Risk of gastric cancer is associated with the *MUC1* 568 A/G polymorphism

QIAN XU[1], YUAN YUAN[1], LI-PING SUN[1], YUE-HUA GONG[1], YING XU[1], XIU-WEN YU[1], NAN-NAN DONG[1], G. DAVID LIN[2], PAUL N. SMITH[3] and RACHEL W. LI[4]

[1]Cancer Control Laboratory of Cancer Institute and General Surgery, The First Affiliated Hospital of China Medical University, Shenyang 110001, P.R. China; [2]Faculty of Health, University of Canberra, ACT 2601; [3]Department of Surgery, The Canberra Hospital, ACT 2606; [4]The Medical School, The Australian National University, ACT 2601, Australia

Received June 25, 2009; Accepted August 5, 2009

DOI: 10.3892/ijo_00000449

Abstract. Identifying the genetic variants that alter MUC1 protein expression may further our understanding of the risk for development of gastric cancer (GC). We used PCR-SSPs to identify the genotype of *MUC1* A/G polymorphism at its 568 site of exon 2 and immunohistochemistry to detect MUC1 protein expression in GC patients and non-cancer subjects and analyzed the association between this polymorphism and MUC1 protein expression. We found that the frequency of AA genotype was significantly high in the GC patients and the risk for GC in AA genotype carriers increased 1.81-fold. Moreover, we found a significant underexpression of MUC1 protein in GC as compared to non-cancer subjects, which was negatively correlated to AA genotype of *MUC1* (r=-0.1790, P=0.004). Furthermore, this study provides a possible mechanistic insight that the *MUC1* A/G polymorphism at its 568 site disrupts the physiological functions of MUC1 which is important to the physiological protection of gastric mucosa. Thus we have provided evidence that may identify the *MUC1* A/G polymorphism at 568 site, as a potential genetic factor which leads to an increase in susceptibility for GC through alteration of *MUC1* gene and MUC1 expression in the population that carry the A allele.

Correspondence to: Dr Yuan Yuan, Cancer Control Laboratory of Cancer Institute and General Surgery, the First Affiliated Hospital of China Medical University, Heping District, Nanjing North Street 155#, Shenyang 110001, P.R. China
E-mail: yyuan@mail.cmu.edu.cn

Abbreviations: MUC1, mucin 1; SNP, single nucleotide polymorphism; GC, gastric cancer; GS, superficial gastritis; GA, atrophic gastritis; PCR-SSPs, sequence-specific primers-polymerase chain reaction; VNTR, variable number tandem tepeats; OR, odds ratio; CI, confidence interval; PBS, phosphate-buffered saline

Key words: gastric carcinoma, mucin 1, mucins, single nucleotide polymorphism, protein expression

Introduction

A number of factors, including hereditary, environmental, occupation and social factors are now recognized as potential contributors to the development of gastric cancer (GC). The contributions of bacterial factors to the disease pathogenesis have been illustrated by the results from our previous study and others, in which most evidence suggests that host factors are paramount in determining progression to GC (1). Cancer susceptibility represents a continuum of interactions between the host and environment. The risk of developing GC is increased by up to 3-fold in individuals with an immediate blood relative with gastric cancer and 10% of cases show familial clustering (2). Susceptibility to *Helicobacter pylori* and to GC appears to be associated to with MUC1 allele length (3,4). Gene polymorphism has been reported to influence or modify individual sensitivity towards micro and macro environmental factors. However, the contribution of genetic factors, such as gene polymorphism, to GC risk is not fully understood.

Secreted mucins are considered the first line of defense for epithelial tissues since they act as physical barriers between the extracellular milieu and the mucosal surface. Tethered transmembrane mucins are the second line of defense, acting as sensors of the different disturbances occurring environmentally and signaling these messages to the inner cytoplasmic milieu (5). MUC1 is a highly polymorphic membrane-associated mucin that is often aberrantly expressed in cancer (6). It possesses a centrally located tandem repeat (TR) domain (7-10) comprised of 20-120 or more repeat units of 60 nucleotides, which encode 20 amino acids. The repeating units include several serine and threonine residues, which carry most of the glycosylation and this glycosylation, as well as the general pattern of expression, is altered in cancerous cells. *MUC1* gene is a member of the mucin family and encodes a membrane-bound, glycosylated phosphoprotein. The protein is anchored to the apical surface of many epithelia by a transmembrane domain, with the degree of glycosylation varying with cell type. It also contains a 20 amino acid variable number tandem repeat (VNTR) domain, with the number of repeats varying from 20 to 120 in different individuals. MUC1

also functions in a cell signaling capacity (11,12). Due to its large conformation, MUC1 has an anti-adhesive role (13), which may block cell-cell interactions by destabilizing cell-cell and cell-matrix connections. MUC1 restricts access of *H. pylori* to the epithelial surface, hence reducing exposure of the host to proinflammatory bacterial products (3).

However, it was reported that overexpression, aberrant intracellular localization, and changes in glycosylation of this protein were associated with carcinomas (14). MUC1 was significantly overexpressed in prostate cancer with poor clinical outcomes after radical surgery. Strawbridge reported that genetic variation in *MUC1* altered prostate cancer risk and progression (15) with identifying five haplotype-tagging single nucleotide polymorphisms (SNP) that describe inherited genetic variation in and around MUC1. The fact that susceptibility to GC appears to be associated with *MUC1* allele length has also been reported (16-18). Clinically, the protective function of mucosa varies between individuals. The observations that MUC1 plays a role in the progression to GC highlight the importance of understanding all the aspects of the normal variation of this gene. It has been reported that MUC1 possesses two variations of the length and structure resulting from *MUC1* A/G SNP at 568 site (10,19). Association between MUC1 and gastricintestinal disease has been attracting the attention of more researchers recently. However, whether the incidence of GC is related to the protective functional variation of MUC1 or the protective functional variation of MUC1 is a consequence of the variation of *MUC1* in length and structure remains unknown. The identification of thousands of single nucleotide polymorphisms (SNPs) in the human genome has prompted a rise in population-based research studies designed to link gene-specific SNPs to complex disease states, pharmacogenetic applications, and individual predisposition and susceptibility to diseases (20). We previously found that the distribution frequency of AA genotype was statistically higher in the GC group than AG, GG genotypes in a population from North-Eastern China (1). In the present study, we hypothesized that the association between 568 site A/G polymorphism in *MUC1* and the function of MUC1 protein was influenced by *MUC1* +568 A/G SNP, which may be responsible for individuals' susceptibility to GC. If such an effect could be demonstrated, it would provide useful information for better understanding the role of *MUC1* polymorphism and MUC1 protein in the modification of individuals' (or a population's) potential risk towards GC and provide support for the design of a MUC1-based tumor vaccine and pharmaceuticals. In this study, we investigated the distribution of *MUC1* A/G polymorphism at 568 site and the association between the A/G polymorphism and susceptibility to GC from the North-Eastern Chinese population. We analyzed the effect of the A/G polymorphism on MUC1 protein expression with the aim of discovering the potential genetic risk factors of GC.

Materials and methods

Patients. This research project was approved by the Ethics Committee of the China Medical University. Gastric tissue specimens and venous blood samples were prospectively collected from patients who underwent gastroscopy in hospitals located in the North-Eastern area of China between 2002 and 2005. All of the samples were collected with letters of consent from the patients and medical histories were taken by questionnaire and the record was computerized. A total of 379 patients from the Shenyang city and Zhuang River region where there was a high incidence of GC were included in this study. Blood samples were taken from 138 patients with GC with an average age of 57.6±11.5 years ranging from 30 to 84 years old. The male/female ratio of the GC patients was 2:1. One hundred and thirty-one patients with superficial gastritis (GS) and 110 patients with atrophic gastritis (GA) were used as control groups. The control population groups had a similar male/female ratio to the GC group and had an average age of 56.6±12.0 years ranging from 30 to 80 years old. The GC group and control group had no statistical difference in terms of gender and age composition (P=0.75 and P=0.43, respectively). Whole blood from individuals was collected and blood clots were allowed to form by incubating clot-activating tubes at room temperature for 60 min. Serum was separated from the clots by centrifugation. Each clot was transferred to a 2 ml centrifuge tube and stored at -80˚C until DNA extraction. Both blood samples and gastric biopsies were analyzed in all cases. The biopsy specimens from the gastroscopies were paraffin-embedded and stained by H&E staining for histological diagnosis. Samples that had an unclear or mixed histology and samples from patients that had received immunoregulatory therapy were excluded from further immunohistochemical analysis. Thus, we prioritized samples from 262 (of 379) patients, including 71 patients with GC (47 male and 24 female, male: female ratio 2:1) at an average age of 59.6±10.6, ranging from 34 to 80 years old; 105 patients with GS; and 86 patients with GA, which served as control groups. The control groups had a similar male/female ratio to the GC group and were at an average age of 55.7±12.4 years, ranging from 30 to 80 years old. The GC group and control groups showed no statistical difference in terms of gender composition (P=0.97); but a slight difference in age composition (P=0.028). Therefore we performed an age correction and gender OR value when analyzing the statistical results.

Genomic DNA from clotted blood for MUC1 single neucleotide polymorphism genotyping. Genomic DNA was purified using a method described (21,22) with some modifications. Briefly, each frozen clot (500 μl) was thawed rapidly at room temperature and placed on ice. The thawed clots were transferred to a centrifuge tube with 800 μl of TE buffer (triethanolamide), mixed well and centrifuged at 10,000 x g for 5 min to disperse the clots. Following clot disruption, 400 μl of TE, 25 μl of 10% SDS and 5 μl of 20 mg/ml proteinase K were added to the residual clot material and incubated overnight. The supernatant was collected and an equal volume of phenol was added. The tube was vortexed for 5 sec to mix and placed on a rotator for 15 min to extract DNA. After rotation, the tube was vortexed for 3 sec and centrifuged at 10,000 x g for 15 min. The supernatant was poured off and the second extraction was performed with the addition of an equal volume of a mixture of phenol and chloroform (1:1) to the tube to further extract DNA. Following centrifugation, the supernatant was poured off and the third extraction was performed with the addition of an equal volume of chloroform.

Following centrifugation, the supernatant was absorbed and two volumes of protein precipitation solution (two volumes of absolute ethanol containing 10% 3 M sodium acetate) were added and incubated for 1 h at -20˚C. Each sample was vortexed at high speed for 20 sec and then centrifuged at 10,000 x g for 10 min. After centrifugation, the resulting DNA pellet was rinsed with 75% ethanol and centrifuged at 10,000 x g for 5 min. The 75% ethanol was decanted and the tube inverted on clean absorbent paper for 10 sec. The resulting DNA was reconstituted in a TE buffer and stored at -20˚C until use.

PCR-SSPs for MUC1 568 A/G SNP polymorphism determination. Genomic DNA was amplified using sequence-specific primers - polymerase chain reaction (PCR-SSPs). The primers and reagents were purchased from Takara Inc., Japan. The forward primer 5'-CTA TGG GCA GAG AGA AGG AG-3' (primer 1), the reverse SSPs 5'-AGC TTG CAT GAC CAG AAC CC-3' (primer 2) and 5'-AGC TTG CAT GAC CAG AAC CT-3 ' (primer 3) were used in combination with the consensus forward primer, i.e. primer 1, leading to expected PCR product sizes of 233 bp (23). The total PCR reaction volume was 26 μl containing 2.6 μl of 10X PCR buffer, 1.024 μl of 2.5 mM dNTP mixture, 0.8 μl of each upstream and downstream primer at 10 pmol/μl, 0.6 U of rTaq DNA polymerase, 10 ng of the template DNA and an appropriate amount of ddH_2O. PCR reaction was performed by pre-denaturing at 95˚C for 1 min, 5 cycles with 95˚C for 25 sec, 70˚C for 25 sec, 72˚C for 30 sec, then 21 cycles with 95˚C for 35 sec, 65˚C for 30 sec, 72˚C for 30 sec, and then, 4 cycles 95˚C for 50 sec, 55˚C for 60 sec, 72˚C for 90 sec (23,24). Twelve μl of PCR product in 2% agarose gel was electrophoresized at 150 V for 30 min separation, then stained by ethidium bromide for 5 min and observed. PCR products on the two different alleles (A allele and G allele) of *MUC1* were commercially sequenced using an ABI377 DNA Sequencer (ABI, USA).

Tissue array-based immunohistochemical staining for detection of MUC1 protein expression in situ. Immunohistochemical analysis was performed in 5-μm-thick sections from sequentially sliced samples of formalin-fixed and paraffin-embedded specimens according to the method described (25) with slight modification. Briefly, tissue sections were deparaffinized and rehydrated. Endogenous peroxidase was blocked using 3% hydrogen peroxide in methanol for 10 min and then the sections were washed with phosphate-buffered saline (PBS), pH 7.4. The sections were incubated with non-immunized horse serum for 20 min at room temperature and washed before being incubated with a specific antibody, Human Milk Fat Global-1, HMFG-1, against MUC1 (1:200 dilution) purchased from Neomarkers Inc. Fremont, USA overnight at 4˚C. Then the sections were washed and incubated with biotinylated secondary antibodies (goat anti-mouse antibody, Maixin Inc., Fujian, China) and streptavidin-biotin peroxidase. After three washes with PBS, the sections were visualized with 3,3'-diaminobenzidine tetra-hydrochloride and counterstained with haematoxylin. Primary antibodies were replaced with PBS buffer as a negative control. Immunohistochemical results were judged by HSCORE (histological score) (26). The HSCORE was calculated using two indices of proportion (P*i*) and intensity (*i*). The proportion (P*i*) was estimated after taking into account the percentage of positive cells. The intensity (*i*) was judged as 0 (no staining), or 1+ (light brown staining), or 2+ (brown staining), or 3+ (heavy brown staining). The MUC1 HSCORE was derived by summing the proportion of cells staining at each intensity multiplied by the intensity of staining.

$$\text{HSCORE} = \Sigma Pi \times i$$

Where i= 0, 1, 2, 3 and Pi varies from 0.0 to 1.0, MUC1 HSCOREs ranged from a minimum of zero in cases with no staining to a maximum of 3.0 in cases in which all the cells stained with maximal intensity. We judged HSCORE >0.0 as positive while HSCORE = 0.0 as negative, and HSCORE ≥2.0 as highly positive. The MUC1 HSCORE was determined by two independent observers.

Statistics. All statistical analysis was performed using the SPSS (13.0) statistical software program (SPSS, Chicago, USA). χ^2 test was used to determinate the significant difference in genotype distribution between the GC and the control populations. Non-conditional logistic regression adjusted by age and gender was used to estimate odds ratios and 95% confidence intervals. χ^2 test and Fisher's exact probabilities were used to compare the positive and highly positive rate of MUC1 protein expression. Pearson correlation coefficient was used to test the correlation between the genotype and MUC1 protein expression.

Results

The fragments type of MUC1 gene A/G polymorphism at 568 site. Previously, we demonstrated an obviously increased risk of developing GC for those who carried both AG+GG genotype and *H. pylori*-IgG-positive, both AA genotype and *H. pylori*-IgG-negative, and both AA genotype and *H. pylori*-IgG-positive, when compared with both AG+GG genotype and *H. pylori*-IgG-negative, in a population from the North-Eastern region of China (1). This finding prompted us to examine *MUC1* A/G polymorphism at 568 site in a population from the same region. As indicated in Fig. 1A, we found that the amplified size of *MUC1* was 233 bp. According to the presence of 233 bp band, the test samples were divided into AA, AG, and GG genotypes. We further sequenced genomic DNA of two alleles to confirm the PCR products (Fig. 1B and C). To test whether this population was a representative group for the frequencies of *MUC1* A/G genotype, a Hardy-Weinberg genetic equilibrium law was performed for the genotype frequencies of the GC and the control groups. The results showed there was no significant difference between the actual number and the theoretical number using χ^2 test (P=0.27, P=0.26, respectively). Since the distribution frequencies of the three genotypes were consistent with the Hardy-Weinberg genetic equilibrium law, it demonstrated that this study population was a representative group.

The distribution of MUC1 A/G polymorphism is associated with the risk of GC. We next sought to identify a potential

Table I. Distribution frequencies of SNP genotype at *MUC1* gene 568 site.

Groups	MUC1 genotype		χ^2	P	OR[a] (95% CI)
	AG+GG (%)	AA (%)			
Control	61 (25.3)	180 (74.7)	1.00		
GC	22 (15.9)	116 (84.1)	4.19	0.031	1.81 (1.06-3.12)

[a]OR value was adjusted by age and gender.

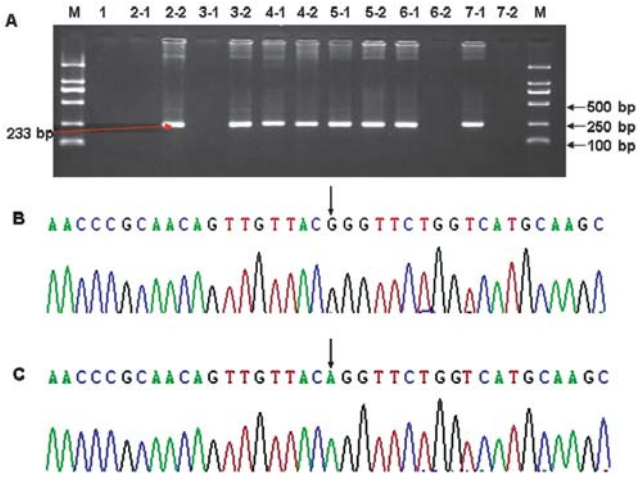

Figure 1. Analysis of *MUC1* genotypes. (A) Identification of *MUC1* genotypes using PCR-SSPs. M, DNA marker DL2000; 1, negative control; 2-1, 3-1, 4-1, 5-1, 6-1, 7-1 were added to primer 1, 2; 2-2, 3-2, 4-2, 5-2, 6-2, 7-2 were added to primer 1, 3; 2 and 3: AA homozygous; 4 and 5: AG heterozygote; 6, 7: GG homozygous. (B and C) DNA sequencing analysis of different allelic PCR products containing *MUC1* A/G polymorphism. The arrow indicates the A/G polymorphism at 568 site.

correlation between *MUC1* A/G polymorphism at 568 site and the risk of GC. We compared the distribution frequencies of AA, AG, and GG genotypes in the *MUC1* gene between GC subjects and control subjects. The distribution frequencies of AA, AG, GG genotypes in the control groups were 74.7% (180/241), 21.2% (51/241) and 4.1% (10/241) respectively. The frequency of allele A was 85.3% (205.5/241) and allele G was 14.7% (35.5/241), which was in agreement with the frequencies of the Human Genome Project (http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=4072037). In the GC groups, we found that the distribution frequencies of AA, AG, GG genotype were 84.1% (116/138), 13.0% (18/138) and 2.9% (4/138). Interestingly, we noted that the frequency of A allele was 90.6% (125/138) and the frequency of G allele was 9.4% (13/138) in GC subjects. The results suggested that the distribution frequency of AA genotype in the GC group was significantly higher than in the control group (84.1 vs. 74.7%, P=0.031). For the statistical analysis, we combined AG and GG genotypes into one group (AG+GG) because of the scarcity of GG genotype in the subjects. We hypothesized that *MUC1* A/G polymorphism at 568 site might be the factor that increases the susceptibility of this population to GC. To test this hypothesis, we used logistic regression adjusted by age and gender (95% CI: 1.06-3.12, Table I) to analyze our data and confirmed that the risk of GC was increased 1.81-fold in AA genotype compared to AG+GG genotype in the subjects studies.

MUC1 protein expression in situ in subclassification of GC. We next asked whether this *MUC1* A/G polymorphism at 568 site has an influence on the expression and function of MUC1 protein that consequently may be associated with the susceptibility of this population to GC. To address this question, we first detected the protein expression of MUC1 from the samples using an immunohistochemical analysis *in situ*. Fig. 2 shows representative immunohistochemistry staining for the MUC1 protein expression in the GC, subclassifications of GC and control specimens. The MUC1 protein was extensively expressed in the gastric mucosa of the control tissue from superficial gastritis (Fig. 2A) and atrophic gastritis (Fig. 2B). The expression of MUC1 protein were mainly detected in the membrane (thick black arrow). We also observed that MUC1 protein was located in cytoplasm (thin black arrow). We assessed the potential association between MUC1 protein expression and subclassifications of GC diagnosed according to Lauren classification (27). The expression of MUC1 protein in the intestinal adenocarcinoma was highly positive as demonstrated in Fig. 2D, while it was not clearly seen in the cell of well-differentiated tubular adenocarcinoma (Fig. 2C). However, we observed a negative or weak expression of MUC1 protein in diffuse carcinoma (Fig. 2E and F). We examined the association between MUC1 protein expression and GC including subclassifications of GC, and control samples. The MUC1 HSCORE and MUC1 positive rates of these samples are summarized in Table II. Association between MUC1 protein expression and GC is confirmed by a comparison of GC and subclassifications of GC with non-cancer samples. The percentage of positive and highly positive MUC1 expression in GC was significantly lower than that in control samples (P=0.000 and P=0.039 respectively). Comparison between two subclasses of GC showed that the positive rate of diffuse carcinoma (80.0%) was significantly lower than that of intestinal adenocarcinoma (100.0%, P=0.022, Fisher's exact probabilities). While the difference between highly positive rate of diffuse carcinoma and intestinal adenocarcinoma was not statistically significant (P=0.06), the results showed a trend for an association between MUC1 protein expression and subclasses of GC.

MUC1 gene polymorphism affected MUC1 protein expression in stomach. We next addressed the question of whether the SNP affects MUC1 expression in GC patients by analyzing

Table II. Expression of MUC1 protein *in situ* in the GC patients and control subjects.

Groups	n	MUC1 HSCORE				Protein expression	
		0.0	0.0-0.9	1.0-1.9	2.0-3.0	Positive (%)	Highly positive (%)
Non-cancer	191	1	37	64	89	99.5	46.6
GC	71	9	22	17	23	87.3[a]	32.4[b]
Intestinal adenocarcinoma	26	0	6	8	12	100.0	46.2
Diffuse carcinoma	45	9	16	9	11	80.0[c]	24.4

Comparison was performed using χ^2 test. [a]P=0.000, GC group vs. non-cancer group; [b]P=0.039, GC group verses non-cancer group; [C]Comparison was performed between GC and diffuse carcinoma using Fisher's extract probabilities, P=0.022.

Table III. Expression of MUC1 protein in different gastric diseases.

Groups	n	AA		AG+GG		P-value[a]	OR[b] (95% CI)
		Positive (%)	Highly positive (%)	Positive (%)	Highly positive (%)		
GC	71	51/58 (87.9)	19/58 (32.8)	11/13 (84.6)	4/13 (30.8)		
Non-cancer	191	140/141 (99.3)	56/141 (39.7)	50/50 (100.0)	36/50 (72.0)	0.000	3.76 (1.87-1.53)
Superficial gastritis	105	76/77 (98.7)	28/77 (36.4)	28/28 (100.0)	18/28 (64.3)	0.011	3.13 (1.28-7.69)
Atrophic gastritis	86	64/64 (100.0)	28/64 (43.8)	22/22 (100.0)	18/22 (81.8)	0.002	5.88 (1.75-20.00)

[a]Comparison of the highly positive rates of MUC1 protein expression in AA and AG+GG genotypes was performed on using non-conditional logistic regression. [b]OR value was adjusted by age and gender.

the total specimens of GC and controls. The association between *MUC1* gene polymorphism and MUC1 protein expression in the samples from GC and non-cancer controls are summarized in Table III. With consideration of the factors of age and gender, we found that there was a significant negative correlation between AA genotype and MUC1 protein expression (r=-0.1790, P=0.004). In non-cancer control samples (total n=191, superficial gastritis n=105, atrophic gastritis n=86), the highly positive rate of MUC1 expression was significantly lower in AA genotype carriers when compared with the samples from AG+GG carriers (P=0.000), while there was no difference in the positive rate of MUC1 expression between genotype carriers of AA and AG+GG. In the subclasses of superficial and atrophic gastritis, the highly positive rates of MUC1 expression were decreased in AA genotype carriers when compared with those in AG+GG carriers (P=0.011 and 0.002, respectively). These findings show that the *MUC1* gene A/G polymorphism at 568 site leads to decreased MUC1 protein levels in non-cancerous gastric mucosa. We did not find a significant difference in MUC1 expression in AA and AG+GG genotypes of GC samples in this study.

Prediction of splice acceptor at MUC1 polymorphism 568 site. Ligtenberg and colleagues reported that ACAG was a splice acceptor (10). The stem-loop structure of variant A (with ACGG sequence, Fig. 3) is thought to be stable and cannot be spliced. However, when the polymorphism site transferred from G to A in the variant B, the ACGG is transferred to ACAG which is a splice acceptor site. When we used the Spliceview in http://www.itb. cnr.it/sun/webgene, we obtained an extra splice acceptor ACAG at the position of 636, when we substituted G in the position of 634 of *MUCI* with A. *MUC1* has 4401 bases and the polymorphic site is located at the position of 634. However, there are 66 bases before the initiator (http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=4072037) and thus the position of 568 was used in the literature and in our study. We summarize the prediction of splice acceptor at *MUC1* polymorphism 568 site in Fig. 3.

Discussion

The identification of genetic factors capable of modulating cancer development has the potential to unravel disease heterogeneity and aid diagnostic and prevention strategies. We report, for the first time, the distribution frequencies of AA, AG, and GG genotypes in the North-East region of China. The results presented here demonstrate that the frequency of AA genotype in patients with GC is significantly higher than that of the non-cancer group. This is the first study that shows a clear association of the *MUC1* A/G polymorphism at 568 site with an increased risk of GC. The result of a significantly reduced frequency in blood DNA of

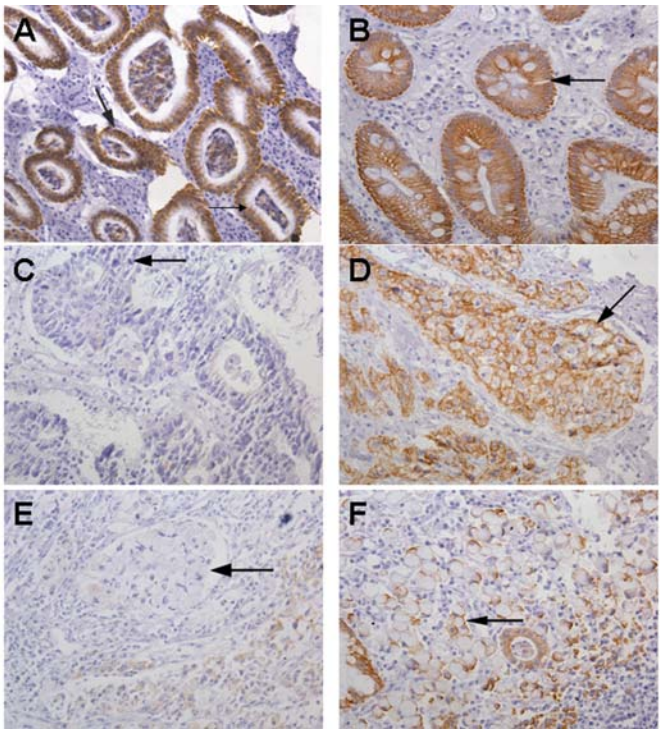

Figure 2. Representative immunohistochemical staining for MUC1 protein in GC, GC subclassifications and control specimens. (A) Positive expression of MUC1 protein in superficial gastritis (400x), the thick black arrow indicates MUC1 protein located in membrane and the thin black arrow indicates MUC1 protein located in cytoplasm. (B) Positive expression of MUC1 protein in atrophy gastritis (400x), the black arrow indicates the MUC1 protein location. (C) Negative expression of MUC1 protein in intestinal adenocarcinoma (400x), the black arrow indicates the loss of MUC1 in the cell of well-differentiated tubular adenocarcinoma. (D) Positive expression of MUC1 protein in intestinal adenocarcinoma (400x), the black arrow indicates MUC1 protein located in the cell of well-differentiated tubular adenocarcinoma. (E) Negative expression of MUC1 protein in diffuse-type carcinoma (400x), the black arrow indicates the loss of MUC1 in the cell of signet ring cell carcinoma. (F) Positive expression of MUC1 protein in diffuse-type carcinoma (400x), the black arrow indicates MUC1 protein located in the cell of signet ring cell carcinoma.

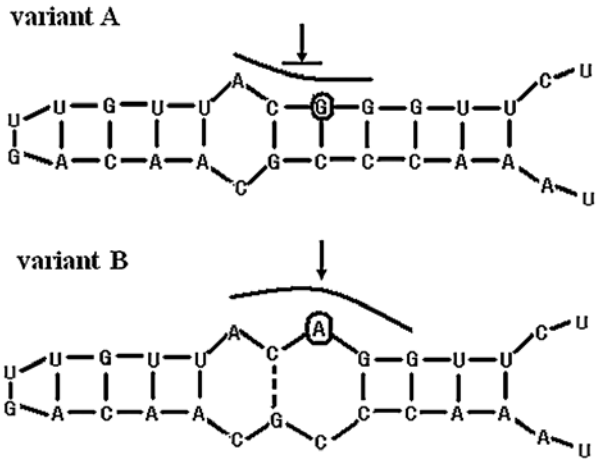

Figure 3. Putative schematic diagram of two variants, variant A and variant B. The +568 polymorphism site is encircled. Upper is a stem-loop structure of variant A. The stem-loop structure of variant A is thought to be stable and can not be spliced. The stop arrow of variant A indicates ACGG can not be spliced. When the polymorphism site is transferred from G to A, the ACGG is transferred to ACAG which is a splice acceptor site. The arrow of variant B indicates ACAG can be spliced.

MUC1 568G allele (G to A) provides the evidence that *MUC1* A/G polymorphism at 568 site might be the factor which leads to an increase in the susceptibility of the population in the North-East region of China to GC. Our result confirms our initial hypothesis that the *MUC1* A/G polymorphism at 568 site influences the expression and/or functions of MUC1 protein that consequently may be associated with the susceptibility of a population that carries the A allele to GC.

MUC1 is a large molecular weight (2000 kDa) glycoprotein, its space structure is composed of the core chain region and some glycosylated side chains. MUC1 is a transmembrane molecule with a large extracellular domain which contains a variable number of tandem repeats (VNTR) sequences of 20 amino acids. The *MUC1* gene in 1q21-24 which encodes the MUC1 protein, is now known to have two exon polymorphisms, which have been reported as a VNTR polymorphism in exon 2 and its upstream as an A/G single nucleotide polymorphism (SNP) within exon 2 at 568 site (10,23,28). The Human Genome Project has reported (http://www.ncbi.nlm.nih.gov/SNP/snp_ref.cgi?rs=4072037) the distribution frequencies from the people of representative countries and regions in *MUC1* A/G polymorphism at 568 site. The study showed that the distribution frequencies of AA, AG, and GG genotypes in 45 Han Chinese in Beijing were 64.4, 31.1 and 4.4%, that of allele A was 80% and allele G was 20%. The allele frequencies of Japanese people were similar to the observation of the Chinese people, while Europeans and Africans were different from the Chinese and Japanese (allele frequencies were 57.9 and 42.1% in American and 49.2 and 50.8% in African populations, respectively). Less specific information is available on the association between *MUC1* A/G single nucleotide polymorphism and the risk of gastric cancer (GC). Several cancer-related studies have reported an association between the *MUC1* A/G polymorphism at 568 site and alterations in DNA and its encoding proteins have recently been linked with diseases. Janssen and colleagues reported this polymorphism had association with the KL-6 expression in the serum, which was MUC1 antibody in the lung and was an indicator of pulmonary fibrosis (23). Imbert and colleagues found that this polymorphism was associated with the dry eye syndrome, and MUC1 secreted by allele A carriers in the cornea and conjunctiva functioned in an inferior manner when compared to allele G carriers, which were more susceptible to dry eye syndrome (29,30). Silva (17) predicted that *MUC1* polymorphism might define different susceptibility backgrounds for the development of conditions that precede gastric carcinoma: chronic atrophic gastritis and intestinal metaplasia. The results of these and other studies provide increasing evidence for the pathological and clinical significance that the *MUC1* A/G polymorphism at 568 site may have. This study provides useful information for a better understanding of a population's susceptibility to GC in the North-East region of China where a high incidence rate of GC has been observed. Whether this polymorphism can serve as a potential predictor of GC among other Asian populations who share the same color of skin requires additional studies, likely with large sample sizes and multiple-centers.

Analysis of the association between the *MUC1* A/G polymorphism at 568 site and the increase of risk to GC,

demonstrating an increased risk to GC, implicates the underlying mechanisms. This polymorphism may be attributed to the decrease of the MUC1 expression in individuals with AA genotype and the further decrease of the physiologically protective functions to the gastric mucosa. MUC1 protein possesses multiple functions. Its participation in the 'mucus-bicarbonate barriers' could enable the pepsin on the surface epithelial cells to lose their function of decomposing protein, thereby preventing digestive enzymes (pepsin) from destroying themselves (31-33). When external factors stimulate gastric mucosa, the mucin could prevent mucosa from stomach acid and digestive enzymes (pepsin), which is beneficial to epithelial repair (31,34). MUC1 protein is also viscous and gluey, so it formed a certain thick mucous layer on the stomach surface, with the function of lubricating and protecting (35). It has been reported that the weaker protective function of AA genotype as compared to that of AG+GG genotype was caused either by the impact on the physiology, or by the length extending out of the cell membrane (29,36). In short, the increase of the genetic susceptibility to GC in AA genotype carriers could be a result of the alteration of MUC1 protein expression and function encoded from an altered *MUC1* gene by this *MUC1* A/G polymorphism at 568 site. We hypothesized that this SNP might play a role in the regulation of *MUC1* and its MUC1 protein expression. We here provide an experimental basis for the association between *MUC1* gene polymorphism and GC. The detection of this polymorphism may have the clinical significance of its use as a screening indicator for GC.

Ligtenberg and colleagues noted the possible importance and potential prognostic value of this polymorphism and suggested a genetic basis for the variable splicing event (10). The secondary structure of the pre-mRNA was also predicted to be different, where only the G allele forms a physiologically stable stem loop structure as demonstrated by Ligtenberg and colleagues who noted the possible importance of this difference. In the putative effect of secondary structure predicted by Ligtenberg, if a G was present at +8, it could form a stem-loop structure of the region surrounding the splice acceptor site which was much more stable than the one formed if an A was present at this position. The allele G at the polymorphic site and splice to the upstream splice acceptor site resulted in variant A, whereas allele A at this position and splice to an acceptor site located 27 bp further downstream, resulted in splice variant B (37). *MUC1* which coded MUC1 protein possessed a splice acceptor site in the second exon and the effect of the polymorphism on the MUC1 expression may be related to the alteration in the splicing site. Whatever the mechanism, there appears to be an alteration in the splice receptor site. We speculate that the secondary structure of the pre-mRNA can make the variant B splice acceptor site accessible to the splicing mechanism. When G transferred to A at position +8 and expressied as the B variant, the mRNA differed from 27 bp could translate MUC1 proteins. This MUC1 protein is altered (9 amino acids), which may results in individuals with AG+GG genotype different from the AA genotype. This allele difference in genotype could be a cancer-related change in splicing, which alters the MUC1 expression.

In conclusion, we used PCR-SSPs and immunohistochemistry to identify the genotype of *MUC1* A/G polymorphism at its 568 site of exon 2 and association between this polymorphism and the level of MUC1 protein expression in GC patients and non-cancer subjects from North-East China. We report for the first time that the frequency of AA genotype was significantly high in the GC patients and the risk for GC in AA genotype carriers is increased by a factor of 1.81. Moreover, our study provides a possible mechanistic insight in that the SNP disrupts the physiological functions of MUC1 which is important to the physiological protection of gastric mucosa. Additional studies, such as mapping, sequencing, splicing acceptor site and functional studies will be needed to define the molecular mechanisms underlying our observed association.

Acknowledgements

This study is supported by grants from the National Natural Science Foundation of China (Ref No. 30572131), the Foundation of the Key Laboratory of Cancer Intervention in Liaoning Province (Ref No. 20060913).

References

1. Xu Q, Sun LP, Gong YH, Xu Y, Dong NN and Yuan Y: The relationship between the polymorphism of MUC1 and susceptibility to gastric cancer in Liaoning region. Yi Chuan 30: 1163-1168, 2008.
2. Fox JG and Wang TC: Inflammation, atrophy, and gastric cancer. J Clin Invest 117: 60-69, 2007.
3. McGuckin MA, Every AL, Skene CD, *et al*: Muc1 mucin limits both *Helicobacter pylori* colonization of the murine gastric mucosa and associated gastritis. Gastroenterology 133: 1210-1218, 2007.
4. Silva F, Carvalho F, Peixoto A, *et al*: MUC1 polymorphism confers increased risk for intestinal metaplasia in a Colombian population with chronic gastritis. Eur J Hum Genet 11: 380-384, 2003.
5. Ponce-Bravo S, Ledesma-Montes C and Morales-Sanchez I: MUC-1 mucin in normal human salivary glands detected by HMFG-1 and HMFG-2 monoclonal antibodies. APMIS 116: 93-98, 2008.
6. Taylor-Papadimitriou J, Burchell JM, Plunkett T, *et al*: MUC1 and the immunobiology of cancer. J Mammary Gland Biol Neoplasia 7: 209-221, 2002.
7. Gendler SJ, Cohen EP, Craston A, Duhig T, Johnstone G and Barnes D: The locus of the polymorphic epithelial mucin (PEM) tumour antigen on chromosome 1q21 shows a high frequency of alteration in primary human breast tumours. Int J Cancer 45: 431-435, 1990.
8. Gendler SJ, Lancaster CA, Taylor-Papadimitriou J, *et al*: Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. J Biol Chem 265: 15286-15293, 1990.
9. Lan MS, Batra SK, Qi WN, Metzgar RS and Hollingsworth MA: Cloning and sequencing of a human pancreatic tumor mucin cDNA. J Biol Chem 265: 15294-15299, 1990.
10. Ligtenberg MJ, Gennissen AM, Vos HL and Hilkens J: A single nucleotide polymorphism in an exon dictates allele dependent differential splicing of episialin mRNA. Nucleic Acids Res 19: 297-301, 1991.
11. Julian J and Carson DD: Formation of MUC1 metabolic complex is conserved in tumor-derived and normal epithelial cells. Biochem Biophys Res Commun 293: 1183-1190, 2002.
12. Macao B, Johansson DG, Hansson GC and Hard T: Auto-proteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin. Nat Struct Mol Biol 13: 71-76, 2006.
13. Goulart LR, Vieira GS, Martelli L, Inacio J, Goulart IM and Franco JG Jr: Is MUC1 polymorphism associated with female infertility? Reprod Biomed Online 8: 477-482, 2004.
14. Creaney J, Segal A, Sterrett G, *et al*: Overexpression and altered glycosylation of MUC1 in malignant mesothelioma. Br J Cancer 98: 1562-1569, 2008.

15. Strawbridge RJ, Nister M, Brismar K, Li C and Lindstrom S: Influence of MUC1 genetic variation on prostate cancer risk and survival. Eur J Hum Genet 16: 1521-1525, 2008.
16. Carvalho F, Seruca R, David L, *et al*: MUC1 gene polymorphism and gastric cancer - an epidemiological study. Glycoconj J 14: 107-111, 1997.
17. Silva F, Carvalho F, Peixoto A, *et al*: MUC1 gene polymorphism in the gastric carcinogenesis pathway. Eur J Hum Genet 9: 548-552, 2001.
18. Vinall LE, King M, Novelli M, *et al*: Altered expression and allelic association of the hypervariable membrane mucin MUC1 in *Helicobacter pylori* gastritis. Gastroenterology 123: 41-49, 2002.
19. Ho SB, Niehans GA, Lyftogt C, *et al*: Heterogeneity of mucin gene expression in normal and neoplastic tissues. Cancer Res 53: 641-651, 1993.
20. Adkins KK, Strom DA, Jacobson TE, Seemann CR, O'Brien DP and Heath EM: Utilizing genomic DNA purified from clotted blood samples for single nucleotide polymorphism genotyping. Arch Pathol Lab Med 126: 266-270, 2002.
21. Brisson-Noel A, Aznar C, Chureau C, *et al*: Diagnosis of tuberculosis by DNA amplification in clinical practice evaluation. Lancet 338: 364-366, 1991.
22. Brisson-Noel A, Gicquel B, Lecossier D, Levy-Frebault V, Nassif X and Hance AJ: Rapid diagnosis of tuberculosis by amplification of mycobacterial DNA in clinical samples. Lancet 2: 1069-1071, 1989.
23. Janssen R, Kruit A, Grutters JC, Ruven HJ, Gerritsen WB and van den Bosch JM: The mucin-1 568 adenosine to guanine polymorphism influences serum Krebs von den Lungen-6 levels. Am J Respir Cell Mol Biol 34: 496-499, 2006.
24. Grutters JC, Sato H, Pantelidis P, *et al*: Increased frequency of the uncommon tumor necrosis factor -857T allele in British and Dutch patients with sarcoidosis. Am J Respir Crit Care Med 165: 1119-1124, 2002.
25. Byrd JC, Ho JJ, Lamport DT, *et al*: Relationship of pancreatic cancer apomucin to mammary and intestinal apomucins. Cancer Res 51: 1026-1033, 1991.
26. Berchuck A, Soisson AP, Clarke-Pearson DL, *et al*: Immunohistochemical expression of CA 125 in endometrial adenocarcinoma: correlation of antigen expression with metastatic potential. Cancer Res 49: 2091-2095, 1989.
27. Lauren P: The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification. Acta Pathol Microbiol Scand 64: 31-49, 1965.
28. Pratt WS, Islam I and Swallow DM: Two additional polymorphisms within the hypervariable MUC1 gene: association of alleles either side of the VNTR region. Ann Hum Genet 60: 21-28, 1996.
29. Imbert Y, Darling DS, Jumblatt MM, *et al*: MUC1 splice variants in human ocular surface tissues: possible differences between dry eye patients and normal controls. Exp Eye Res 83: 493-501, 2006.
30. Imbert Y, Foulks GN, Brennan MD, *et al*: MUC1 and estrogen receptor alpha gene polymorphisms in dry eye patients. Exp Eye Res 88: 334-338, 2009.
31. Hogan DL, Ainsworth MA and Isenberg JI: Review article: gastroduodenal bicarbonate secretion. Aliment Pharmacol Ther 8: 475-488, 1994.
32. Turnberg LA: Gastric mucosal defence mechanisms. Scand J Gastroenterol (Suppl) 110: 37-40, 1985.
33. Turnberg LA: Gastric mucus, bicarbonate and pH gradients in mucosal protection. Clin Invest Med 10: 178-180, 1987.
34. Gutierrez-Cabano CA: Thioctic acid protection against ethanol-induced gastric mucosal lesions involves sulfhydryl and prostaglandin participation. Acta Gastroenterol Latinoam 27: 31-37, 1997.
35. MacAdam AB, Shafi ZB, Marriott C, Martin GP and James SL: Anti-mucus polyclonal antibody production, purification and linkage to the surface of albumin microspheres. Int J Pharm 195: 147-158, 2000.
36. Hilkens J, Buijs F, Hilgers J, *et al*: Monoclonal antibodies against human milk-fat globule membranes detecting differentiation antigens of the mammary gland and its tumors. Int J Cancer 34: 197-206, 1984.
37. Wreschner DH, Hareuveni M, Tsarfaty I, *et al*: Human epithelial tumor antigen cDNA sequences. Differential splicing may generate multiple protein forms. Eur J Biochem 189: 463-473, 1990.