI, and TMB in COAD tissues. These findings

suggested that peptides from MUCmut might be potential neoantigens for COAD.

#### Characteristics of neoantigens of MUCmut in COAD

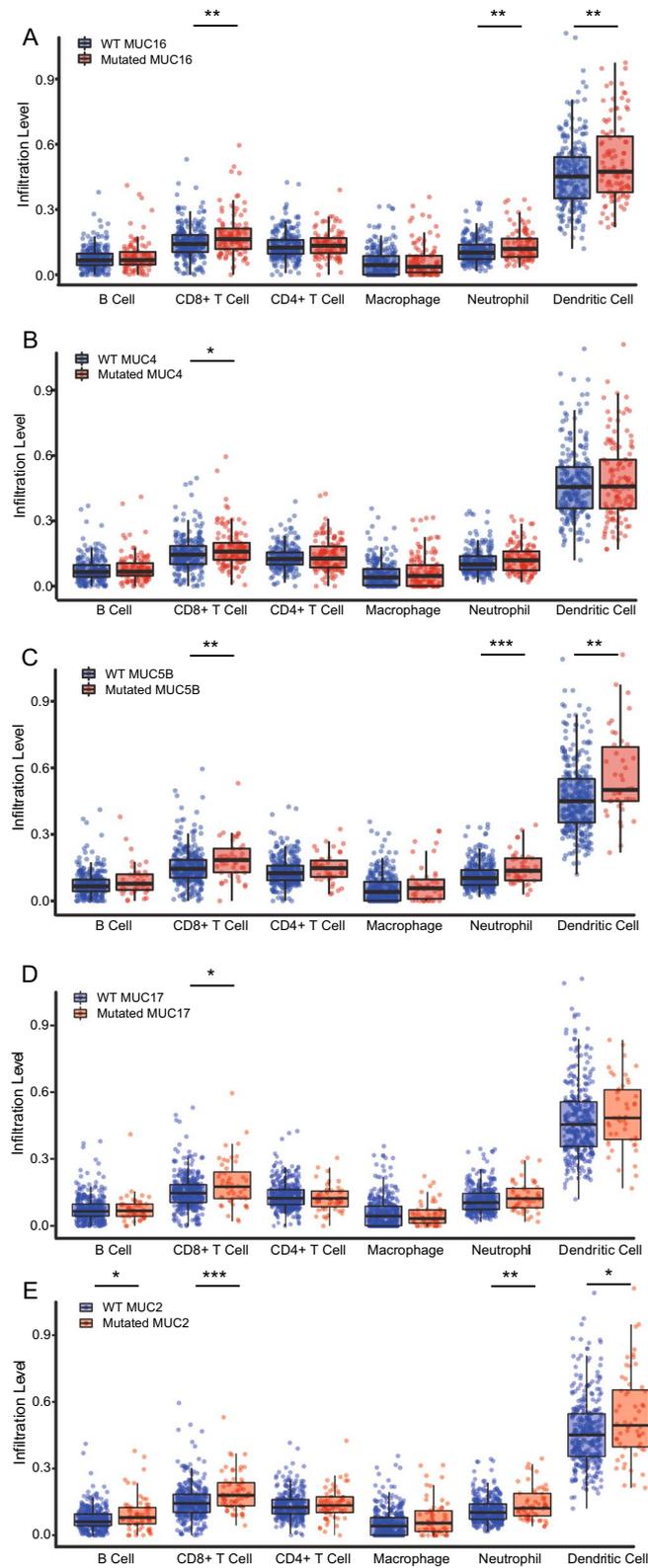
To discover the potential neoantigens profiling of mutated genes in COAD, we comprehensively analyzed all mutations of the MUC family in 531 COAD samples obtained from the TSNAdb database [27] using the method of DAI to select high-confidence TSA. We first confirmed that MUC16 was ranked as the top 6 most frequently mutated genes (Fig. 3A), and was moved to the top 2 generate-neoantigens across all mutated genes in COAD (Additional file 5: Fig. S5). Next, 1243 SNVs of MUCmut neoantigens by TSNAD v1.0 NetMHCpan v2.8, and 452 SNVs of MUCmut neoantigens by TSNAD v1.0 NetMHCpan v4.0 were identified (Additional file 7: Table S2). The order for the number of SNV neoantigens was 282 for MUC16mut, 38 for MUC17mut, 35 for MUC5mut, 28 for MUC4mut, 27 for MUC6mut and 17 for MUC5ACmut (Fig. 7B). The two most binding HLA alleles of MUCmut neoantigens were HLA-C\*12:03 and HLA-A\*02:01 (Fig. 7A), which also ranked as the top 2 HLA alleles of predicted neoantigens in COAD by TSNAdb1.0/NetMHCpan v2.8 (Additional file 5: Fig. 5B). Next, we use TSNAdb v2.0 to identify higher confidence neoantigens of MUCmut. The results showed that 325 potential neoantigens of MUC mutated genes involving 157 INDELS, 1 TS, and 57 SNVs were identified (Additional file 8: Table S3). Among 157 INDELS, the top 3 MUCmut neoantigens were MUC4, 5B, and 5AC (Fig. 7D and Additional file 7: Table S2). Moreover, we applied the tool of DAI to select the higher-confidence affinity neoantigens and obtained 10 high-confidence SNV MUCmut neoantigens out of 57 SNVs (Fig. 7C).

#### Discussion

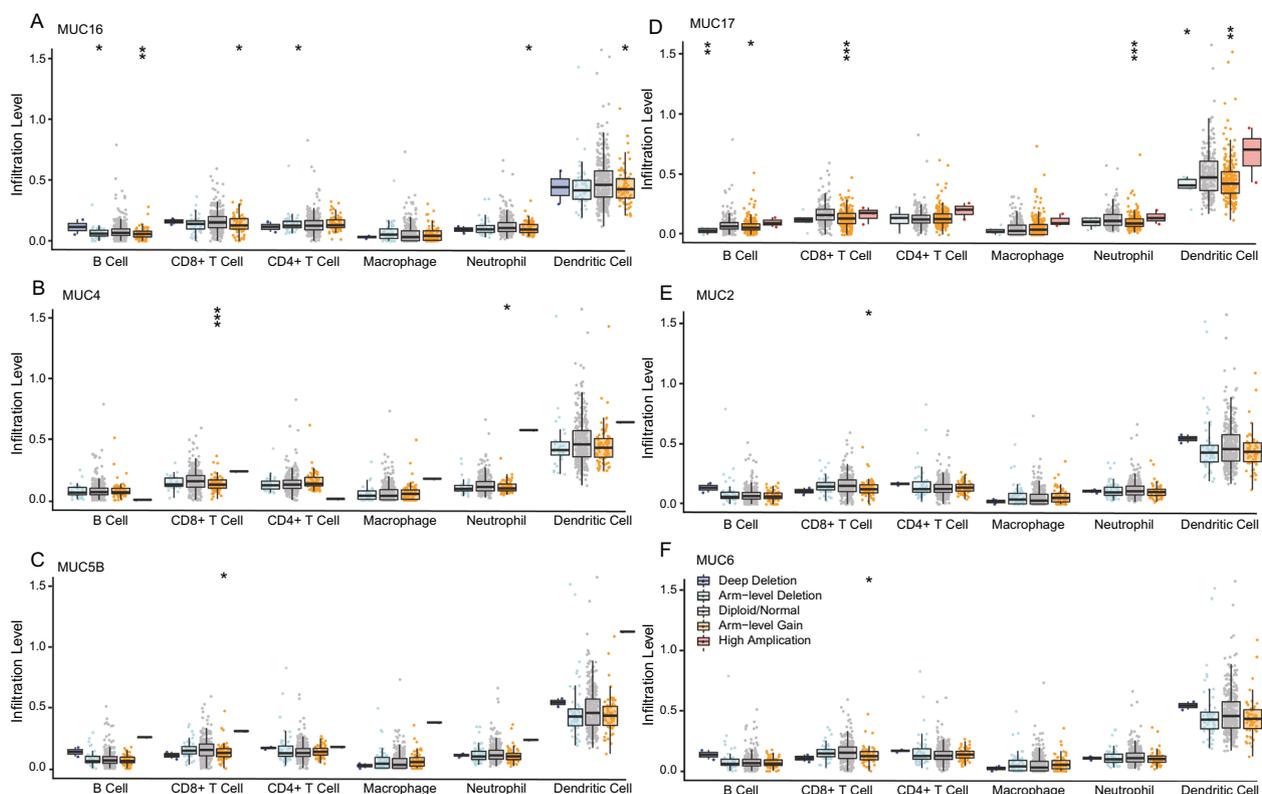
Identification of immunogenic neoantigens from numerous sources is a crucial step in the development of effective immunotherapies [23]. However, the reported studies of the identification of potential neoantigens in COAD are rare. To address whether the MUC family can potentially target tumor-specific immunotherapy in COAD, we performed a thorough screen across the MUCmut and a detailed analysis of immune infiltration estimation algorithms of MUCmut in COAD. After deep analysis, our

(See figure on next page.)

**Fig. 4** Five MUC genes comprehensive analysis of tumor-infiltrating immune cells (B cell, CD8<sup>+</sup>T cell, CD4<sup>+</sup>T cell, Macrophage, Neutrophil, Dendritic cell) in COAD using the Timer database. **A** MUC16; **B** MUC4; **C** MUC5B; **D** MUC17; **E** MUC2. The mutation module compares the levels of immune infiltrates with or without the presence of a given mutation. Box plots are generated for each immune subset, to compare the distributions of immune infiltration levels under different gene mutation statuses, with statistical significance estimated using a two-sided Wilcoxon rank-sum test



**Fig. 4** (See legend on previous page.)



**Fig. 5** SCNA module provides the comparison of tumor infiltration levels among tumors with different somatic copy number alterations for a given gene. SCNAs are defined by GISTIC 2.0, including deep deletion (−2), arm-level deletion (−1), diploid/normal (0), arm-level gain (1), and high amplification (2). Box plots are presented to show the distributions of each immune subset at each copy number status in COAD, i.e., **A** MUC16; **B** MUC4; **C** MUC5B; **D** MUC17; **E** MUC2; **F** MUC6. The infiltration level for each SCNA category is compared with the normal using a two-sided Wilcoxon rank-sum test

data revealed the potential value of MUC mutated genes as a predictor of prognosis and immunological biomarkers. Next, we performed a typical workflow for neoantigen prediction as following steps: (i) mutation calling, (ii) HLA typing, (iii) neoantigen filtering and prioritization based on HLA binding affinity. Through our analysis with two neoantigen-related databases (TSNadb and NetMHCpan) plus DAI, we obtained 10 high-confidence SNVs and 157 INDELS neoantigens from 6 MUC mutate genes (MUC16, MUC17, MUC4, MUC6, MUC2, and MUC5B) that may serve as candidate targets for neoantigen vaccines.

Currently, the initial stage in the process of detecting possible neoantigens from NGS data is mapping

tumor-specific genetic abnormalities using WES of the tumor and normal DNA. In this study, with the integrated analysis of the 6261 COAD samples, somatic genomic alterations for the MUC family include SNVs and gene fusions which might promote the production of tumor neoantigens. The top six mutated family members were MUC16, MUC17, MUC5B, MUC2, MUC4, and MUC6 with the mutation frequencies from 14 to 2.9%. The high mutation of MUC16, MUC4 and MUC6 had a strong association with PFS or DFS in COAD, which was consistent with previous studies for MUC16 [49] and MUC4 [50].

Currently, multiple potential factors have been identified for predicting the clinical outcome of

(See figure on next page.)

**Fig. 6** Immune and molecular correlation with the mutation or expression of MUC family. **A** immunoinhibitors; **B** immunostimulators; **C** MHC molecules; **D** chemokine; **E** receptor. These immunomodulators were collected from Charoentong’s study. **F** MUC16 expression was high associated with immune and molecular subtypes, respectively in COAD, Kruskal–Wallis Test: P value=1.96e−06; n=C1 332, C2 85, C3 9, C4 12, C6 3; **G** MUC 5B expression was high associated with immune and molecular subtypes, respectively in COAD, Kruskal–Wallis Test: P value=9.5e−08; n=C1N 226, GS 49, HM-SNV 6, HM-INDEL 20

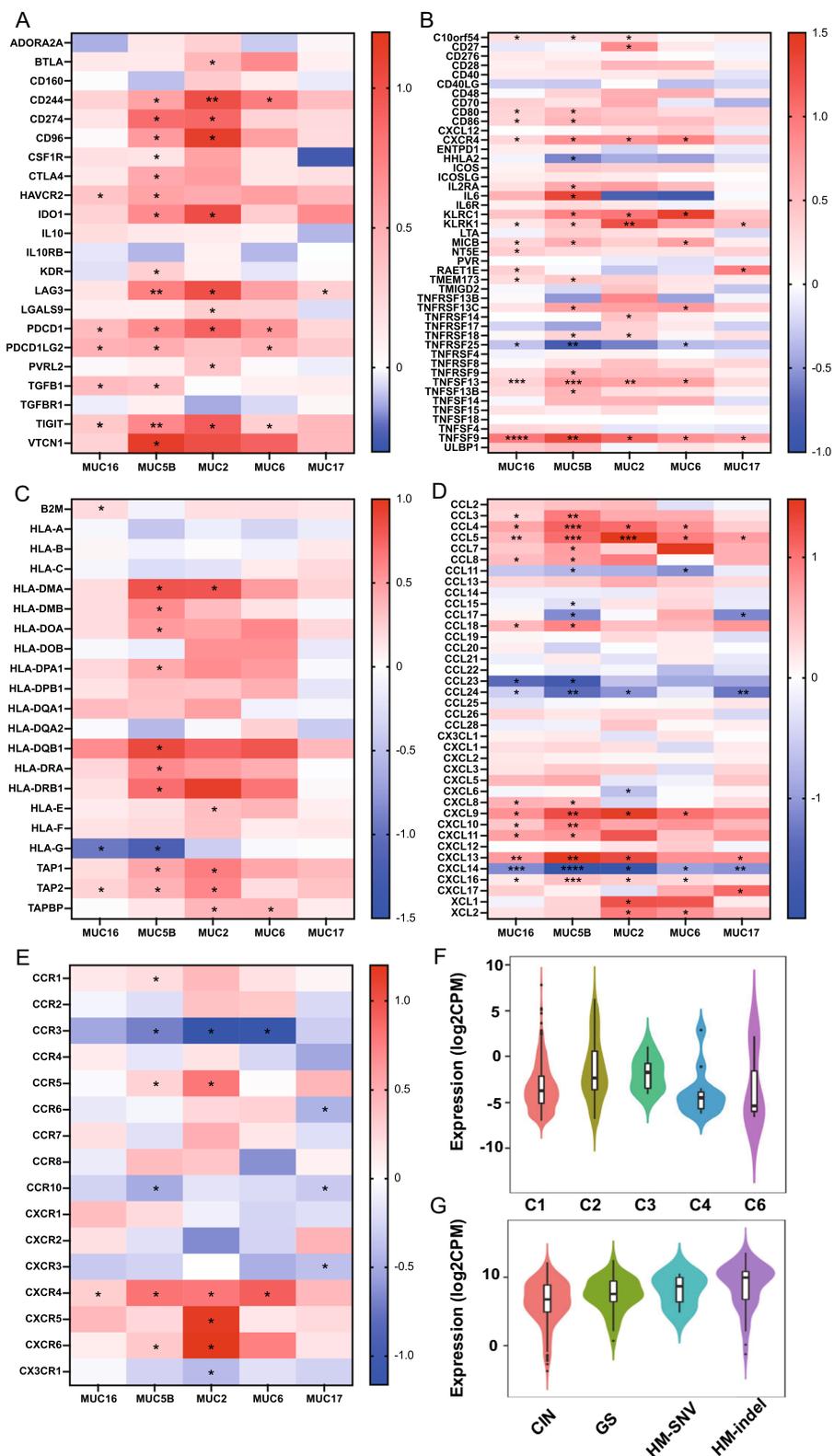
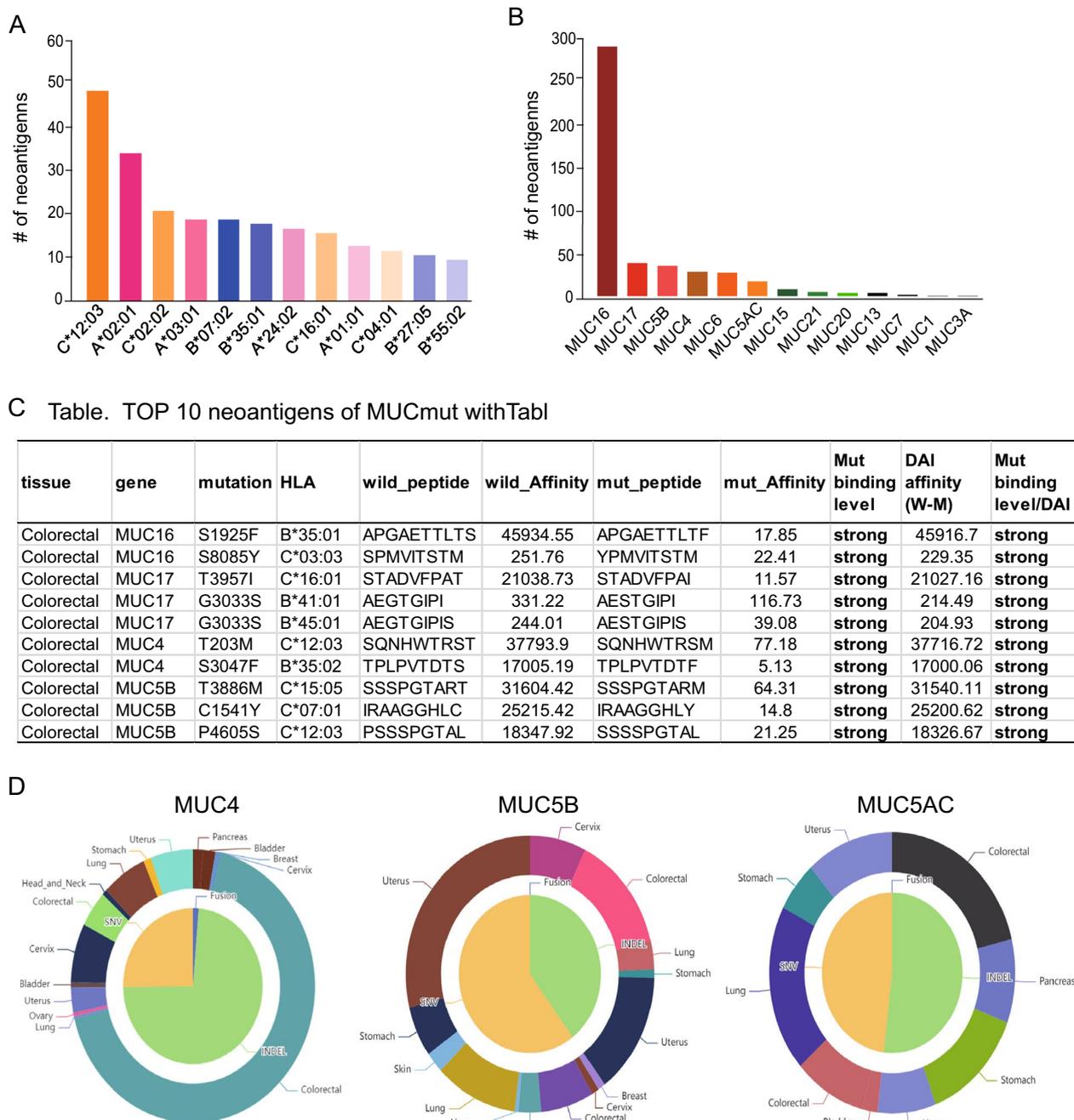


Fig. 6 (See legend on previous page.)



**Fig. 7** Predicted neoantigens of mutated MUC family for COAD The top 12 HLA alleles **A** in MUCmut **B** with the number of predicted neoantigens are displayed in COAD, with the detailed neoantigen information listed **C** The binding level 'Strong' indicates strong binding with IC50 < 150 nM based on Mutate binding level or DAI > 150 nM. **D** The INDEL mutational signature of MUCmut in COAD

immunotherapy, including the tumor mutation burden (TMB) [51], MSI status [7], neoepitope load [52–55], PD-L1 level [55], CD8<sup>+</sup> T-cell density [56], interferon-γ gene signature [57], and MHC and T-cell receptor repertoire [58, 59]. The TMB and MSI were applied as clinical biomarkers for a potential response to ICI

immunotherapy in colorectal and other solid tumors [51]. Our study showed a strong association between several MUC-mutated genes with TMB and MSI in COAD. This finding was in line with previous studies, showing that the cancer-associated mucins play a role in immune modulation and metastasis [60]. Tumor-infiltrating

lymphocytes in the tumor microenvironment have been indicated to be efficient in predicting prognosis and immunotherapeutic efficacy for cancer [58]. In the present study, we found that the infiltration levels of several effector TIICs, such as activated CD8<sup>+</sup> T cells, activated CD4<sup>+</sup> T cells, activated B cells, T Neutrophil cells, dendritic cells, and T helper cells, were significantly upregulated in COAD patients with MUC mutated genes. In addition, ICB targeting these ICPs, such as PD1-blocking antibodies, pembrolizumab, and nivolumab, have shown efficacy in patients with metastatic MSI-H CRC, and they have been granted accelerated FDA approval [7]. Our study showed a strong relationship between MUC mutated genes and key ICPs, such as CD274 (PD-L1), CTLA4, HAVCR2, IDO1, LAG3, PDCD1, TGFβ1 and TIGIT, as well as MSI and TMB in COAD. This indicates that it can be of value in further investigations on developing novel immunotherapeutic strategies as tumor neoantigens to target MUC mutated genes, which might benefit COAD patients. Intriguing, cancer mutations can form neo-epitopes recognized by T cells on HLA molecules, which contributes to the clinical success of immunotherapy [18, 52, 61–63], and creates enthusiasm for neo-epitope vaccines [14, 64]. In our analyses, MHC molecules include HLA-DMA, HLA-G, TAP1, TAP2, and TAPBP. were highly related to identified mutations of the MUC family.

Based on analyses with TSNAD v1.0 NetMHCpan v4.0 [28, 65, 66], we obtained 452 SNV unique neoantigens involving 13 MUC genes in 531 COAD patients (Additional file 6: Table S1). Recent clinical studies have shown a correlation between the burden of strong-binding neopeptides with an affinity for MHC-I of <500 nM (referred to as neoantigens) and patient outcomes in advanced melanoma and lung cancer [52, 61, 62]. After we defined the binding level 'Strong' indicates strong binding with IC<sub>50</sub> < 150 nM based on Mutate binding level or 'Weak' indicates weak binding with 150 nM < IC<sub>50</sub> < 500 nM, we obtained 271 'Strong' neoantigens and 181 'Weak' neoantigens of MUC. Additionally, an important future field is to discovery of neoantigens created by cancer-specific INDELS (insertions and deletions), fusion genes, and splice variants that have a lower degree of similarity to self-antigens than SNV-derived neoantigens.

Similar to SNV-encoded neoantigens, INDEL encoded neoantigens are more common in cancers with MSI-H, which is determined by the absence of DNA mismatch repair (MMR) mechanisms. As the evolution of MMR-deficient cancers is mainly triggered by mutations that inactivate tumor suppressor genes containing coding microsatellites, frameshift peptide neoantigens are more frequently shared among MMR-deficient cancers than missense mutation-derived neoantigens [67]. In this

study, through comprehensive analyses by TSNAD v2.0 NetMHCpan v4.0, 157 INDELS, 1 TS, and 57 SNVs were identified as potential neoantigens for MUC mutated genes showing high affinity with MHC-I. Interestingly, 94 INDELS were for MUC4, which is consistent with a previous study that found that MUC4 was highly alternative splicing in CRC [68].

Mechanistically, the difference in predicted affinity for any paired wild-type/mutant peptides termed as differential agretopicity index (DAI) is a broad and better indicator of neuropeptide dissimilarity from self and a feature of immunogenicity [69–71]. In this study, among TSNAD v2.0 NetMHCpan v4.0 analyzed data, of 59 SNVs, 10 mutant neopeptides selected in the group of both binding level 'Strong' and DAI strong were 2 MUC16, 3 MUC17, 3 MUC5B, 2 MUC4. In Prof. Rosenberg's clinical study [13] on the identification of tumor neoantigen using reactive tumor-infiltrating lymphocytes from 75 patient gastrointestinal cancers, one patient of MUC4 mutant neoantigen was recognized by CD8<sup>+</sup> T cells and another patient with MUC6 mutant neoantigen were recognized by CD4<sup>+</sup> T cells. All these results greatly expand the universe of target cancer antigens and identify new tools for anti-MUC immunotherapy in COAD.

## Conclusion

In conclusion, we identified neoantigen candidates of MUC mutant genes in COAD by the combination of a comprehensive analysis and set up a prediction model for initiating the selection of high-confidence neoantigens derived from MUC mutations. These candidates could be potential targets for neoantigen-based immunotherapy. The validation for the strong immunogenic effects of these potential antigens should be further evaluated in a large patient cohort before being utilized in clinical setting. This study could promote neoantigen-based immunotherapy for broader application in COAD.

## Abbreviations

COAD	Colon adenocarcinoma
SNVs	Single nucleotide variants
CNV	Copy number variation
TMB	Tumor mutation burden
TSI	Tumor-specific immunoreactivity
TME	Tumor microenvironment
ICP	Immune checkpoint
dMMR	Deficient mismatch repair
ICIs	Immune-checkpoint inhibitors
CRC	Colorectal cancer
pMMR	Proficient mismatch repair
MSI-L	Microsatellite instability-low
TAP	Transporter associated with antigen processing
MHC	Major histocompatibility complex
OS	Overall survival
DFS	Disease-free survival rate
PFS	Progression-free survival rate
DSS	Disease-specific survival rate
TCIA	The Cancer Immunome Atlas

TME Tumor microenvironment  
MDSCs Myeloid-derived suppressor cells

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-024-05103-z>.

**Additional file 1: Figure S1.** Relations between the abundance of tumor-infiltrating lymphocytes (TILs) and the mutation of MUC16. The immune-related signatures types from Charoentong's study, the relative abundance of TILs in COAD was inferred by using gene set variation analysis (GSVA) based on gene expression profile.

**Additional file 2: Figure S2.** Relations between the abundance of TILs and the mutation of MUC5B. The immune-related signatures types from Charoentong's study, the relative abundance of TILs in COAD was inferred by using GSVA based on gene expression profile.

**Additional file 3: Figure S3.** Relations between the abundance of TILs and the mutation of MUC17. The immune-related signatures types from Charoentong's study, the relative abundance of TILs in COAD was inferred by using GSVA based on gene expression profile.

**Additional file 4: Figure S4.** Relations between the abundance of TILs and the mutation of MUC6. The immune-related signatures types from Charoentong's study, the relative abundance of TILs in COAD was inferred by using GSVA based on gene expression profile.

**Additional file 5: Figure S5.** The top 20 genes and HLA alleles with the number of predicted neoantigens are displayed in COAD by TSNAdb1.0/NetMHCpan v2.8.

**Additional file 6: Table S1.** Mutation Rates of MUC family within the TCGA Dataset.

**Additional file 7: Table S2.** MUCmut neoantigens predicted by TSNAD v1.0/NetMHCpan v2.8.

**Additional file 8: Table S3.** MUCmut neoantigens by TSNAD v2.0/NetMHCpan v4.0.

## Acknowledgements

We gratefully acknowledge for TCGA, TIMER, cBioPortal, and TSNAdb, which made the genomic data and clinical data of COAD available.

## Author contributions

All authors made a significant contribution to this study, whether that is in the conception by MC, study design by MC and HXA, execution by XZ, ZHM and MC, acquisition of data by LY and XRF, analysis by XZ, ZHM, MC and HXA, and interpretation by ZGH, MC and HXA, or all these areas; took part in drafting, revising, or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

This work was supported by the Natural Science Foundation Committee of China (Grant No. 82172009), Shanxi Bethune Hospital of Shanxi Medical University (Grant No. 2023GZR204), Natural Science Foundation of Shanxi Province of China (Grant No. 202303021221200).

## Availability of data and materials

The datasets used and/or analyzed during this study are available from the corresponding author upon reasonable request.

## Declarations

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

## Author details

<sup>1</sup>Clinical Central Research Core, Shanxi Bethune Hospital, Shanxi Medical University, Taiyuan, Shanxi, China. <sup>2</sup>The Center Laboratory, Shanghai Medical College, Zhongshan Hospital (Xiamen Affiliated) of Fudan University, Fudan University, Xiamen, China. <sup>3</sup>Cancer Center and Department of Breast and Thyroid Surgery, School of Medicine, Xiang'an Hospital of Xiamen University, Xiamen University, Xiamen, China. <sup>4</sup>Shanxi Bethune Hospital, Shanxi Medical University, Taiyuan, Shanxi, China. <sup>5</sup>Department of Chemistry and Key Laboratory for Preparation and Application of Ordered Structural Materials of Guangdong Province, Chemistry and Chemical Engineering Guangdong Laboratory, Shantou University, Guangdong, China. <sup>6</sup>Department of Colorectal and Anal Surgery, Shanxi Bethune Hospital, Shanxi Medical University, Taiyuan, Shanxi, China. <sup>7</sup>The Cancer Center, Shanxi Bethune Hospital, Shanxi Medical University, Taiyuan, Shanxi, China.

Received: 5 January 2024 Accepted: 15 March 2024

Published online: 27 March 2024

## References

- Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. *Lancet*. 2019;394:1467–80.
- Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. *CA Cancer J Clin*. 2021;71:7–33.
- Couzin-Frankel J. Breakthrough of the year 2013 Cancer immunotherapy. *Science*. 2013;342:1432–3.
- Ribas A, Wolchok JD. Cancer immunotherapy using checkpoint blockade. *Science*. 2018;359:1350–5.
- Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med*. 2018;24:1550–8.
- Lichtenstern CR, Ngu RK, Shalapour S, Karin M. Immunotherapy inflammation and colorectal cancer. *Cells*. 2020;9:618–35.
- Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, Diaz LA Jr. Immunotherapy in colorectal cancer: rationale, challenges and potential. *Nat Rev Gastroenterol Hepatol*. 2019;16:361–75.
- Ciardello D, Vitiello PP, Cardone C, Martini G, Troiani T, Martinelli E, Ciardiello F. Immunotherapy of colorectal cancer: Challenges for therapeutic efficacy. *Cancer Treat Rev*. 2019;76:22–32.
- Wang QJ, Yu Z, Griffith K, Hanada K, Restifo NP, Yang JC. Identification of T-cell receptors targeting KRAS-mutated human tumors. *Cancer Immunol Res*. 2016;4:204–14.
- Veatch JR, Jesernig BL, Kargl J, Fitzgibbon M, Lee SM, Baik C, Martins R, Houghton AM, Riddell SR. Endogenous CD4(+) T cells recognize neoantigens in lung cancer patients, including recurrent oncogenic KRAS and ERBB2 (Her2) driver mutations. *Cancer Immunol Res*. 2019;7:910–22.
- Chandran SS, Ma J, Klatt MG, Dündar F, Bandlamudi C, Razavi P, Wen HY, Weigelt B, Zumbo P, Fu SN, et al. Immunogenicity and therapeutic targeting of a public neoantigen derived from mutated PIK3CA. *Nat Med*. 2022;28:946–57.
- Tran E, Ahmadzadeh M, Lu YC, Gros A, Turcotte S, Robbins PF, Gartner JJ, Zheng Z, Li YF, Ray S, et al. Immunogenicity of somatic mutations in human gastrointestinal cancers. *Science*. 2015;350:1387–90.
- Parkhurst MR, Robbins PF, Tran E, Prickett TD, Gartner JJ, Jia L, Ivey G, Li YF, El-Gamil M, Lalani A, et al. Unique neoantigens arise from somatic mutations in patients with gastrointestinal cancers. *Cancer Discov*. 2019;9:1022–35.
- Carreno BM, Magrini V, Becker-Hapak M, Kaabinejadian S, Hundal J, Petti AA, Ly A, Lie WR, Hildebrand WH, Mardis ER, Linette GP. Cancer immunotherapy: a dendritic cell vaccine increases the breadth and diversity of melanoma neoantigen-specific T cells. *Science*. 2015;348:803–8.
- Ott PA, Hu Z, Keskin DB, Shukla SA, Sun J, Bozym DJ, Zhang W, Luoma A, Giobbie-Hurder A, Peter L, et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature*. 2017;547:217–21.
- Sahin U, Derhovanessian E, Miller M, Kloke BP, Simon P, Lower M, Bukur V, Tadmor AD, Luxemburger U, Schrors B, et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature*. 2017;547:222–6.
- Lennerz V, Fatho M, Gentilini C, Frye RA, Lifke A, Ferel D, Wolfel C, Huber C, Wolfel T. The response of autologous T cells to a human melanoma

- is dominated by mutated neoantigens. *Proc Natl Acad Sci U S A*. 2005;102:16013–8.
18. Tran E, Robbins PF, Lu YC, Prickett TD, Gartner JJ, Jia L, Pasetto A, Zheng Z, Ray S, Groh EM, et al. T-cell transfer therapy targeting mutant KRAS in cancer. *N Engl J Med*. 2016;375:2255–62.
  19. Simnica D, Kobold S. Neoantigen T-cell receptor gene therapy in pancreatic cancer. *N Engl J Med*. 2022;387:573.
  20. Fang Y, Mo F, Shou J, Wang H, Luo K, Zhang S, Han N, Li H, Ye S, Zhou Z, et al. A pan-cancer clinical study of personalized neoantigen vaccine monotherapy in treating patients with various types of advanced solid tumors. *Clin Cancer Res*. 2020;26:4511–20.
  21. Palmer CD, Rappaport AR, Davis MJ, Hart MG, Scallan CD, Hong SJ, Gitlin L, Kraemer LD, Kounlavouth S, Yang A, et al. Individualized, heterologous chimpanzee adenovirus and self-amplifying mRNA neoantigen vaccine for advanced metastatic solid tumors: phase 1 trial interim results. *Nat Med*. 2022;28:1619–29.
  22. Kloor M, Reuschenbach M, Pauligk C, Karbach J, Rafiyan MR, Al-Batran SE, Tariverdian M, Jager E, von Knebel DM. A Frameshift peptide neoantigen-based vaccine for mismatch repair-deficient cancers: a phase I/IIa clinical trial. *Clin Cancer Res*. 2020;26:4503–10.
  23. Barbier AJ, Jiang AY, Zhang P, Wooster R, Anderson DG. The clinical progress of mRNA vaccines and immunotherapies. *Nat Biotechnol*. 2022;40:840–54.
  24. Borden ES, Buetow KH, Wilson MA, Hastings KT. Cancer neoantigens: challenges and future directions for prediction, prioritization, and validation. *Front Oncol*. 2022;12: 836821.
  25. Segal NH, Parsons DW, Peggs KS, Velculescu V, Kinzler KW, Vogelstein B, Allison JP. Epitope landscape in breast and colorectal cancer. *Cancer Res*. 2008;68:889–92.
  26. Sjöblom T, Jones S, Wood LD, Parsons DW, Lin J, Barber TD, Mandelker D, Leary RJ, Ptak J, Silliman N, et al. The consensus coding sequences of human breast and colorectal cancers. *Science*. 2006;314:268–74.
  27. Wu J, Zhao W, Zhou B, Su Z, Gu X, Zhou Z, Chen S. TSNAdb: a database for tumor-specific neoantigens from immunogenomics data analysis. *Genomics Proteomics Bioinformatics*. 2018;16:276–82.
  28. Hoof I, Peters B, Sidney J, Pedersen LE, Sette A, Lund O, Buus S, Nielsen M. NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics*. 2009;61:1–13.
  29. Jurtz V, Paul S, Andreatta M, Marcatili P, Peters B, Nielsen M. NetMHCpan-4.0: improved peptide-MHC Class I interaction predictions integrating eluted ligand and peptide binding affinity data. *J Immunol*. 2017;199:3360–8.
  30. Dhanda SK, Mahajan S, Paul S, Yan Z, Kim H, Jespersen MC, Jurtz V, Andreatta M, Greenbaum JA, Marcatili P, et al. IEDB-AR: immune epitope database-analysis resource in 2019. *Nucleic Acids Res*. 2019;47:W502–6.
  31. Apostolopoulos V, Yuriev E, Ramsland PA, Halton J, Osinski C, Li W, Plebanski I, Paulsen H, McKenzie IF. A glycopeptide in complex with MHC class I uses the GalNAc residue as an anchor. *Proc Natl Acad Sci U S A*. 2003;100:15029–34.
  32. Vlad AM, Muller S, Cudic M, Paulsen H, Otvos L Jr, Hanisch FG, Finn OJ. Complex carbohydrates are not removed during processing of glycoproteins by dendritic cells: processing of tumor antigen MUC1 glycopeptides for presentation to major histocompatibility complex class II-restricted T cells. *J Exp Med*. 2002;196:1435–46.
  33. Wu X, Yin Z, McKay C, Pett C, Yu J, Schorlemer M, Gohl T, Sungsuwan S, Ramadan S, Baniel C, et al. Protective epitope discovery and design of MUC1-based vaccine for effective tumor protections in immunotolerant mice. *J Am Chem Soc*. 2018;140:16596–609.
  34. Quoix E, Ramlau R, Westeel V, Papai Z, Madroszyk A, Riviere A, Koralewski P, Breton JL, Stoelben E, Braun D, et al. Therapeutic vaccination with TG4010 and first-line chemotherapy in advanced non-small-cell lung cancer: a controlled phase 2B trial. *Lancet Oncol*. 2011;12:1125–33.
  35. Butts C, Socinski MA, Mitchell PL, Thatcher N, Havel L, Krzakowski M, Nawrocki S, Ciuleanu TE, Bosquee L, Trigo JM, et al. Tecemotide (L-BLP25) versus placebo after chemoradiotherapy for stage III non-small-cell lung cancer (START): a randomised, double-blind, phase 3 trial. *Lancet Oncol*. 2014;15:59–68.
  36. Kufe DW. Mucins in cancer: function, prognosis and therapy. *Nat Rev Cancer*. 2009;9:874–85.
  37. Kesari MV, Gaopande VL, Joshi AR, Babanagare SV, Gogate BP, Khadilkar AV. Immunohistochemical study of MUC1, MUC2 and MUC5AC in colorectal carcinoma and review of literature. *Indian J Gastroenterol*. 2015;34:63–7.
  38. Imai Y, Yamagishi H, Fukuda K, Ono Y, Inoue T, Ueda Y. Differential mucin phenotypes and their significance in a variation of colorectal carcinoma. *World J Gastroenterol*. 2013;19:3957–68.
  39. Shanmugam C, Jhala NC, Katkooi VR, Wan W, Meleth S, Grizzle WE, Manne U. Prognostic value of mucin 4 expression in colorectal adenocarcinomas. *Cancer*. 2010;116:3577–86.
  40. Kocer B, Soran A, Erdogan S, Karabeyoglu M, Yildirim O, Eroglu A, Bozkurt B, Cengiz O. Expression of MUC5AC in colorectal carcinoma and relationship with prognosis. *Pathol Int*. 2002;52:470–7.
  41. Renaud F, Vincent A, Mariette C, Crepin M, Stechly L, Truant S, Copin MC, Porchet N, Leteurtre E, Van Seuningen I, Buisine MP. MUC5AC hypomethylation is a predictor of microsatellite instability independently of clinical factors associated with colorectal cancer. *Int J Cancer*. 2015;136:2511–21.
  42. Betge J, Schneider NI, Harbaum L, Pollheimer MJ, Lindtner RA, Kornprat P, Ebert MP, Langner C. MUC1, MUC2, MUC5AC, and MUC6 in colorectal cancer: expression profiles and clinical significance. *Virchows Arch*. 2016;469:255–65.
  43. Hu X, Wang Z, Wang Q, Chen K, Han Q, Bai S, Du J, Chen W. Molecular classification reveals the diverse genetic and prognostic features of gastric cancer: a multi-omics consensus ensemble clustering. *Biomed Pharmacother*. 2021;144: 112222.
  44. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, Li B, Liu XS. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res*. 2020;48:W509–w514.
  45. Bjerregaard AM, Nielsen M, Jurtz V, Barra CM, Hadrup SR, Szallasi Z, Eklund AC. An analysis of natural T cell responses to predicted tumor neoepitopes. *Front Immunol*. 2017;8:1566.
  46. Topalian SL, Drake CG, Pardoll DM. Immune checkpoint blockade: a common denominator approach to cancer therapy. *Cancer Cell*. 2015;27:450–61.
  47. Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. *N Engl J Med*. 2017;377:2500–1.
  48. Peng M, Mo Y, Wang Y, Wu P, Zhang Y, Xiong F, Guo C, Wu X, Li Y, Li X, et al. Neoantigen vaccine: an emerging tumor immunotherapy. *Mol Cancer*. 2019;18:128.
  49. Huang L, Zheng S, Fu J, Zhang M, Ge X, Mu N. Bioinformatics analysis based on TCGA: MUC16 mutation correlates with clinical outcome in gastric cancer. *Dis Markers*. 2022;2022:6734105.
  50. Lu S, Catalano C, Huhn S, Pardini B, Partu L, Vymetalkova V, Vodickova L, Levy M, Buchler T, Hemminki K, et al. Single nucleotide polymorphisms within MUC4 are associated with colorectal cancer survival. *PLoS ONE*. 2019;14: e0216666.
  51. Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, Barron DA, Zehir A, Jordan EJ, Omuro A, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet*. 2019;51:202–6.
  52. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, Lee W, Yuan J, Wong P, Ho TS, et al. Cancer immunology. mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science*. 2015;348:124–8.
  53. Asaoka Y, Ijichi H, Koike K. PD-1 blockade in tumors with mismatch-repair deficiency. *N Engl J Med*. 1979;2015:373.
  54. Chan TA, Wolchok JD, Snyder A. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med*. 1984;2015:373.
  55. Nishino M, Ramaiya NH, Hatabu H, Hodi FS. Monitoring immune-checkpoint blockade: response evaluation and biomarker development. *Nat Rev Clin Oncol*. 2017;14:655–68.
  56. van der Leun AM, Thommen DS, Schumacher TN. CD8(+) T cell states in human cancer: insights from single-cell analysis. *Nat Rev Cancer*. 2020;20:218–32.
  57. Gocher AM, Workman CJ, Vignali DAA. Interferon-gamma: teammate or opponent in the tumour microenvironment? *Nat Rev Immunol*. 2022;22:158–72.
  58. Azimi F, Scolyer RA, Rumcheva P, Moncrieff M, Murali R, McCarthy SW, Saw RP, Thompson JF. Tumor-infiltrating lymphocyte grade is an independent predictor of sentinel lymph node status and survival in patients with cutaneous melanoma. *J Clin Oncol*. 2012;30:2678–83.

59. Jardim DL, Goodman A, de Melo GD, Kurzrock R. The challenges of tumor mutational burden as an immunotherapy biomarker. *Cancer Cell*. 2021;39:154–73.
60. Bhatia R, Gautam SK, Cannon A, Thompson C, Hall BR, Aithal A, Banerjee K, Jain M, Solheim JC, Kumar S, et al. Cancer-associated mucins: role in immune modulation and metastasis. *Cancer Metastasis Rev*. 2019;38:223–36.
61. Snyder A, Makarov V, Merghoub T, Yuan J, Zaretsky JM, Desrichard A, Walsh LA, Postow MA, Wong P, Ho TS, et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. *N Engl J Med*. 2014;371:2189–99.
62. Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, Sucker A, Hillen U, Foppen MHG, Goldinger SM, et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. *Science*. 2015;350:207–11.
63. McGranahan N, Furness AJ, Rosenthal R, Ramskov S, Lyngaa R, Saini SK, Jamal-Hanjani M, Wilson GA, Birkbak NJ, Hiley CT, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science*. 2016;351:1463–9.
64. Delamarre L, Mellman I, Yadav M. Cancer immunotherapy. Neo approaches to cancer vaccines. *Science*. 2015;348:760–1.
65. Falk K, Rötzschke O, Stevanović S, Jung G, Rammensee HG. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature*. 1991;351:290–6.
66. Lundegaard C, Lund O, Nielsen M. Prediction of epitopes using neural network based methods. *J Immunol Methods*. 2011;374:26–34.
67. Chen PJ, Hussmann JA, Yan J, Knipping F, Ravisankar P, Chen PF, Chen C, Nelson JW, Newby GA, Sahin M, et al. Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell*. 2021;184:5635–52.
68. Althenayyan S, AlMuhanna MH, AlAbdulrahman A, Alghanem B, Alsagaby SA, Alfahed A, Alasiri G, Aziz MA. Alternatively spliced isoforms of MUC4 and ADAM12 as biomarkers for colorectal cancer metastasis. *J Pers Med*. 2023;13(1):135–52.
69. Duan F, Duitama J, Al Seesi S, Ayres CM, Corcelli SA, Pawashe AP, Blanchard T, McMahan D, Sidney J, Sette A, et al. Genomic and bioinformatic profiling of mutational neoepitopes reveals new rules to predict anticancer immunogenicity. *J Exp Med*. 2014;211:2231–48.
70. Ghorani E, Rosenthal R, McGranahan N, Reading JL, Lynch M, Peggs KS, Swanton C, Quezada SA. Differential binding affinity of mutated peptides for MHC class I is a predictor of survival in advanced lung cancer and melanoma. *Ann Oncol*. 2018;29:271–9.
71. Balachandran VP, Luksza M, Zhao JN, Makarov V, Moral JA, Remark R, Herbst B, Askan G, Bhanot U, Senbabaoglu Y, et al. Identification of unique neoantigen qualities in long-term survivors of pancreatic cancer. *Nature*. 2017;551:512–6.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.