87400_Auto_Edited.docx

WORD COUNT	3748	TIME SUBMITTED	20-OCT-2023 09:55AM
		PAPER ID	103603138

Name of Journal: *World Journal of Gastrointestinal Surgery*

Manuscript NO: 87400

Manuscript Type: ORIGINAL ARTICLE

Basic Study

Mutational separation and clinical outcomes of TP53 and CDH1 in gastric cancer

TP53 and CDH1 Mutations in Gastric Cancer

Abstract

BACKGROUND

Gastric cancer (GC) is a deadly tumor with the 5th highest occurrence rate and highest mortality rate globally. Owing to its heterogeneity, the underlying mechanism of GC remains unclear, and chemotherapy offers little benefit to individuals.

AIM

To investigate the clinical outcomes of *TP53* and *CDH1* mutations in GC.

METHODS

In this study, 202 gastric adenocarcinoma tumor tissues and their corresponding normal tissues were collected. Using the target-capture technique, 490 genes were identified. Through test and Wilcoxon rank sum test, somatic mutations, microsatellite instability, and clinical statistics including overall survival were detected, compared, and calculated.

RESULTS

The mutation rates of 32 genes, including *TP53*, *SPEN*, *FAT1*, and *CDH1*, exceeded 10%. *TP53* mutations had a slightly lower overall occurrence rate (33%), similar to the findings of other studies. The *TP53* mutation rate was significantly higher in the advanced stages (stage III/IV) than in the early stages (stage I/II) ($p<0.05$). In contrast,

CDH1 mutations were significantly associated with diffuse GC. TP53 is related to the poor prognosis of advanced-stage tumors; nevertheless, CDH1 corresponds to a diffuse type of cancer. Moreover, *TP53* is exclusively mutated into *CDH1*, which is primarily caused by two distinct GC mechanisms.

CONCLUSION

Different somatic mutation patterns in *TP53* and *CDH1* indicate two major mechanisms of GC.

Key Words: Gastric cancer; TP53 mutation; CDH1 mutation; Clinical outcomes; Somatic mutations; Diffuse gastric cancer

Liu HL, Peng H, Huang CH, Zhou HY, Ge J. Mutational separation and clinical outcomes of TP53 and CDH1 in gastric cancer. *World J Gastrointest Surg* 2023; In press

Core Tip: Mutational separation of TP53 and CDH1 in gastric cancer reveals distinct mechanisms. TP53 mutations are associated with advanced-stage tumors and poor prognosis, while CDH1 mutations are linked to diffuse gastric cancer. The study highlights the heterogeneity of gastric cancer and provides insights into potential targeted therapies based on specific mutation patterns. Understanding the mutational landscape of TP53 and CDH1 can aid in personalized treatment approaches for patients with gastric cancer.

INTRODUCTION

Gastric cancer (GC) is one of the most severe malignancies globally, with the 5th leading incidence rate and the highest mortality rate (1). Global Cancer Statistics in 2018 revealed that GC was the second most prevalent malignant tumor in China, with high morbidity and mortality rates and the 3rd leading cause of cancer-related deaths globally after lung cancer (2,3). Although it remains unclear, the pathogenesis of GC is

caused by several factors, including genetic background and external environment (4). Although the standardized treatment for GC is continually improving, its overall incidence and mortality rates remain high. The poor prognosis of GC patients is attributed to limited therapeutic interventions (5,6). However, detecting hidden symptoms at the early stages remains difficult; hence, most patients are diagnosed at advanced stages (7,8). The current treatment for GC is primarily surgical resection combined with preoperative or postoperative adjuvant chemotherapy or radiochemotherapy (5,9). Chemotherapy remains the primary method for postoperative treatment of advanced GC. D2 gastrectomy is the recommended treatment for GC followed by postoperative adjuvant chemotherapy (10). Nonetheless, the tumor response rate to postoperative chemotherapy is low and patients respond differently to chemotherapy (11,12). This difference in response to chemotherapy among patients occurs because GC is a heterogeneous disease that can manifest as differences in gene levels, biological features, and drug sensitivity (13,14). With the rapid developments in molecular biology, genomics, bioinformatics, and high-throughput next-generation sequencing, studies on the pathogenesis, biological markers, targeted sequencing, and treatment of GC are advancing. Specifically, the advancement of individualized treatments and precision medicine for tumors underscore the need to understand the biological characteristics of GC.

With more in-depth research on the molecular basis of GC, the activation of oncogenes or the inactivation of tumor suppressor genes caused by somatic gene mutations has been shown to modulate the development of malignant tumors. Therefore, identifying potential driver genes and mutations associated with GC is key to understanding the mechanism of GC occurrence and development, as well as to formulating a follow-up treatment scheme. In this study, target-capture sequencing technology was used to sequence 490 genes in 202 cases of gastric adenocarcinoma (GAC) and adjacent tissue samples to detect somatic mutations.

MATERIALS AND METHODS

Samples

This study enrolled 202 patients with GAC comprising 135 males and 67 females who underwent surgery at the Department of Gastrointestinal Surgery, Xiangya Hospital, Central South University, China, between January 1st, 2014, and December 31st, 2015. The primary GAC tumor tissues and matched non-cancerous (NC) tissues located at least 5 cm away from the tumor core were obtained after surgical resection and immediately processed and then stored for subsequent use. Before surgery, none of the recruited patients received chemotherapy or radiotherapy. Histopathological diagnosis was carried out preoperatively and confirmed by surgery based on the World Health Organization Classification of Tumors (15). The tumor stage was defined following the 8th IASLC/ AJCC staging system (16).

Experiments

DNA was extracted from cancer and NC tissues using a customized panel from Roche NimbleGen, Inc. The customized panel included exons and hotspots of 490 genes, with a total length of 1 Mb. An X10 sequencer (Illumina Inc.) was used for sequencing in PE150 mode. The adjacent tissues were ≥5 cm away from the cancerous tissues. None of the specimens were treated with radiotherapy or chemotherapy before surgery. All patients underwent curative resection; 27 IA-stage patients did not receive comprehensive treatment, 22 IB-stage patients received S1 chemotherapy, and patients with stage II and above received SOX chemotherapy. After treatment, the patient was followed up *via* phone calls and internet contact.

Sequencing data analysis

Mapping and somatic mutation calling: Quality control was performed on raw sequencing data using FastQC (17); and the sequences were trimmed for adapters and low-quality bases using Trimmomatic, version 0.38, HEADCROP:3 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36 (18). The trimmed reads were then aligned to GRCh37/hg19 using bwa mem version 0.7.17 (19). PICARD (20) was

used to add read groups and mark duplicates. The Genome Analysis Toolkit (GATK) version 3.8 (21) was used for realignment of the indel area and base quality recalibration. GATK was also used for germline and somatic variant calling with Haplotype Caller and MuTect2, respectively. Variants were annotated using an ANNOVAR (22).

Microsatellite instability (MSI) detection: Five commonly used MSI sites, BAT25, BAT26, NR21, NR24, and MONO27, were used for detection. MSI-high (MSI-H) and MSI-low (MSI-L) were selected if at least two loci or one locus between the cancer and NC tissues was correspondingly unstable. All sites were stable between tissues, and the samples were microsatellite-stable (MSS).

Statistical analysis

All statistical analyses were performed using SPSS software package (version 23.0 (SPSS Inc., Chicago, IL, USA) and R (R Core Team (2018). R: Language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.). Quantitative data are presented as mean ± standard deviation (SD). Pearson's χ^2 test was used to compare the difference among ranked data, whereas the one way analysis of variance test was performed to compare the differences among quantitative data. Survival analyses were performed using the Kaplan–Meier method and compared using the log-rank test. $P<0.05$ indicated statistically significant differences. For survival analysis, overall survival (OS) was defined as the period from the date of pathological diagnosis to the date of death or the date of the last follow-up. The cause of death in this study was aggravation of GC. Written informed consent was obtained from each patient prior to the surgery. This study was approved by the Research Ethics Committee of Central South University, China. All specimens were handled and anonymized according to the ethical and legal guidelines.

RESULTS

Sample and sequencing statistics

The average age of the cohort was 55.53 ± 10.25 years, range: 26–82 years. The tumor diameters were < 5 cm in 158 patients and ≥ 5 cm in 44 patients. There were 39 and 163 cases in the medium-to high-differentiation and low-differentiation groups, respectively. In terms of the TNM stage, 94 and 108 patients were classified as stage I/II and stage III/IV, respectively. Local lymph node metastasis was detected in 127 patients and in 75 patients without metastasis.

In all tumor and NC samples, the average sequencing base was 2.17 Gb and 1.19 Gb, and the mean sequencing depths were 829 and 457 respectively. In the target area, each pair of samples exhibited a mean somatic mutation of 23.1, which is consistent with that in other studies, *e.g.*, The Cancer Genome Atlas had 23.5 mutations in the area (23). Moreover, the types and composition of the mutations were similar to those reported in previous studies (24,25). Among all the mutations, point mutations made up the majority, among which missense mutations accounted for the largest fraction (24). Small indels primarily comprise of frameshift mutations. Among the point mutations, the order of mutation type sorted by proportion was C>T, followed by T>C, and T>G, which is consistent with the results reported in previous studies (25-27). Simultaneously, the ratio of C>T mutations was related to age (28,29). Older patients had a larger ratio of C>T mutations, probably due to somatic methylation and lifespan.

MSI and mutations

Among the 202 samples, 9 MSI-H, 19 MSI-L, and 172 MSS were detected, whereas the MSI states of the rest two samples could not be determined. Two MSS samples filtered out all single nucleotide variants under standard criteria (variation quality: PASS, location: exon, mutation frequency > 0.01); thus, 200 samples were used for mutation-related analysis. The proportion of MSI samples was consistent with that in previous research (30). The prevalence of MSI-H GC in Asians is commonly <10% of all GC cases (31), which is lower than most occurrence rates reported in Western studies.

Tumor mutation burden (TMB) was calculated as the total number of somatic mutations divided by the capture size in Mb. The TMB values in the MSI-H and MSI-L samples were significantly higher than those in the MSS samples (Wilcoxon rank-sum test, both $p<0.01$) (Figure 1). In the MSS and MSI samples, the values of TMB were 19.0 and 55.0, respectively, on average i.e. 52.5 and 56.1 for MSI-H and MSI-L, respectively. Nevertheless, no significant difference was noted in TMB values between the MSI-H and MSI-L groups. The proportion of somatic point mutations in the MSI samples was significantly higher than that in the MSS samples, which is consistent with previous findings (32). The increase in somatic mutations caused by MSI was not statistically significant according to the pathological classification (Lauren classification) or clinical stage (TNM stages).

Nearly all the patients had somatic mutations in the target area. The most commonly mutated gene was *TP53*, as previously discovered; however, the mutation rate of *TP53* was 33%, which is far lower than that reported in other studies (23,26,27,33,34). In total, 32 genes had mutation rates greater than 10%, with 10 of these genes (including *KMT2B*, *SPEN*, and *LRP1B*) having mutations greater than 20% (Figure 2). Among the MSS and MSI samples, 31.8% (54/170) and 42.8% (12/28), respectively, had somatic mutations in *TP53*, whereas the difference was not significant between groups (test $p<0.3$). Moreover, no obvious differences in the ratio of gene mutations were noted in the tumor differentiation levels or pathological types.

After co-analyzing the top 15 somatically mutated genes, *TP53* did not co-mutate with other genes, but exclusively mutated with *CDH1*. TP53 is the most frequently mutated somatic protein in GAC that modulates tumorigenesis. Co-analysis results suggested that TP53 mutations may be a special molecular type that does not interact with other genes during gastric cancer occurrence. In addition, *CDH1* is an important gene related to GAC. Therefore, further investigation is necessary because mutated TP53 and CDH1 may indicate two distinct patterns in the pathogenesis of GAC. Additionally, all genes, including *FAT1*, *MGA*, and *ZFHX3*, were co-mutated, except for

TP53 and *CDH1*. In addition, , mutations in *TP53* and *CDH1* may present another pattern.

Driver genes

Driver genes were identified using OncodriveCLUST in maftools (Mayakonda A, Lin DC, Assenov Y, Plass C and Koeffler HP. 2018. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Research. PMID: 30341162), followed by strict additional filtering criteria to focus on the top driver genes (Figure 4). A total of 59 genes (FDR < 0.05) were detected, whereas only 7 genes, including *KMT2B*, *SPEN*, *FAT1*, *MGA*, *MED12*, *KIF1B*, and *ERBB2*, overlapped with the top 20 somatically mutated genes. Most of the top driver genes, including *HOXB13*, *AKT3*, *CHEK1*, *FGFR3*, and *CALR*, were not included in the list of the top somatically mutated genes. Meanwhile, the top mutated genes including *TP53* and *CDH1* were not shown to be important in the driver gene clusters because more regional clusters were identified based on the position of gene mutations. Low HOXB13 expression is responsible for poor tumor differentiation, metastasis (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6004642/), and poor prognosis in GC (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6004642/). *AKT3*, which has a somatic mutation frequency of 7.5%, has been identified as an important driver gene. *AKT3* is an isoserine/threonine protein kinase that regulates TP53 activity through acetylation. *CHEK1* (also called CHK1), a gene involved in the DNA damage checkpoint pathway, cooperates with MMR deficiency to trigger chromosomal instability in MMR-deficient colorectal cancer cells (https://europepmc.org/article/pmc/2735479). Among the 8 samples with CHEK1 mutations, the numbers of MSI-H, MSI-L, and MSS were 2, 2, 4, which were significantly different from the whole cohort (test $p < 0.05$). High CALR expression was observed in 20 of 30 patients with GC and was responsible for positive serosal invasion, lymph node metastasis, perineural invasion, and poor survival in GC (https://pubmed.ncbi.nlm.nih.gov/29441976/) (https://pubmed.ncbi.nlm.nih.gov/190

50968/), and is a good biomarker of prognosis in gastric cancer (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6775705/).

Survival analysis

The OS of MSS patients was not significantly better than that of MSI patients (P = 0.215), whereas in patients with early GC (pT1), the OS of MSS was significantly greater than that of MSI (P = 0.034). The non-significant difference in OS between patients with MSI and MSS conflicts with the one reported in other reports. Several GC studies have suggested a positive relationship between MSI-H phenotype, mismatch repair deficiency (MMRD), and better prognosis (35-37). Although the statistical outcome was not significant, survival analysis of 27 stage IA patients suggested that patients with MSI had a better prognosis than those with MSS, whereas the MSS group had a better prognosis in patients at stage IB. Retrospective Asian studies have confirmed the hypothesis that patients with MSS benefit from adjuvant 5-fluorouracil-based chemotherapy, whereas patients with MSI-H stage II or III GC do not (38-41). Postoperative adjuvant chemotherapy was administered to patients at stage IB and advanced stages; therefore, we speculated that adjuvant chemotherapy contributed to the differences in OS between patients with MSS and MSI at different stages.

Survival analysis revealed that somatic mutations in *TP53* were significantly correlated with a 5-year OS rate of 52.34% and median survival time of 60.00 mo. Specifically, samples with *TP53* somatic mutations had a significantly lower 5-year OS rate than those without *TP53* somatic mutations (39% vs. 58%, p=0.01; Figure 5a). The samples were successfully classified based on the pathological stage, and TP53 did not affect the OS rate in the early (I/II) or middle-late (III/IV) stages (Figure 5b and 5c). Thus, the *TP53* mutation ratio was not associated with the pathological type, and the decrease in OS rate was associated with high mutation rate of *TP53* in middle-late cases. The mean OS rate of samples with *TP53* mutations was lower in diffuse GC than in those without *TP53* mutations (30.87 vs. 37.49, Wilcoxon P = 0.064). In contrast, *CDH1* somatic mutation was not significantly associated with OS (63.83% vs. 49.36%, p=0.33,

Figure 5d). Among the patients without *TP53* mutations, those with *CDH1* mutations seemed to have higher survival rates; however, the difference was not significant. Moreover, those with both *TP53* and *CDH1* mutations had the worst 5-year OS rates ($P = 0.02$; Figure 5e). Nine patients had both *TP53* and *CDH1* mutations, of which one case was a highly differentiated intestinal type, whereas the other eight cases had Lauren diffuse GC.

Further investigation of poorly differentiated diffuse cases led to the identification of 38 *TP53* and 36 *CDH1* mutations. Low-differentiation and diffuse GC had a higher percentage of *CDH1* somatic mutations, which was significantly different from other types of gastric cancer. A survival analysis revealed an overall poor prognosis in patients with poorly differentiated diffuse gastric cancer, and the presence of *P53* mutations resulted in the worst prognosis, whereas that of CDH1 had no significant effect.

DISCUSSION

This study analyzed surgical cases between January 2014 and December 2015 at Xiangya Hospital and fully evaluated the survival time after treatment. Table 1 shows the pathological types, clinical stages, and other clinical statistics of the 202 patients. These results are similar to those of other epidemiological studies of GC in China. However, there is a lack of large-scale molecular genetics research on GC in China because most of the existing studies have obtained specimens from Europe, America, Korea, and Vietnam. Our study revealed approximately 24.1 somatic mutations on average in the 1M capture area of GC samples, corroborating previous studies (23,26,27,33,34). MSI is associated with a number of somatic mutations, but may not be directly related to the pathological type and prognosis. We found that the mutation frequencies of 32 genes, including *TP53*, *KMT2B*, *SPEN*, *FAT1*, and *CDH1*, in GC exceeded 10%. These results are inconsistent with those of previous studies. Our study provides molecular genetics data on Chinese patients with GC.

Our analysis of driver genes differed from previous studies in that *TP53* and *CDH1* were not identified as important driver genes. Driver genes are typically those whose mutations increase the net cell growth under specific microenvironmental conditions in cells *in vivo*. *TP53* and *CDH1* did not co-mutate with other genes, showing unique biological features in our GC samples.

Other studies suggested that *TP53*, a crucial gene associated with GC development, has a somatic mutation rate of approximately 50% (27,33,34). In this work, the overall mutation rate of *TP53* was 33%; however, it was significantly higher in stage III/IV(41%) than stage I /II (23%, χ^2 test $p<0.01$). The mutation type of *TP53* was consistent with that reported in previous studies (25-27), mainly single nucleotide mutations such as C/T and T/C. Mutations located in exon 5 accounted for approximately 36% of all *TP53* mutations, consistent with other studies (23,42), hence the overall lower *TP53* mutation rate may be due to the composition of the clinical samples. Somatic mutations occurred exclusively in *TP53* and *CDH1*. *TP53* did not co-mutate with the other genes. Therefore, *TP53* mutations might represent a unique type of GC tumorigenesis. Notably, the number and probability of lymph node metastases in patients with P53 mutations have increased, which may promote a high incidence of *TP53* mutations in stage III/IV GC. Consequently, *TP53* mutations can modulate lymph node metastasis.

CDH1 is closely associated with GC and is a causative gene of diffuse hereditary lung cancer. In diffuse GC, the *CDH1* mutation rate (25%) was higher than non-diffuse type (11%, χ^2 test $p<0.05$), consistent with previous findings (43). In contrast to *TP53* mutations, *CDH1* mutations play a significant role in the development of diffuse GC via distinct mechanisms. As there is a clear connection between *TP53* and tumor development, tumors with poor prognoses may be more likely to be in advanced stages. Although it has no direct effect on prognosis, *CDH1* mutation is a significant contributor to diffuse GC. *TP53* and *CDH1* mutations may indicate two different types of GC at the molecular level, which warrants further study.

CONCLUSION

GC is the second leading malignant tumor in China, and its incidence is much higher than that in Western countries. This study provides crucial molecular data on GC among Chinese patients because large-scale Chinese genetic evidence is lacking. This study revealed that *TP53* and *CDH1* mutations represent two important pathways for GC occurrence and development. The pathogenesis of GC in Han Chinese people (in the middle and lower reaches of the Yangtze River), as well as the diagnosis and treatment of GC, would benefit from our findings.

ARTICLE HIGHLIGHTS

Research background

GC is the second leading malignant tumor in China, and its incidence is much higher than that in Western countries. This study provides crucial molecular data on GC among Chinese patients because large-scale Chinese genetic evidence is lacking. This study revealed that *TP53* and *CDH1* mutations represent two important pathways for GC occurrence and development. The pathogenesis of GC in Han Chinese people (in the middle and lower reaches of the Yangtze River), as well as the diagnosis and treatment of GC, would benefit from our findings.

Research motivation

One of the challenges to the design of effective treatments for gstric cancer is heterogeneity , its poses an obstacle for the uniform therapy plan irrespective of specific subtypes of tumors in clinical practice.

Research objectives

TP53 and CDH1 have been reported closely related to gastric cancer.we aim to investigate the clinical outcomes of *TP53* and *CDH1* mutations in gastric cancer.

Research methods

202 primary gastirc cancer tissues and matched non-cancerous tissues were sampled *via* surgery.After DNA extraction for cancer tissue and non-cancerous tissue, DNA was captured using customized panel from Roche NimbleGen Inc. The customized panel includes exon and hotspot of 490 genes in total length of 1 Mb. X10 sequencer from Illumina Inc.

Research results

The mutation rate of 32 genes exceeded 10% including TP53, SPEN, FAT1 and CDH1 *etc.* We found that TP53 mutation had a slightly lower overall occurrence rate (33%) while the mutation type was similar to other studies. TP53 mutation rate was significantly higher in advanced stage (III/IV) than early stage (I /II) ($p<0.05$). On the other hand, we also found that CDH1 mutation is significantly related to diffuse GC. TP53 was related to the poor prognosis of advanced stage of tumor, nevertheless, CDH1 was corresponding to diffuse type of cancer. Moreover, TP53 was only exclusively mutated with CDH1, which implies the major reason of two different GC mechanisms.

Research conclusions

Different somatic mutation patterns of TP53 and CDH1 indicated the two major mechanisms of GC.

Research perspectives

Understanding the mutational landscape of TP53 and CDH1 would positively affect the pathogenesis study of GC in Chinese Han people (in the middle and lower reaches of the Yangtze River), as well as guiding the diagnosis and treatment of GC.

ACKNOWLEDGEMENTS

We are sincerely grateful to Yinming Han and Zhiliang Fu for their help with the data sequencing.

87400_Auto_Edited.docx

ORIGINALITY REPORT

4%
SIMILARITY INDEX

PRIMARY SOURCES

1	github.com Internet	25 words — 1%
2	Yukun Cheng, Jian Li, Fangjie Yao, Li Long et al. "Dissection of loci conferring resistance to stripe rust in Chinese wheat landraces from the middle and lower reaches of the Yangtze River via genome-wide association study", Plant Science, 2019 Crossref	24 words — 1%
3	Margherita Ratti, Andrea Lampis, Jens C. Hahne, Rodolfo Passalacqua, Nicola Valeri. "Microsatellite instability in gastric cancer: molecular bases, clinical perspectives, and new treatment approaches", Cellular and Molecular Life Sciences, 2018 Crossref	20 words — 1%
4	link.springer.com Internet	17 words — < 1%
5	Suchir Misra, Andrew Shih, Xiao-Jie Yan, Wentian Li. "A Higher Proportion of Craniosynostosis Genes Are Cancer Driver Genes", Cold Spring Harbor Laboratory, 2019 Crossref Posted Content	13 words — < 1%
6	www.frontiersin.org Internet	13 words — < 1%

7	www.ncbi.nlm.nih.gov Internet	13 words — < 1%
8	Annalisa Di Cello, Maddalena Di Sanzo, Francesca Marta Perrone, Gianluca Santamaria et al. "DJ-1 is a reliable serum biomarker for discriminating high-risk endometrial cancer", Tumor Biology, 2017 Crossref	12 words — < 1%
9	Sandra Prescher, Sebastian Winkler, Leonhard Riehle, Meike Hiddemann et al. "Patient reported experience and adherence to remote patient management in chronic heart failure patients: a posthoc analysis of the TIM-HF2 trial", European Journal of Cardiovascular Nursing, 2022 Crossref	12 words — < 1%
10	quieora.ink Internet	12 words — < 1%
11	www.biorxiv.org Internet	12 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 12 WORDS
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 12 WORDS