Retrospective Study

Why MUC16 mutations lead to a better prognosis: A study based on The Cancer Genome Atlas gastric cancer cohort

Yu-Jie Huang, Zhi-Fei Cao, Jie Wang, Jian Yang, Yi-Jun Wei, Yu-Chen Tang, Yin-Xiang Cheng, Jian Zhou, Zi-Xiang Zhang

ORCID number: Yu-Jie Huang 0000-0003-4773-8515; Zhi-Fei Cao 0000-0002-7940-6235; Jie Wang 0000-0001-5621-0857; Jian Yang 0000-0002-5573-2650; Yi-Jun Wei 0000-0003-1711-5683; Yu-Chen Tang 0000-0002-6123-3049; Yin-Xiang Cheng 0000-0002-3052-5688; Jian Zhou 0000-0002-0749-1354; Zi-Xiang Zhang 0000-0001-9204-3873.

Author contributions: Huang YJ and Cao ZF contributed equally to this work; Huang YJ and Cao ZF were responsible for conceptualization, data curation, methodology, and wrote the original draft; Wang J, Yang J, Wei YJ, Tang YC, Cheng YX and Zhou J were responsible for visualization and software; Zhang ZX was responsible for validation, supervision, reviewed and edited the manuscript; all authors approved the final submission.

Supported by National Natural Science Foundation of China, No. 81902985; The Project of Suzhou People's Livelihood Science and Technology, No. SYS2018037 and No. SYS201739; The Six Talent Peaks Project in Jiangsu Province, No. WSW-059; Postgraduate Research & Practice Innovation Program of Jiangsu Province, No. SJCX20_1073; Medical Research

Abstract

BACKGROUND

MUC16, encoding cancer antigen 125, is a frequently mutated gene in gastric cancer. In addition, MUC16 mutations seem to result in a better prognosis in gastric cancer. However, the mechanisms that lead to a better prognosis by MUC16 mutations have not yet been clarified.

AIM

To delve deeper into the underlying mechanisms that explain why MUC16 mutations signal a better prognosis in gastric cancer.

METHODS

We used multi-omics data, including mRNA, simple nucleotide variation, copy number variation and methylation data from The Cancer Genome Atlas, to explore the relationship between MUC16 mutations and prognosis. Cox regression and random survival forest algorithms were applied to search for hub genes. Gene set enrichment analysis was used to elucidate the molecular mechanisms. Single-sample gene set enrichment analysis and “EpiDISH” were used to assess immune cells infiltration, and “ESTIMATE” for analysis of the tumor microenvironment.

RESULTS

Our study found that compared to the wild-type group, the mutation group had a better prognosis. Additional analysis indicated that the MUC16 mutations appear to activate the DNA repair and p53 pathways to act as an anti-tumor agent. We also identified a key gene, NPY1R (neuropeptide Y receptor Y1), which was
significantly more highly expressed in the MUC16 mutations group than in the MUC16 wild-type group. The high expression of NPY1R predicted a poorer prognosis, which was also confirmed in a separate Gene Expression Omnibus cohort. Further susceptibility analysis revealed that NPY1R might be a potential drug target for gastric cancer. Furthermore, in the analysis of the tumor microenvironment, we found that immune cells in the mutation group exhibited higher anti-tumor effects. In addition, the tumor mutation burden and cancer stem cells index were also higher in the mutation group than in the wild-type group.

CONCLUSION
We speculated that the MUC16 mutations might activate the p53 pathway and DNA repair pathway: alternatively, the tumor microenvironment may be involved.

Key Words: Gastric cancer; MUC16 mutation; Cancer antigen 125; Prognosis; The Cancer Genome Atlas; Gene Expression Omnibus

©The Author(s) 2021. Published by Baishideng Publishing Group Inc. All rights reserved.

Core Tip: Our study utilized multi-omics data from The Cancer Genome Atlas, by analyzing these data, we found that the MUC16 mutations may activate the p53 pathway and DNA repair pathway on the one hand, on the other hand, the tumor microenvironment may be involved, with higher tumor killer cells and lower stromal score together building the unique tumor microenvironment of the mutation group.

INTRODUCTION
As the fourth most common cancer worldwide, gastric cancer is a major cause of cancer-related death[1]. Currently, the number of new gastric cancer cases reported globally accounts for 5.7% of all cancer cases, with 8.2% of patients dying as a result[2]. In China, there were approximately 679100 new cases, with 498000 deaths in 2015. The five-year survival rate of patients with gastric cancer is less than 25% due to resistance to chemotherapy drugs and tumor recurrence[3,4].

MUC16 is a type I transmembrane mucin that encodes cancer antigen 125 (CA-125). It consists of a C-terminal domain, a tandem repeat region, and an extracellular N-terminal section, of which CA-125 is part of the tandem repeat domain. In ovarian cancer, CA-125 is used to monitor cancer progression[5]. Previous studies have shown that MUC16 mutations are associated with a longer survival time in patients with gastric cancer, although the reasons for this are not well understood.

Our study aimed to explore in greater depth the underlying mechanisms as to why MUC16 mutations lead to a better prognosis in patients with gastric cancer.

MATERIALS AND METHODS

Acquisition of patient data
We obtained multi-omics data, including mRNA, methylation, and simple nucleotide variation (SNP) data from The Cancer Genome Atlas (TCGA) (https://portal.gdc.Cancer.gov/) and copy number variation (CNV) data from the University of California Santa Cruz (http://xena.ucsc.edu/) for gastric cancer patients, and we also downloaded the matching clinical data. The mRNA data of a separate gastric cancer cohort (GSE62254) and its corresponding clinical data were obtained from the

Peer review report's scientific quality classification

Grade A (Excellent): A
Grade B (Very good): 0
Grade C (Good): 0
Grade D (Fair): 0
Grade E (Poor): E

Institutional review board statement: This study was a bioinformatics study, and the data used was downloaded from a public database, so the study did not require the approval of the ethics committee.

Informed consent statement: This research did not involve any human or animal experiments, and the data used was downloaded from a public database, so the study did not require the informed consent.

Conflict-of-interest statement: None declared.

Data sharing statement: No additional data are available.

Open-Access: This article is an open-access article that was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution NonCommercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works upon this work non-commercially, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/License/s/by-nc/4.0/

Manuscript source: Unsolicited manuscript

Specialty type: Medicine, research and experimental

Country/Territory of origin: China

Peer review report's scientific quality classification

Grade A (Excellent): A
Grade B (Very good): 0
Grade C (Good): 0
Grade D (Fair): 0
Grade E (Poor): E

...
Tumor mutation burden (TMB) is generally defined as the total number of replacement and insertion/deletion (indel) mutations per basic group in the exon coding region of the assessed gene in a tumor cell's genome. In our study, the following formula was used to calculate TMB: the numbers of mutations/length of exons (38 mb).

Cancer stem cell
As research on tumors has progressed, growing numbers of academics have begun to recognize the heterogeneity of tumors and have proposed the cancer stem cell hypothesis. They argue that the tumor cells, even in the same tumor tissue, can be divided into different clusters, one of which is cancer stem cells[6,7]. Malta et al[8] proposed two cancer stem cell indices, mRNAsi and mDNAsi. The former reflects gene expression and the latter reflects epigenetic features, using a one-class logistic regression machine learning algorithm based on transcriptomic and epigenetic feature set information. For this reason, we also used their calculated mRNAsi in our study.

Acquisition of differentially expressed genes and construction of a model
To investigate differentially expressed genes due to MUC16 mutations, we first classified the TCGA cohort into the MUC16 mutation group and the MUC16 wild-type group and then analyzed them using 'edgeR' (an R package), where we defined logFC>1 (logFC<-1) and false discovery rate (FDR) < 0.05 as differentially expressed genes in these two groups. To validate the GEO cohort, the differentially expressed genes obtained in TCGA were taken to intersect with all genes in the GEO cohort. The intersected genes were used in our subsequent model construction.

We used univariate Cox regression to identify prognostically related genes, where we took P < 0.05 as the cutoff value and ranked the importance of the obtained genes using a random survival forest algorithm (nrep = 100, nstep = 5). We defined genes with relative importance > 0.65 as our target genes. The entire random survival forest algorithm was implemented using the two R packages “randomForestSR” and “randomSurvivalForest”.

Estimation of immune cells infiltration
Two different methods were applied to estimate the infiltration of immune cells according to the type of data; for mRNA data, we used single-sample gene set enrichment analysis (ssGSEA) to calculate the infiltration of 28 different immune cells and EpiDISH was used for methylation data, to estimate the infiltration of six different immune cells.

Gene set enrichment analysis
Gene set enrichment analysis (GSEA) was performed to elucidate molecular mechanisms and pathways using two different tools to achieve this. In the case of multiple genes, we used the “clusterProfiler” (an R package). For a single gene, we used javaGSEA v. 4.0 software based on the data from C2 (c2.cp.kegg.v7.1.symbols).

Assessment of immune score, tumor purity, and stromal score
In order to evaluate the immune score, tumor purity, and stromal score, we analyzed the TCGA cohort using ESTIMATE (an R package). The ESTIMATE algorithm was designed by Yoshihara et al[9] to calculate stromal score and immune score in tumor tissues using expression data. In recent years, this algorithm has been used in tumors such as breast cancer and glioblastoma multiforme and has shown unique appeal[10-12]. The Kruskal-Wallis test compared stromal and immune scores in the MUC16 mutation group and the MUC16 wild-type group.

Statistical analysis
Continuous variables between the two groups were compared utilizing the Wilcoxon rank-sum test, and the Kruskal-Wallis test applied to compare more than two categories. Kaplan-Meier survival analysis was performed and tested using the log-rank test. Heatmaps and correlation matrices were plotted using “pheatmap” and “corrplot” packages, respectively, and “maftools” was used to analyze somatic mutation profiles. We completed all statistical analyses using R software (version 3.6.2). All P values are two-sided and P < 0.05 is considered statistically significant.
RESULTS

**MUC16 mutations in relation to prognosis and TMB**

We obtained 437 samples of gastric cancer patients containing mutation data in TCGA and visualized them using maftools. We found that the top ten mutations in the TCGA cohort were TTN, TP53, MUC16, LRP1B, SYNE1, ARID1A, CSMD3, FAT4, FLG, and PCLO (Figure 1A). In the mutation types analysis, we observed that missense mutations were predominant and that C>T was the most common SNP class (Supplementary Figure 1A). Supplementary Figure 1B shows the correlation between the top 20 mutations.

Following further survival analysis of the 389 samples with complete survival data, we noted that only MUC16 ($P = 0.020$) had a strong correlation with prognosis (Figure 1B). The MUC16 mutation group had a better prognosis, and after calculating the TMB, a significantly higher TMB score was found in the MUC16 mutation group than in the wild-type group (Figure 1C).

**MUC16 mutations and the tumor microenvironment**

We obtained three scores of immune score, tumor purity, and stromal score in the TCGA cohort using ESTIMATE (Figure 2A-C). The Wilcoxon test was applied to evaluate the differences between the MUC16 mutation group and the wild-type group. We found that the MUC16 mutation group differed from the wild-type group in the stromal score, in that the wild-type group exhibited a higher stromal score, while there was no significant difference between an immune score and tumor purity. The results showed that the MUC16 mutation group demonstrated a higher cancer stem cell index when we introduced the previously calculated cancer stem cell index (Figure 2D). Thus, we inferred that when the stromal score was too high, and the cancer stem cell index too low, it might impact prognosis.

In addition, we calculated the infiltration of 28 different immune cells using ssGSEA, and the data showed that activated CD4T cells and CD56dimNK cells were more abundant in the mutation group, but effector memory CD4T cells and plasmacytoid dendritic cells were more enriched in the wild-type group compared to the MUC16 wild-type group (Figure 2E).

**Differentially expressed genes and drug sensitivity**

To explore the differentially expressed genes, we combined the mRNA data from TCGA, which we divided into the MUC16 mutation group and the wild-type group and analyzed them using edgeR. We defined logFC>1 (logFC<-1) and FDR < 0.05 as the differentially expressed genes in these two groups and intersected with all genes in the GEO cohort (Figure 3A), and then further screened genes associated with prognosis using univariate Cox, with $P < 0.05$ in univariate Cox analysis as input genes for the random survival forest algorithm. The random survival forest algorithm was used to further define and analyze the data, and the genes with relative importance greater than 0.65 were defined as our target genes. Finally, the gene NPY1R (neuropeptide Y receptor Y1) was obtained (Figure 3B-C).

Next, we attempted to predict possible drug targets for NPY1R by browsing the Genomics of Drug Sensitivity in Cancer (https://www.cancerrxgene.org/) database and downloading the relevant data. We analyzed the relationship between the expression of NPY1R and the IC$_{50}$ (natural log half maximal inhibitory concentration) values of several targeted drugs in gastric cancer cell lines. A positive correlation between NPY1R expression and IC$_{50}$ value implied increased drug resistance in gastric cancer cell lines. Conversely, it implied that the drug could inhibit the expression of NPY1R. Three drugs that showed a strong negative correlation with NPY1R expression were obtained, including WH-4-023, WZ-184, and Roscovitine (Figure 3D).

**Prognosis and GSEA analysis**

We used the TCGA cohort as a training group with the median value of NPY1R as a cutoff value and divided the TCGA cohort into high-risk and low-risk groups. The GEO cohort was used as external validation to test our model. Kaplan-Meier survival analysis showed that high expression of NPY1R led to poorer prognosis, whether in the TCGA cohort or the GEO cohort (Figure 4A and B). Moreover, the frequency of poorer prognosis was higher in the high-risk group (Figure 4C-F).

GSEA was used to elucidate molecular mechanisms and pathways. The Kyoto Encyclopedia of Genes and Genomes enrichment (KEGG) showed that the low-risk group showed more enrichment on DNA repair, while the high-risk group showed little enrichment, whether in the TCGA cohort or the GEO cohort (Figure 4G-H).
Figure 1 MUC16 mutations in relation to prognosis and tumor mutation burden. A: Waterfall plot showing the mutation information for each gene. Each color represents a type of mutation; B: Kaplan-Meier survival analysis was used to assess prognosis, with red representing the mutation group and blue representing the wild-type group; C: Boxplot showing the tumor mutation burden scores in the two groups. TMB: Tumor mutation burden.

NPY1R and immune infiltration
To further explore the relationship between NPY1R and immune infiltration, we used EpiDISH to evaluate the infiltration of six immune cells based on methylation data. The percentage of six immune cells in each sample is shown in the plots (Figure 5A). Next, we divided them into the high-risk group and the low-risk group according to the median value of NPY1R, and the Wilcoxon test was applied to compare whether there was a difference in immune cell infiltration between the two groups. The results were visualized in violin plots.

We found that the contents of B cells and CD4T cells were higher in the high-risk group than in the low-risk group. However, natural killer (NK) cells and neutrophils were more abundant in the low-risk group (Figure 5B). We also plotted the correlation matrix of these six immune cells. In the correlation matrix, we noted that CD4T cells negatively correlated with neutrophils and CD8T cells, and B cells also showed a negative correlation with NK cells (Figure 5C). At the same time we plotted the heatmap to show the contents of these six immune cells between the two groups (Figure 5D).
Figure 2 MUC16 mutations and the tumor microenvironment. A-C: Violin plot showing tumor purity, immune score, and stromal score, respectively; D: Boxplot showing the difference in cancer stem cell index between the two groups; E: The infiltration of 28 immune cells in tumors was assessed using single-sample gene set enrichment analysis, red represents the mutation group and blue represents the wild-type group.

**MUC16 mutations and CNVs**

To explore the relationship between MUC16 mutations and CNVs, we downloaded CNV data. The two groups were divided similarly, according to MUC16 mutation status. Using the χ² test, we assessed whether there was a difference in CNV between the two groups, and we defined $P < 0.05$ as differential CNVs, displayed by the circle plot (Figure 6A). The circle plot illustrates CNVs at chromosomal locations. The relationship between CNV and mRNA expression was further analyzed.
Huang YJ et al. MUC16 mutations in GC prognosis

Figure 3 Differentially expressed genes and drug sensitivity. A: Volcano plot showing differentially expressed genes between the MUC16 mutation group and the MUC16 wild-type group, up-regulated genes are in red, down-regulated genes are in green, and black represents genes that are not differentially expressed; B: Error rate for the data as a function of the classification tree and out-of-bag importance values for the predictors; C: Genes of relative importance greater than 0.65; D: The plot showing the correlation between NPY1R expression and the sensitivity of gastric cancer cell lines to various drugs, with the three most negatively correlated drugs being WH-4-023, WZ-184, and Roscovitine. NPY1R: Neuropeptide Y receptor Y1.

We assessed this using the Kruskal-Wallis test: $P < 0.05$ was defined as the differential. In particular, we found that inhibitor of growth gene 2 (ING2), with its CNV gain and deletion, mRNA expression also changed, and that ING2 is involved in DNA damage repair (Figure 6B and C).

In addition, we undertook GSEA to utilize differential CNVs and used the clusterProfiler to perform enrichment analysis and visualize the results. KEGG pathways showed enrichment in colorectal cancer, B-cell receptor signaling pathways, apoptosis-multi-species, apoptosis, and p53 signaling pathways (Figure 7A). GO analysis seemed to be more focused on the two processes of apoptosis and anti-apoptosis (Figure 7B-D).

DISCUSSION

MUC16 (OMIM606154) is the most frequently mutated gene in gastric cancer and the prognosis of the mutation group was better than that of the wild-type group. However, it is not well understood why the mutation in MUC16 leads to a better prognosis[5]; thus, we explored the possible mechanisms through multi-omics data, including SNP, CNV, mRNA, and methylation data.
We first analyzed the gene mutation landscape in the TCGA cohort, and found that MUC16 had the third-highest mutation frequency, after TTN and TP53, and was dominated by missense mutations. Survival analysis showed that the MUC16 mutation group had a better prognosis than the wild-type group. Moreover, we found higher TMB scores in the mutation group. As a marker for predicting immune checkpoint blockade, TMB has shown unique prognostic value in immunotherapy in several tumors and was included in the National Comprehensive Cancer Network (NCCN) guidelines for non-small cell lung cancer in 2019[13-16]. Studies in melanoma, neuroendocrine cervical cancer, and colorectal cancer have also found that patients...
with high TMB scores have a better prognosis after treatment with immune checkpoint inhibitors (ICI)\cite{15,17-19}. Therefore, we speculated that the MUC16 mutation group is more likely to benefit from ICI treatment.

We also investigated the tumor microenvironment and found no statistically significant differences in the remaining immune score and tumor purity, except for the stromal score, which was higher in the MUC16 wild-type group. Thus, a higher stromal score seems to predict a worse prognosis, and Gong et al\cite{20} also found that a
Huang YJ et al. MUC16 mutations in GC prognosis

Figure 6 MUC16 mutations and copy number variations. A: The circle plot shows the location of differential copy number variations in the chromosome; B and C: The left boxplot shows the relationship between ING2 (inhibitor of growth gene 2) gain and deletion and the mRNA expression, the right boxplot shows the differential expression of ING2 in the MUC16 mutation and wild-type groups. ING2: Inhibitor of growth gene 2.

A higher stromal score was always associated with poor prognosis in patients with melanoma. mRNAsi in the MUC16 mutation and wild-type groups was also compared, and mRNAsi was found to be much higher in the mutation group than in the wild-type group. Previous studies have shown that cancer stem cells appear to be associated with tumor recurrence and drug resistance in bladder cancer; therefore, more in-depth studies are needed on the relationship between mRNAsi and gastric cancer[21].

Additionally, in assessing the infiltration of 28 immune cells, we found that activated CD4T cells and CD56dimNK cells were more abundant in the mutation group, whereas effector memory CD4T cells and plasmacytoid dendritic cells were more abundant in the wild-type group. Several studies have shown that CD4T cells are useful for tumor suppression. Belisle et al[22] found that in ovarian cancer, MUC16 specifically binds to CD56dimNK cells, thereby causing immunosuppression, and a similar situation occurred in our study. It has been suggested that plasmacytoid dendritic cells may contribute to the development of Tregs in a colorectal cancer-
Figure 7 Gene set enrichment analysis. A: Kyoto Encyclopedia of Genes and Genomes enrichment results of the differential copy number variations; B-D: Gene Ontology enrichment results of the differential copy number variations.
resistant environment, thus leading to a poor prognosis[23]. The MUC16 mutation group exhibited higher tumor-killing cells and lower immunosuppression cells; thus, we speculate that the infiltration of immune cells may also be a non-negligible factor contributing to better prognosis in the mutation group than the wild-type group.

In the subsequent analysis, we applied the random survival forest algorithm and Cox regression to investigate the differentially expressed genes associated with prognosis due to MUC16 mutations. After screening, we found NPY1R, which belongs to the G protein-coupled receptor superfamily. It encodes a transmembrane protein that functions as neuropeptide Y (NPY, a neurotransmitter), and peptide YY (PYY, a gastrointestinal hormone). It has been suggested that NPY1R has a unique role in gastric acid secretion and anti-tumorigenesis[24,25].

We noted that high NPY1R expression was always associated with a poor prognosis in the study of NPY1R and breast cancer, where Liu et al[26] found that high NPY1R expression was associated with a poorer prognosis. There, immune infiltration based on methylation data showed that NK cells—neutrophils—were more evident in the low NPY1R expression group, while there were more B cells and CD4T cells in the high expression group. NK cells act as tumor-killing cells, leading to a better prognosis. In addition, neutrophils potentially influence the clinical outcome of gastric cancer by modulating immune response; it has been shown that enriched neutrophils can induce NK cell activation through receptor-ligand interaction and interleukin 18 production[27-29]. It has also been pointed out that CD4T may be involved in breast tumor progression[30] and that B cells may be associated with a better prognosis of the esophagogastric junction[31].

As we discovered that high expression of NPY1R was associated with worse outcomes, we hypothesized that if we could reduce the expression of NPY1R, then we might be able to improve the prognosis; thus, we have identified a drug that has the potential to benefit patients with high NPY1R expression: Roscovitine.

Roscovitine is a selective cyclin-dependent kinase inhibitor. Iseki et al[32] found that it inhibited the growth of gastric cancer cells. Roscovitine has been reported in many articles to increase p53, and p53 also enhances the effects of Roscovitine[32-34]. Interestingly, our GSEA analysis showed that the NPY1R low expression group had more DNA repair pathways and p53 pathways, while it was rarely present in the NPY1R high expression group. A similar result emerged when further KEGG and GO analyses were performed on differential CNVs, which were more involved in the p53 pathway, B-cell response pathway in the MUC16 mutation group, while GO analysis showed that they were more enriched in p53-mediated apoptosis and anti-apoptosis. This may suggest that the MUC16 mutation may activate the p53 pathway and DNA damage repair, leading to a better prognosis. Therefore, we speculate that the use of Roscovitine may be able to increase p53 expression to promote apoptosis in tumor cells, and perhaps also inhibit NPY1R expression, which may be a possible direction for patients with high NPY1R expression. Also, when examining the differences in CNVs between the wild-type group and the mutation group, we found that ING2, with its CNV gain and deletion, mRNA expression also changed. ING2, a member of the ING protein family, contains a highly conserved plant homeodomain and nuclear localization sequences[35,36]. In addition, ING2 is involved in DNA damage repair. It is considered to be a tumor suppressor gene. The expression of ING2 is decreased in several cancers including lung cancer, hepatocellular carcinoma, and breast cancer[37-39]. As a tumor suppressor gene, we found higher expression in the mutation group, suggesting that the high expression of ING2 in the mutation group may indicate stronger DNA damage repair.

Therefore, we speculate that the MUC16 mutation may activate the p53 pathway and DNA repair pathway on the one hand, and the immune microenvironment may also be involved on the other hand, with higher tumor killer cells and fewer tumor suppressor cells jointly constructing the tumor immune microenvironment of the mutation group. Perhaps, it is these characteristics exhibited by the mutation group that makes their prognosis better.

**CONCLUSION**

Using multi-omics data, we investigated possible mechanisms leading to a better prognosis of MUC16 mutations from different aspects, and the results implied that the immune microenvironment and DNA repair are possible underlying mechanisms. More importantly, we also identified a gene—NPY1R—that was significantly differentially expressed in the MUC16 mutation group and the MUC16 wild-type group, and...
also discovered a potential drug — Roscovitine — that inhibits NPY1R expression and may even improve the prognosis of gastric cancer patients.

**ARTICLE HIGHLIGHTS**

**Research background**

*MUC16* is a frequently mutated gene in gastric cancer and the *MUC16* mutations seem to result in a better prognosis in gastric cancer. Unfortunately, the mechanism that leads to a better prognosis with *MUC16* mutations is less clear.

**Research motivation**

Among gastric cancer patients, *MUC16* mutations were associated with better prognosis; however, the mechanism of this is not well understood.

**Research objectives**

To explore in greater depth the underlying mechanisms as to why *MUC16* mutations lead to a better prognosis in gastric cancer patients.

**Research methods**

Based on gastric cancer data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO), which were performed to explore the relationship between *MUC16* mutations and prognosis, Cox regression and random survival forest algorithms were applied to search for hub genes. Gene set enrichment analysis was used to elucidate the molecular mechanisms. Single sample gene set enrichment analysis and EpiDISH were used to assess immune cells infiltration, and ESTIMATE analyzed the tumor microenvironment.

**Research results**

Our study shows that *MUC16* mutations appear to activate the DNA repair and p53 pathways and act as an anti-tumor agent. We also identified a key gene, NPY1R (neuropeptide Y receptor Y1), with high expression of NPY1R predicting a poorer prognosis, which was confirmed in a separate GEO cohort. Further susceptibility analysis also revealed that NPY1R might be a potential drug target for gastric cancer. In analyzing the tumor microenvironment, we found that immune cells in the mutation group exhibited higher anti-tumor effects. In addition, the tumor mutation burden and cancer stem cells index were also higher in the mutation group.

**Research conclusions**

By analyzing the gastric cancer data from TCGA and GEO, we speculate that the *MUC16* mutation may activate the p53 pathway and DNA repair pathway on the one hand. On the other hand, the tumor microenvironment may be involved, with higher tumor killer cells and lower stromal scores together building the unique tumor microenvironment in the mutation group.

**Research perspectives**

In gastric cancer patients, *MUC16* mutations predict a better prognosis.

**REFERENCES**

1. Zhang J, Piao HY, Wang Y, Meng XY, Yang D, Zhao Y, Zheng ZC. To Develop and Validate the Combination of RNA Methylation Regulators for the Prognosis of Patients with Gastric Cancer. *Onco Targets Ther* 2020; 13: 10785-10795 [PMID: 33122917 DOI: 10.2147/OTT.S276239]
2. Chen X, Zhang D, Jiang F, Shen Y, Li X, Hu X, Wei P, Shen X. Prognostic Prediction Using a Stemness Index-Related Signature in a Cohort of Gastric Cancer. *Front Mol Biosci* 2020; 7: 570702 [PMID: 33134315 DOI: 10.3389/fmolb.2020.570702]
3. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ, He J. Cancer statistics in China, 2015. *CA Cancer J Clin* 2016; 66: 115-132 [PMID: 26809342 DOI: 10.3322/caac.21338]
4. Van Cutsem E, Sargent D, Topal B, Haustermans K, Prenen H. Gastric cancer. *Lancet* 2016; 388: 2654-2664 [PMID: 27156933 DOI: 10.1016/S0140-6736(16)30354-3]
5. Li X, Pasche B, Zhang W, Chen K. Association of MUC16 Mutation With Tumor Mutation Load and Outcomes in Patients With Gastric Cancer. *JAMA Oncol* 2018; 4: 1691-1698 [PMID: 30098163 DOI: 10.1001/jamaoncol.2018.2805]
6 Kim C, Gao R, Sei E, Brandt R, Hartman J, Hatschek T, Crosetto N, Foukakis T, Navin NE. Chemo-resistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. Cell 2018; 173: 879-893. e13 [PMID: 29681456 DOI: 10.1016/j.cell.2018.03.041]

7 Karayaz M, Cristea S, Gillespie SM, Patel AP, Mylvaganam R, Luo CC, Specht MC, Bernstein BE, Michor F, Ellisen LW. Unravelling subclonal heterogeneity and aggressive disease states in TNBC through single-cell RNA-seq. Nat Commun 2018; 9: 3588 [PMID: 30181541 DOI: 10.1038/s41467-018-06052-0]

8 Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, Kamińska B, Huelsken J, Omborg L, Gevaert O, Colaprico A, Czerwińska P, Mazurek S, Mishra L, Heyn H, Krasnitz A, Godwin AK, Lazar AJ. Cancer Genome Atlas Research Network, Stuart JM, Hoadley KA, Laird PW, Nossmehr H, Wiznerowicz M. Machine Learning Identifies Stemness Features Associated with Oncogenic Differentiation. Cell 2018; 173: 338-354. e15 [PMID: 29625051 DOI: 10.1016/j.cell.2018.03.034]

9 Yoshihara K, Shmahoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, Mills GB, Verhaak RG. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 2013; 4: 2612 [PMID: 24113773 DOI: 10.1038/ncomms3612]

10 Jia D, Li S, Li D, Xue H, Yang D, Liu Y. Mining TCGA database for genes of prognostic value in glioblastoma microenvironment. Aging (Albany NY) 2018; 10: 592-605 [PMID: 29676997 DOI: 10.18632/aging.101415]

11 Priediezig N, Watters RJ, Lucas PC, Basudan A, Bhargava R, Horne W, Kolls JK, Fang Z, Rosenzweig MQ, Bruksfi AM, Weiss KR, Oesterreich S, Lee AV. Exome-capture RNA sequencing of decade-old breast cancers and matched decalcified bone metastases. JCI Insight 2017; 2 [PMID: 28878133 DOI: 10.1172/jci.insight.95703]

12 Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, Bin J, Liao Y, Rao J, Liao W. Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures. Cancer Immunol Res 2019; 7: 737-750 [PMID: 30842092 DOI: 10.1158/2326-6066.CIR-18-0436]

13 Carbone DP, Reck M, Paz-Ares L, Creeelan B, Horn L, Steins M, Felip E, Van den Heuvel MM, Ciuleanu TE, Badin F, Ready N, Hiltermann TJN, Nair S, Juergens R, Peters S, Minenza E, Wang X, Heiskanen-Di Fabio RJ, Villaruz LC, Havel L, Krejci J, Corral Jaime J, Chang H, Geese WJ, Bhagavatheeswaran P, Chen AC, Socinski MA; CheckMate 026 Investigators. First-Line Nivolumab in Stage IV or Recurrent Non-Small-Cell Lung Cancer. N Engl J Med 2017; 376: 2415-2426 [PMID: 28636851 DOI: 10.1056/NEJMoa1613493]

14 Hellmann MD, Ciuleanu TE, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, Minenza E, Linnardou H, Burgers S, Salman P, Borghaei H, Ramalingam SS, Brahmer J, Reck M, O’Byrne KJ, Geese WJ, Green G, Chang H, Szustakowski J, Bhagavatheeswaran P, Healey D, Fu Y, Nathan F, Paz-Ares L. Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. N Engl J Med 2018; 378: 2093-2104 [PMID: 29658845 DOI: 10.1056/NEJMoa1801946]

15 Chan TA, Wolchok JD, Snyder A. Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. N Engl J Med 2015; 373: 1984 [PMID: 26559922 DOI: 10.1056/NEJMoa1508163]

16 Yarchoan M, Hopkins A, Jaffe EM. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. N Engl J Med 2017; 377: 2500-2501 [PMID: 29262725 DOI: 10.1056/NEJMoa1713444]

17 Van Allen EM, Miao D, Schilling B, Shukla SA, Blank C, Zimmer L, Sucker A, Hillen U, Foppen MHG, Golginger SM, Utkal J, Hassel JC, Weide B, Kaehler KC, Loquai C, Mohr P, Gutmer R, Dammert R, Gabriel S, Wu CJ, Schadendorf D, Garraway LA. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. Science 2015; 350: 207-211 [PMID: 26359337 DOI: 10.1126/science.aad0095]

18 Sharabi A, Kim SS, Kato S, Sanders PD, Patel SP, Sanghvi P, Weihe E, Kurzrock R. Exceptional Response to Nivolumab and Stereotactic Body Radiation Therapy (SBRT) in Neuroendocrine Cervical Carcinoma with High Tumor Mutational Burden: Management Considerations from the Center For Personalized Cancer Therapy at UC San Diego Moores Cancer Center. Oncology 2017; 22: 631-637 [PMID: 28530027 DOI: 10.1634/theoncologist.2016-0517]

19 Domingo E, Camps C, Kaisaki PJ, Parsons MJ, Mouradov D, Makino S, Palmieri M, Ward RB, Hawkins NJ, Gibbs P, Asktudh H, Oukrif D, Wang H, Wood J, Tomlinson E, Bark Y, Kaur K, Johnstone EC, Patels C, Church DN, Novelli M, Danielen HE, Sherlock J, Kerr D, Kerr R, Sieber O, Taylor J, Tomlinson I. Mutation burden and other molecular markers of prognosis in colorectal cancer treated with curative intent: results from the QUASAR 2 clinical trial and an Australian community-based series. Lancet Gastroenterol Hepatol 2018; 3: 635-643 [PMID: 30042065 DOI: 10.1016/S2468-253X(18)30117-1]

20 Gong Q, Wan Q, Li A, Yu Y, Ding X, Lin L, Qi X, Hu L. Development and validation of an immune and stromal prognostic signature in uveal melanoma to guide clinical therapy. Aging (Albany NY) 2020; 12: 20254-20267 [PMID: 33100273 DOI: 10.18632/aging.103779]

21 Aghaeehanki N, Rashitchizadeh N, Shadpour P, Allameh A, Mahmodi M. Cancer stem cells as a therapeutic target in bladder cancer. J Cell Physiol 2019; 234: 3197-3206 [PMID: 30471107 DOI: 10.1002/jcp.26916]

22 Belisle JA, Gubbels JA, Raphael CA, Migneault M, Rancourt C, Connor JP, Patankar MS. Peritoneal natural killer cells from epithelial ovarian cancer patients show an altered phenotype and bind to the tumour marker MUC16 (CA125). Immunology 2007; 122: 418-429 [PMID: 17617155 DOI: 10.1111/j.1365-2567.2007.02860.x]
Huang YJ et al. MUC16 mutations in GC prognosis

10.1111/j.1365-2567.2007.02660.x]

23 Gai XD, Song Y, Li C, Lei YM, Yang B. Potential role of plasmaeytoid dendritic cells for FOXP3+ regulatory T cell development in human colorectal cancer and tumor draining lymph node. Pathol Res Pract 2013; 209: 774-778 [PMID: 24080284 DOI: 10.1016/j.prp.2013.08.011]

24 Kermami M, Eliaissi A. Gastric acid secretion induced by paraventricular nucleus microinjection of orexin A is mediated through activation of neuropeptide Yergic system. Neuroscience 2012; 226: 81-88 [PMID: 22968171 DOI: 10.1016/j.neuroscience.2012.08.052]

25 Lee HM, Udupi V, Englander EW, Rajaraman S, Coffey BJ Jr, Greeley GH Jr. Stimulatory actions of insulin-like growth factor-I and transforming growth factor-alpha on intestinal neutrosenin and peptide YY. Endocrinology 1999; 140: 4065-4069 [PMID: 10465277 DOI: 10.1210/endo.140.9.6969]

26 Liu L, Xu Q, Cheng L, Mu C, Xiao L, Xu D, Gao Y, Wang J, Song H. NPY IR is a novel peripheral blood marker predictive of metastasis and prognosis in breast cancer patients. Oncol Lett 2015; 9: 891-896 [PMID: 25624911 DOI: 10.3892/ol.2014.2721]

27 Herbst RS, Soria JC, Kowanetz M, Fine GD, Hamid O, Gordon MS, Sosman JA, McDermott DF, Powderly JD, Gettinner SN, Kohut HE, Horn L, Lawrence DP, Rost S, Leabman M, Xiao Y, Mokratin A, Kopeen H, Hegde PS, Mellman I, Chen DS, Hodi FS. Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients. Nature 2014; 515: 563-567 [PMID: 25428504 DOI: 10.1038/nature14011]

28 Amano K, Hiroyama M, Azuma E, Iwamoto S, Keida Y, Komada Y. Neutrophils induced licensing of natural killer cells. Mediators Inflamm 2015; 2015: 747680 [PMID: 25977601 DOI: 10.1155/2015/747680]

29 Sporri R, Joller N, Hilbi H, Oxenius A. A novel role for neutrophils as critical activators of NK cells. J Immunol 2008; 181: 7121-7130 [PMID: 18981133 DOI: 10.4049/jimmunol.181.10.7121]

30 Tan W, Zhang W, Strasner A, Grivennikov S, Cheng JQ, Hoffman RM, Karin M. Tumour-infiltrating regulatory T cells stimulate mammary cancer metastasis through RANKL-RANK signalling. Nature 2011; 470: 548-553 [PMID: 21326202 DOI: 10.1038/nature09707]

31 Knief J, Reddemann K, Petrova E, Herhahn T, Wellner U, Thorns C. High Density of Tumor-infiltrating B-Lymphocytes and Plasma Cells Signifies Prolonged Overall Survival in Adenocarcinoma of the Esophagogastric Junction. Anticancer Res 2016; 36: 5339-5345 [PMID: 27798897 DOI: 10.21873/anticanceres.11107]

32 Iseki H, Ko TC, Xue XY, Seapan A, Hellmich MR, Townsend CM Jr. Cyclin-dependent kinase inhibitors block proliferation of human gastric cancer cells. Surgery 1997; 122: 187-94; discussion 194 [PMID: 9288122 DOI: 10.1016/s0039-6060(97)90008-8]

33 Wesierska-Gadek J, Wandl S, Kramer MP, Pickem C, Krystof V, Hajek SB. Roscovitine up-regulates p53 protein and induces apoptosis in human HeLa(S)3 cervix carcinoma cells. J Cell Biochem 2008; 105: 1161-1171 [PMID: 18845603 DOI: 10.1002/jcb.21003]

34 Pappskárová M, Krystof V, Jordá R, Dzubák P, Hajdich M, Wesierska-Gadek J, Strnad M. Functional p53 in cells contributes to the anticancer effect of the cyclin-dependent kinase inhibitor roscovitine. J Cell Biochem 2009; 107: 428-437 [PMID: 19308936 DOI: 10.1002/jcb.22139]

35 He GH, Helbing CC, Wagner MJ, Sensen CW, Riabowol K. Phylogenetic analysis of the ING family of PHD finger proteins. Mol Biol Evol 2005; 22: 104-116 [PMID: 15536280 DOI: 10.1093/molbev/msi256]

36 Ohkouchi C, Kumamoto K, Saito M, Ishigame T, Suzuki SI, Takenoshita S, Harris CC. ING2, a tumor associated gene, enhances PAI1 and HSPA1A expression with HDAC1 and mSin3A through the PHD domain and Cterminal. Mol Med Rep 2017; 16: 7367-7374 [PMID: 28948462 DOI: 10.3892/mmr.2017.7553]

37 Okano T, Gemma A, Hosoya Y, Hosomi Y, Nara M, Kokubo Y, Yoshimura A, Shibuya M, Nagashima M, Harris CC, Kudoh S. Alterations in novel candidate tumor suppressor genes, ING1 and ING2 in human lung cancer. Oncol Rep 2006; 15: 545-549 [PMID: 16465410]

38 Zhang HK, Pan K, Wang H, Weng DS, Song HF, Zhou J, Huang W, Li J, Chen MS, Xia JC. Decreased expression of ING2 gene and its clinicopathological significance in hepatocellular carcinoma. Cancer Lett 2008; 261: 183-192 [PMID: 18160212 DOI: 10.1016/j.canlet.2007.11.019]

39 Walzak AA, Veldhoen N, Feng X, Riabowol K, Helbing CC. Expression profiles of mRNA transcript variants encoding the human inhibitor of growth tumor suppressor gene family in normal and neoplastic tissues. Exp Cell Res 2008; 314: 273-285 [PMID: 17720155 DOI: 10.1016/j.yexcr.2007.07.029]
