A Novel Tri-Allelic Insertion/Deletion Polymorphism in the Promoter of $p21^{Waf1/Cip1}$ and the Association with Gastric Cancer

Zhiqiang Chen,^{1,2,*} Ying Dong,^{1,*} Ningju Wang,^{3,*} Hanmei Liang,¹ Yong Du,⁴ Xiaofeng Ye,³
 Wei Jia,⁴ Zhongyi He,¹ Yideng Jiang,¹ Yinxue Yang,^{3,4} and Wenjun Yang^{1,3}

Aims: $p21^{Waf1/Cip1}$ is a cyclin-dependent kinase inhibitor that is pivotal in arresting cellular growth in terminal cell differentiation and apoptosis. Thus, the existence of natural variants of $p21^{Waf1/Cip1}$ could be linked to specific cancer. The purpose of this report was to identify a novel tri-allelic insertion/deletion (INDEL) polymorphism (rs4135235) involving a poly-T sequence in the promoter region of $p21^{Waf1/Cip1}$ gene and to explore its role in gastric cancer (GC). **Method:** A total of unrelated 676 subjects (376 GC patients; 300 cancer-free controls) were enrolled in the study, and genomic DNA was obtained from each subject for genotyping. PCR-directed sequencing technique was used to detect the genotypes of the polymorphism. TA cloning was used to confirm the existence of three alleles. Odds ratios (ORs) and 95% confidence intervals (CIs) were estimated by unconditional logistic regression analysis. **Results:** Six genotypes (9T/9T, 10T/10T, 11T/11T, 9T/10T, 10T/11T, and 9T/11T) and three alleles (9Ts, 10Ts, and 11Ts) were identified among all study subjects. GC cases were different from the control group with over-representation of 9T/11T heterozygotes (19.7% vs. 12.3%) and under-representation of 10T/10T homozygotes (18.4% vs. 20.7%). Compared with those carrying 10T/10T, individuals with 9T/11T increased the susceptibility for GC (OR = 1.797, 95%CI = 1.065–3.031). **Conclusion:** Our findings confirmed the existence of a tri-allelic polymorphism in the promoter of $p21^{Waf1/Cip1}$. It has also shown the heterozygous genotype 9T/11T to be a potential risk factor for GC in the Chinese population.

Introduction

$p21^{Waf1/Cip1}$ (OMIM:116899; hereafter, $p21$), a critical cyclin-dependent kinase inhibitor, modulates kinase activity by interacting with cdk/cyclin complexes and arresting cell division (Sherr and Roberts, 1999). Additionally, $p21$ prevents DNA synthesis and regulates DNA replication through interaction with proliferating cell nuclear antigen (Dotto, 2000). In addition to the modulation of cell proliferation, $p21$ has been identified as a p53-dependent protein implicated in cell cycle checkpoints. $p21$ inhibits cell cycle progression through the transactivation of p53 when cells are exposed to DNA-damaging agents (El-deiry *et al.*, 1993; Abbas and Dutta, 2009). Hence, $p21$ acts as a tumor suppressor.

The $p21$ gene (also known as *CDKN1A*) is an important downstream target gene whose expression is directly regu-

lated by the p53 tumor suppressor protein (El-deiry *et al.*, 1993). $p53$ and $p21$ genes were shown to be essential for maintaining the G2 checkpoint in human cells (Bunz *et al.*, 1998). Because mutations or functional inactivation within the $p53$ pathway are an almost universal feature of human cancer, it comes as no surprise that mutations or inherited defects of the $p21$ gene might promote genetic instability, leading to several disorders (Goh *et al.*, 2011). Conceivably, the existence of mutations or natural variants of $p21$ gene could be linked to specific cancers, but evidence to suggest otherwise is strong: the $p21$ gene is rarely mutated in human tumors. Within the $p21$ gene's translation region, two single nucleotide polymorphisms (SNPs) have been discovered, and only one ($p21$ codon 31) was later confirmed in an actual population (Xi *et al.*, 2002; Gartel and Radhakrishnan, 2005; Pietsch *et al.*, 2006; Hrstka *et al.*, 2009). In a polymorphism scanning study of the $p21$ gene

¹Key Laboratory of Reproduction and Heredity, Oncology Department of Cancer Hospital, General Hospital, Ningxia Medical University, Yinchuan, People's Republic of China.

²Radiology Department of General Hospital, Ningxia Medical University, Yinchuan, People's Republic of China.

³Oncology Department of Cancer Hospital and Institute, General Hospital, Ningxia Medical University, Yinchuan, People's Republic of China.

⁴Surgical Laboratory of General Hospital, Ningxia Medical University, Yinchuan, People's Republic of China.

*These authors contributed equally to this work.

promoter, we genotyped a novel tri-allelic insertion/deletion variant (INDEL) involving a poly-T sequence in a Chinese population of 376 gastric cancer (GC) patients and 300 cancer-free controls, which was previously reported to be a mixed T to A substitution INDEL (www.ncbi.nlm.nih.gov/snp/?term=rs4135235) in the -1933 bp area of the *p21* gene promoter region. TA cloning and denaturing high-performance liquid chromatography (DHPLC) methods were further conducted to confirm the variant's alleles and genotypes. We also investigated the relationship between the novel INDEL and GC.

Materials and Methods

Study subjects

Patients ($n=376$) with GC were histopathologically confirmed between 2007 and 2011 in the General Hospital of the Ningxia Affiliated Hospital, Yinchuan City of the Ningxia region of China. Demographic data, including age at diagnosis and gender were obtained from medical records. Cancer-free controls ($n=300$) were selected from participants in the GC screening program in the Ningxia area and were gender- and age-matched with the GC cases. Demographic data were collected using a brief interview.

All subjects enrolled in this study were unrelated to one another. Blood samples from all subjects were collected for DNA preparation and genotyping of the *p21* gene.

Research protocols and related materials were submitted to and approved by Ningxia Medical University Ethical Review Committee, China. All participants provided written informed consent according to regulations governing the humane treatment of human study participants.

Genotyping for *p21*-1933 INDEL polymorphism by PCR direct sequencing

Genomic DNA was extracted from the peripheral blood lymphocytes of the participants using a standard phenol-chloroform extraction. To identify the *p21* polymorphism at the upstream 1933 bp from the transcriptional start point, an 892-bp genomic fragment covering -1933 poly Ts was amplified using the forward primer 5'-TGTGGCTCTGATTG GCTTTC-3' and the reverse primer 5'-GGGAGGACAGGC TTCTTTC-3'. PCR was performed in a 50 μ L reaction mixture that contained DNA (50 ng) primer (0.1 μ M of each), dNTP (0.2 mM) $MgCl_2$ (2.0 mM), 0.5 units of high-fidelity platinum PCR polymerase (Invitrogen Life Technology), and 1 $\times$ reaction buffer. The PCR profile consisted of an initial melting step of 2 min at 95°C, followed by 35 cycles of 45 s at 94°C, 40 s at 61°C, 45 s at 72°C, and a final elongation step of 10 min at 72°C. The polymorphism was identified by directly sequencing the PCR products using ABI 3700 DNA sequencer (PE Applied Biosystems) with 5'-TGTGGCTCTGATTG GCTTTC-3' for forward direction and 5'-TGGTACTGAGCTTCAACAATG-3' for reverse direction. PCR products were purified by MultiScreen-PCR purifying plate (Millipore Company) and used for sequencing. The terminator ready sequencing reaction was performed in a total volume 10 μ L containing 2 μ L of PCR product, 4 pmol of sequencing primer, 2 μ L of Terminator Ready Reaction mixture (PE Applied Biosystems). Cycling conditions were as following: initial denaturing 96°C for 2 min, followed by 30 cycles at 94°C for 10 s, 50°C for 5 s, and 60°C for 4 min. After ethanol precipi-

tation, the reaction products were sequenced by an ABI 3700 DNA sequencer (PE Applied Biosystems) according to the manufacturer's instructions. The sequencing chromatograms were base-called with Phred (version 0.990722), assembled with Phrap (version 0.990329), scanned by PolyPhred (version 3.5), and the results were viewed with Consed (version 9.0). PCR and sequencing primers were designed by Oligo5.0 software. Genotyping quality was verified by repeat genotyping of at least 10% of samples with agreement rate of >99%.

Construction of *pGEM-T* easy cloning

A 892-bp fragment of the *p21* promoter region was amplified with the primers mentioned above. PCR was performed in a 20 μ L reaction mixture containing DNA (50 ng) primer (0.1 μ M of each), dNTP (0.2 mM) $MgCl_2$ (2.0 mM), 0.2 units of high-fidelity platinum polymerase (Invitrogen Life Technology), and 1 $\times$ reaction buffer. PCR reaction conditions were the same as mentioned above. The PCR product was purified with the QIAquick PCR purification kit (QIAGEN) and cloned into the *pGEM-T* Easy cloning vector (Promega) using a TA cloning strategy, according to the manufacturer's protocol. The cloned product was sequenced with an ABI 3700 DNA sequencer (PE Applied Biosystems) according to the manufacturer's instructions.

Analysis for variation by DHPLC

To optimize DHPLC conditions for variation analysis, a 194-bp fragment of the *p21* promoter was amplified from all samples with the forward primer 5'-TAGAGTGTGTCTGGG TACA-3' and the reverse primer 5'-TGGTACTGAGCTTCA CAATG-3'. PCR was performed at a 50 μ L reaction mixture containing 50 ng of DNA, 0.1 μ M of each primer, 0.2 mM of dNTP, 2.0 mM $MgCl_2$, 0.5 unit of Taq DNA polymerase (Promega), and 1 $\times$ reaction buffer. The PCR profile consisted of an initial melting step of 2 min at 95°C, followed by cycles of 45 s at 94°C, 40 s at 56°C, 45 s at 72°C, and a final elongation step of 10 min at 72°C. Meanwhile, each TA clone containing 9T, 10T, or 11T was also used as PCR template. PCR products were directly used as homozygote controls and the mixture of each two fragments were used as heterozygote controls. Detection of polymorphisms was detected in PCR products using DHPLC on a WAVE DNA Fragment Analysis System (Transgenomic). All PCR products were introduced into the mobile phase (5 μ L injection volume) by the autosampler. A DNASep analytical column (Transgenomic) was used as the solid phase and the products were eluted from the column with a binary gradient of 0.1 M triethylammonium acetate (TEAA) and 0.1 M TEAA/25% acetonitrile as the mobile phase (flow rate: 0.9 mL/min). Elution gradients and analysis temperatures were automatically predicted from the target sequences by WAVEMaker 4.0 software (Transgenomic) and eluted products were detected by UV analysis at 260 nm. Nondenaturing analysis was conducted at 55°C and partial denaturing analyses were conducted at the temperatures recommended by WAVEMaker.

Statistical analysis

Descriptive statistical analysis and unconditional logistic regression were performed with the Statistical Analysis

TABLE 1. COMPARISON OF CHARACTERISTICS BETWEEN CASES AND CONTROLS

	Case (n=376)	Control (n=300)	p-Value ^a
Mean age (year±SD)	58.61±10.318	58.83±9.867	0.302
Gender (male/female)	207/169	153/147	0.177

Mean±SD values for continuous variables.

^aTwo-sided χ^2 test.

System software (version 8.01; SAS Institute). Hardy-Weinberg equilibrium (HWE) was tested by comparing expected and observed genotype frequencies using the χ^2 test. Odds ratios (ORs) and 95% confidence intervals (CIs) were estimated by unconditional logistic regression analysis that was adjusted for age and gender. A *p* value of <0.05 was used as the criterion of statistical significance, and all statistical tests were two-sided.

Results

INDEL identification

Subjects (*n*=676), including 376 GC patients and 300 cancer-free controls were studied for genetic polymorphisms. No significant differences were found between GC cases and controls with respect to age and gender (Table 1). As shown in Figure 1, a poly-T sequencing chromatogram was detected with high-fidelity platinum PCR polymerase (Invitrogen) in each sample. Careful analysis of all sequencing pictures revealed that the chromatograms could be divided into two groups: Group 1 with one peak and Group 2 with two peaks once the sequence reached 10 (or 11) Ts or 10 (or 11) Ts as from

both forward and reverse directions, suggesting the presence of a mixed 1-T or 2-T INDEL polymorphism. Group 1 with single peak contains three homozygotes (Fig. 1A–C): 9T, 10T, and 11T. Group 2 with double peaks includes three heterozygotes (Fig. 1D–E): 9T/10T, 10T/11T, and 9T/11T, respectively. Genotyping quality was verified by repeat genotyping of at least 10% of samples with agreement rate of >99% and overall success rate of >99%. All discrepant genotype calls for any sample were re-genotyped or were left out of the final analysis.

TA cloning method

We first selected three samples from each kind of sequencing chromatograms to construct TA clones. Then five or six clones from each sample were selected for sequencing. Sequencing results revealed that all clones from the three kinds of homozygotes contained only 9T, 10T, or 11T. However, more than one clone type was found in the heterozygote groups. Both 9-T and 10-T sequences were found in clones from three samples of group 9T/10T. All clones from samples of group 10T/11T or 9T/11T had 10Ts or 9Ts in addition to 11Ts. All chromatograms containing poly-T clones appeared clear and one-peaked.

DHPLC method

To further confirm the genotypes of the polymorphisms, a 194-bp fragment covering the poly-T sequence from 20% of all samples was amplified and examined by DHPLC. Samples with single peaks indicate homozygous genotypes, and those with two peaks indicate heterozygous genotypes. Compared with the homozygous and heterozygous controls, all samples estimated to be 9T/9T, 10T/10T, or 11T/11T genotypes from

FIG. 1. Sequencing chromatograms of *p21-1933* INDEL polymorphism by PCR direct sequencing. All sequencing chromatograms could be classified into two groups: Group 1 containing three homozygotes with clear and single peak and Group 2 containing three heterozygotes with double peaks. (A) The homozygous sequence of 9T/9T. (B) The homozygous sequence of 10T/10T. (C) The homozygous sequence of 11T/11T. (D) The heterozygous sequence of 9T/10T. (E) The heterozygous sequence of 10T/11T. (F) The heterozygous sequence of 9T/11T.

FIG. 2. Analysis for variation by DHPLC. Three homozygous genotypes were one-peaked. Three heterozygous genotypes were two-peaked. DHPLC, denaturing high-performance liquid chromatography.

Group 1 were one-peaked whereas samples from Group 2 were two-peaked (Fig. 2).

Statistical analysis

Statistical analysis was performed and data for GC cases and controls are presented in Table 2. Genotype distributions in control group were in HWE ($p > 0.05$). GC cases were different from the control group with over-representation of 9T/11T heterozygotes (19.7% vs. 12.3%) and under-representation of 10T/10T homozygotes (18.4% vs. 20.7%). Compared with 10T/10T, the heterozygous 9T/11T genotype was associated with increased risk for development of GC (OR = 1.797, 95%CI = 1.065–3.031), being more susceptible to GC than any other genotypes.

Discussion

GC is one of the most frequently occurring cancers globally, with 989,600 new cases reported annually. In 2008, approximately 738,000 deaths were attributed to GC, representing 10% of total cancer-related deaths (Jemal *et al.*, 2009). Geographical distribution of GC has wide international variation. More than 70% new cases and deaths occur in developing countries (Ferlay *et al.*, 2010). GC is the second leading cause of cancer-related death in China, preceded only by lung cancer

(Bertuccio *et al.*, 2009). Environmental factors, dietary habits, and *Helicobacter pylori* infections are known to be associated with GC risk, but genetic factors may be a critical driver of GC. Recent studies suggest that inherited genetic polymorphisms in the *p53* pathway influence the formation, progression, and/or response to GC therapy (Tahara, 2004). However, such genetic variation studies are primarily focused on SNP. At this time, few INDELs related to GC have been reported. Because methods for studying INDELs are complex, reports about these aspects of cancer risk have lagged behind reports of other variation analysis forms. As the second most abundant genetic variation in humans after SNPs, INDEL variants remain largely under-discovered. Thus, few INDELs have been evaluated in validation studies, and we know relatively little about the influence of small INDEL variations on human biology and disease (Mullaney *et al.*, 2010; Mills *et al.*, 2011).

In our investigation, we identified the tri-allelic genetic variant rs4135235 in *p21* promoter for the first time from 376 GC cases and 300 cancer-free controls (from Yinchuan city, Ningxia region) who were recruited from an area of China with the highest incidences of GC. It was proved that a mixed 1-T or 2-T INDEL polymorphism existed downstream, 1933bp from the transcriptional start point of the *p21* promoter. INDELs that occur within transcription factor binding sites or enhancers are hypothesized to diminish or abolish gene expression. In addition, INDELs also can dramatically alter the phasing and spacing of DNA sequences within promoters (Lu *et al.*, 2012). Thus, different alleles of the poly-T INDEL in the promoter region of *p21* may probably affect the binding ability with certain transcriptional proteins, leading to the alteration of *p21* gene expression. Collectively, our data demonstrated that the *p21* rs4135235 polymorphism that existed in the promoter region of *p21* is a tri-allelic INDEL. The heterozygous genotype 9T/11T of the INDEL might be a potential risk factor for GC. This finding may assist in uncovering the mechanisms behind *p21* gene or *p53* pathway involvement in GC susceptibility.

Acknowledgments

We thank all the participants who participated in this study. This work was supported by National Natural Science

TABLE 2. GENOTYPE FREQUENCIES OF *p21*-1933 POLYMORPHISM IN CASES AND CONTROLS

Genotype	Values in		OR (95%CI) ^a
	Case (n=376)	Control (n=300)	
10T/10T	69 (18.4%)	62 (20.7%)	1.000
11T/11T	51 (13.6%)	38 (12.7%)	1.118 (0.654–1.909)
9T/9T	5 (1.3%)	7 (2.3%)	0.642 (0.194–2.126)
10T/11T	119 (31.6%)	102 (34.0%)	1.022 (0.614–1.700)
9T/11T	74 (19.7%)	37 (12.3%)	1.797 (1.065–3.031) ^b
9T/10T	58 (15.4%)	54 (18.0%)	1.039 (0.614–1.700)

^a Adjusted for age and gender.

^b $p < 0.05$.

OR, odds ratio; CI, confidence interval.

Foundation of China (No. 81160249 and 81260309), Ningxia Natural Science Foundation (No. NZ1234 and NZ13057), and Grant NZ11263 from the General Hospital of Ningxia Medical University, China.

Author Disclosure Statement

No competing financial interests exist.

References

- Abbas T, Dutta A (2009) p21 in cancer: intricate networks and multiple activities. *Nat Rev Cancer* 9:400–414.
- Bertuccio P, Chatenoud L, Levi F, *et al.* (2009) Recent patterns in gastric cancer: a global overview. *Int J Cancer* 125:666–673.
- Bunz F, Dutriaux A, Lengauer C, *et al.* (1998) Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science* 282:1497–1501.
- Dotto GP (2000) p21(WAF1/Cip1): more than a break to the cell cycle? *Biochim Biophys Acta* 1471:M43–M56.
- El-Deiry WS, Tokino T, Velculescu VE, *et al.* (1993) WAF1, a potential mediator of p53 tumor suppression. *Cell* 75:817–825.
- Ferlay J, Shin HR, Bray F, *et al.* (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 127:2893–2917.
- Gartel AL, Radhakrishnan SK (2005) Lost in transcription: p21 repression, mechanisms, and consequences. *Cancer Res* 65:3980–3985.
- Goh AM, Coffill CR, Lane DP (2011) The role of mutant p53 in human cancer. *J Pathol* 223:116–126.
- Hrstka R, Coates PJ, Vojtesek B (2009) Polymorphisms in p53 and the p53 pathway: roles in cancer susceptibility and response to treatment. *J Cell Mol Med* 13:440–453.
- Jemal A, Bray F, Center MM, *et al.* (2009) Global cancer statistics. *CA Cancer J Clin* 61:69–90.
- Lu JT, Wang Y, Gibbs RA, Yu F (2012) Characterizing linkage disequilibrium and evaluating imputation power of human genomic insertion-deletion polymorphisms. *Genome Biol* 13:R15.
- Mills RE, Pittard WS, Mullaney JM, *et al.* (2011) Natural genetic variation caused by small insertions and deletions in the human genome. *Genome Res* 21:830–839.
- Mullaney JM, Mills RE, Pittard WS, *et al.* (2010) Small insertions and deletions (INDELs) in human genomes. *Hum Mol Genet* 19:R131–R136.
- Pietsch EC, Humbey O, Murphy ME (2006) Polymorphisms in the p53 pathway. *Oncogene* 25:1602–1611.
- Sherr CJ, Roberts JM (1999) CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes Dev* 13:1501–1512.
- Tahara E (2004) Genetic pathways of two types of gastric cancer. *IARC Sci Publ* 157:327–349.
- Xi YG, Ding KY, Ren YH, *et al.* (2002) Esophageal cancer in Chinese population: no polymorphism in codon 149 of P21(Waf1/Cip1) cyclin dependent kinase gene. *Oncogene* 21:7745–7748.

Address correspondence to:

Wenjun Yang, PhD

Key Laboratory of Reproduction and Heredity

Oncology Department of Cancer Hospital

General Hospital, Ningxia Medical University

Yinchuan 750004

People's Republic of China

E-mail: yw001@nxmu.edu.cn

Yinxue Yang, MD

Surgical Laboratory of General Hospital

Ningxia Medical University

Yinchuan 750004

People's Republic of China

E-mail: nyfyyyx@126.com