OPEN

Evaluation of a Novel Functional Single-Nucleotide Polymorphism (rs35010275 G>C) in *MIR196A2* Promoter Region as a Risk Factor of Gastric Cancer in a Chinese Population

Ming Xu, PhD, Fulin Qiang, MD, Yan Gao, MD, Meiyun Kang, MD, Meilin Wang, PhD, Guoquan Tao, PhD, Weida Gong, PhD, Haixia Zhu, MD, Dongmei Wu, PhD, Zhengdong Zhang, PhD, and Qinghong Zhao, PhD

Abstract: Single-nucleotide polymorphisms (SNPs) in microRNAs (miRNAs) have been suggested to influence the occurrence and progression of cancer through altering the expression and biological function of miRNAs. The aim of this study was to investigate whether the potential functional SNPs in *MIR196A2* promoter had effect on the susceptibility to gastric cancer (GC) in a Chinese population.

We conducted a 2-stage case–control study (753 cases and 854 controls in testing set; 940 cases and 1061 controls in validation set) to evaluate the association between 2 potential functional SNPs in *MIR196A2* promoter (rs12304647 A>C and rs35010275 G>C) and GC risk. The luciferase reporter assay and electrophoretic mobility shift assay were used to examine the functionality of the important polymorphism.

We found that the rs35010275 C allele was significantly associated with the decreased risk of GC (adjusted odds ratio = 0.85, 95% confidence interval = 0.77–0.94) in the combined case–control studies. The miR-196a expression levels in GC tissues were significantly higher than that in corresponding adjacent normal tissues ($P < 0.001$). Besides, each allele of rs35010275 displayed completely opposite effects to influence the transcription activity of *MIR196A2* promoter via recruiting different transcription factors or complexes.

The functional rs35010275 G>C polymorphism in *MIR196A2* promoter was significantly associated with miR-196a expression and influenced the genetic susceptibility to GC.

(*Medicine* 93(26):e173)

Abbreviations: 3′-UTR = 3′-untranslated region, EMSA = electrophoretic mobility shift assay, GC = gastric cancer, HWE = Hardy–Weinberg equilibrium, MAF = minor allele frequency, SNPs = single-nucleotide polymorphisms.

INTRODUCTION

Gastric cancer (GC) remains the fourth most common cancer worldwide with a huge number of new diagnosed cases every year, especially in Asian countries.[1,2] In recent decades, GC mortality has been reported to decline; however, it is still the second fatal cause in China, following lung cancer.[2] Clinical data have shown that most GC patients were diagnosed with an advanced stage, which means these patients have missed their best treatment period and will confront a poor survival. A better understanding of the molecular mechanisms underlying gastric carcinogenesis may help to identify more accurate diagnostic markers and even more effective treatment strategies for this lethal disease.[3–5]

Classic genetic studies have implicated that aberrant expressions of tumor-suppressor genes or oncogenes are crucial factors for GC occurrence. With the development of epigenomics, alterations of CpG islands and whole genomic methylation were proved to be involved in gastric carcinogenesis.[6–8] Besides, accumulating evidences have indicated that microRNAs (miRNAs), as important factors, also participate in the etiology of GC.[9–11] MiRNAs are a sort of endogenous small noncoding RNAs, which cleaved from 70 to 100 nucleotides hairpin miRNA precursors, and are about 19 to 25 nucleotides long in mature form. In eukaryotes, mature miRNAs mainly affect the expressions of diverse protein-coding genes by targeting their mRNA in 3′-untranslated regions (3′-UTRs), and then leading to a post-transcriptional retardation by either inhibiting mRNAs' translation or accelerating their degradation.[12] It has been reported that miRNAs, including miR-196a, were noticed for their aberrant behaviors on various target mRNAs covering almost all of the important signal pathways.[13–16]

Editor: Bulent Kantarceken.
Received: June 15, 2014; revised: September 9, 2014; accepted: September 13, 2014.
From the Department of Environmental Genomics (MX, YG, MK, MW, HZ, DW, ZZ), Jiangsu Key Laboratory of Cancer Biomarkers, Prevention and Treatment, Cancer Center, Nanjing; Department of Genetic Toxicology (MX, YG, MK, MW, DW, ZZ), The Key Laboratory of Modern Toxicology of Ministry of Education, School of Public Health, Nanjing Medical University, Nanjing; Core Laboratory of Nantong Tumor Hospital (FQ, HZ), Nantong; Department of General Surgery (GT), Huai-An First People's Hospital Affiliated to Nanjing Medical University, Huai-an; Department of General Surgery (WG), Yixing Cancer Hospital, Yixing; and Department of General Surgery (QZ), The Second Affiliated Hospital of Nanjing Medical University, Nanjing, China.

Correspondence: Zhengdong Zhang, Department of Environmental Genomics, School of Public Health, Nanjing Medical University, 818 East Tianyuan Road, Jiangning District, Nanjing 211166, China (e-mail: drzdzhang@gmail.com).

MX, FQ, and YG contributed equally to this work.
This study was partly supported by the National Natural Science Foundation of China (81230068, 81302490, 81373091, and 81473049), the Natural Science Foundation of Jiangsu Province (BK2011194 and BK2012842), the Key Program for Basic Research of Jiangsu Provincial Department of Education (12KJA330002), Jiangsu Provincial Science and Technology Innovation Team, and the Priority Academic Program Development of Jiangsu Higher Education Institutions (Public Health and Preventive Medicine).

The authors have no funding and conflicts of interest to disclose.

ISSN: 0025-7974
DOI: 10.1097/MD.0000000000000173

Recently, single-nucleotide polymorphisms (SNPs) in several miRNAs genes have aroused researchers' concern. The potential role of these SNPs have been identified in the cancer development, such as nasopharyngeal carcinoma,[17] breast cancer,[18,19] hepatocellular carcinoma,[20,21] lung cancer,[22,23] head and neck cancer,[24] and GC.[25,26] The *MIR196A2* rs11614913, as a prestigious biomarker in cancers, including GC, was reported in our previous study.[27] The variant of rs11614913 could strongly influence the binding ability of mature miR-196a to its target mRNAs.[28,29] Overexpression of miR-196a was common in colon cancer and GC.[16,30–32] Considering functional polymorphisms in promoter region could influence gene expression,[33–35] it is rational to hypothesize that there exist some SNPs in *MIR196A2* promoter region that might directly trigger the transcription of *MIR196A2* and finally affect GC susceptibility. In our previous report, rs11614913 SNP was associated with the risk of GC.[36] However, as a variant existing in mature miRNA, it was still difficulty to affect the transcription of miR-196a-2. As an intergenic miRNA gene, *MIR196A2* has recently been reported to possess its own promoter that might alter miR-196a expression level and closely associate with the incidence of GC.[36] To test this hypothesis, we initiated to screen *MIR196A2* promoter region and then focused on 2 potentially functional SNPs (rs12304647 A>C and rs35010275 G>C). Through further functional study, we testified their probable mechanisms and evaluated the possibility of these SNPs as biomarkers for GC susceptibility and clinical outcomes.

MATERIALS AND METHODS

Study Population

This study was approved by the Institutional Review Board of Nanjing Medical University, Nanjing, China, and all subjects participated in this 2-stage study signed a written informed consent form. Seven hundred fifty-three GC cases and 854 cancer-free controls in testing set and additional subjects consisted of 1022 cases and 1061 cancer-free participants in validation set were periodical recruited from the Cancer Clinical Research Base of Nanjing Medical University during March 2006 to January 2010 and January 1999 and December 2006, respectively, which had been described previously in details.[37] In the set for validation, 78 patients (7.6%) with incomplete follow-up information and 4 patients with nonadenocarcinoma were excluded, and finally 940 patients with gastric adenocarcinoma were enrolled. All patients involved in our study were newly diagnosed and histopathologically identified, without previous history of cancer or previous chemotherapy or radiotherapy. Collecting data on patients mainly included age, sex, and tumor site histological type, depth of invasion, lymph node metastasis, distant metastasis, as well as clinical tumor node metastasis (TNM) stage. Histopathology of tumor was classified to diffuse type and intestinal type according to Lauren criteria.[38] The tumor invasion, lymph nodes metastasis, distant metastasis, and clinical TNM stage were recorded according to TNM classification (American Joint Commission on Cancer Staging, 6th).[39] All controls were enrolled in the same period with no genetic relationship to the cases. Besides, the controls were frequency matched to cases by sex and age (±5 years) (Table 1). Each subject donated 5 mL peripheral blood after interviews.

SNP Selection

SNPs in *MIR196A2* promoter region were selected based on HapMap data (http://hapmap.ncbi.nlm.nih.gov/) and dbSNP data (http://www.ncbi.nlm.nih.gov/projects/SNP/). The potentially functional polymorphisms were identified following the criteria: located in the *MIR196A2* 5′-flanking region; and Minor allele frequency (MAF) >0.05 in Chinese population. According to the criteria, 2 SNPs (rs12304647 A>C and rs35010275 G>C) remained in our study.

DNA Isolation and Genotyping

Genomic DNA was extracted from peripheral blood and paraffin sections of tissues for testing set and validation set, respectively. After proteinase K digestion and phenol–chloroform extraction, approximately 10% of DNA samples were randomly selected for agarose electrophoresis as quality control. Genotyping was carried out by fluorescent-based TaqMan SNP Genotyping Assay using ABI 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, CA). Quality control was performed with the quality criteria as our previous study.[37]

RNA Isolation and qRT-PCR

Total RNA were extracted by TRIzol reagent (Invitrogen, Carlsbad, CA) according to the protocol, which were from 66 pairs of patients' tumor and adjacent normal tissues including 753 cases in our testing set. RNA was measured by Nanodrop ND-2000 spectrophotometer (Thermo, Waltham, MA) for its quality and quantity, and then stored at −80°C.

Quantitative reverse transcriptase PCR was carried out by ABI 7900HT Fast Real-Time PCR System (Applied Biosystems). The PCR of each sample was normalized against an U6 internal control. The primer sequences were mentioned in supplementary data (see supplementary Digital Content 1, http://links.lww.com/MD/A80). All PCR reactions were conducted in triplicate.

Construction of Promoter–Reporter Plasmids

The 5′-flanking region sequence of human *MIR196A2* gene was obtained from the *Homo sapiens* chromosome 12 (NC_000012.11) after a blast search, starting from the first base of pri-miRNA. Luciferase reporter plasmids and corresponding variants were constructed by PCR amplification of nucleotides −1000, −700, −500 to transcriptional start site of *MIR196A2* promoter from human genomic DNA. All amplified fragments were cloned into pGL3-basic vectors (Promega, Madison, WI) and sequenced to confirm the orientation and integrity of each construct's inserts.

Cell Culture

Gastric epithelium cell line (GES-1) and SGC-7901 are both cultured in Dulbecco's Modified Egale medium/high glucose (Invitrogen) culture medium, supplementing with 10% heat-inactivated fetal bovine serum obtained from GIBCO (Burlington, ON, Canada), 10 mmol/L 4-(2-hydroxyethyl)-1-piperazineethanesulfonicaci, 2 mmol/L L-glutamine, 1 mmol/L pyruvate sodium, 100 U/mL penicillin, and 100 μg/mL streptomycin at 37°C in a humidified atmosphere containing 5% CO_2.

Transfection and Luciferase Assay

For transfection, 1×10^6 cells were seeded in individual well of a 24-well culture plate. Cells were transfected via lipofectamine-2000 transfection reagent with 0.5 μg constructed luciferase reporter plasmids mentioned above. The pRL-SV40 (as internal control) was transiently cotransfected into cells. Twenty-four hours after transfection, all cells were

TABLE 1. Demographic Characteristics and Clinical Features

Variables	Testing Set			Validation Set		
	Cases (n = 753) n (%)	Controls (n = 854) n (%)	P^*	Cases (n = 940) n (%)	Controls (n = 1061) n (%)	P^*
Age, y						
≤65	432 (57.4)	473 (55.4)	0.424	604 (64.3)	660 (62.2)	0.343
>65	321 (42.6)	381 (44.6)		336 (35.7)	401 (37.8)	
Sex						
Male	512 (68.0)	564 (66.0)	0.406	724 (77.0)	805 (75.9)	0.546
Female	241 (32.0)	290 (34.0)		216 (23.0)	256 (24.1)	
Tumor sites						
Cardia	295 (39.2)			318 (33.8)		
Noncardia	458 (60.8)			622 (66.2)		
Histological types						
Diffuse	437 (58.0)			541 (57.6)		
Intestinal	316 (42.0)			399 (42.4)		
Depth of invasion†						
T1	130 (17.3)			148 (15.7)		
T2	130 (17.3)			200 (21.3)		
T3	381 (50.6)			545 (58.0)		
T4	112 (14.8)			45 (4.8)		
Lymph node metastasis						
N0	297 (39.4)			374 (39.8)		
N1/N2/N3	456 (60.6)			566 (60.2)		
Distant metastasis						
M0	655 (87.0)			882 (93.8)		
M1	98 (13.0)			58 (6.2)		
TNM stages						
I	202 (26.8)			261 (27.8)		
II	165 (21.9)			186 (19.8)		
III	266 (35.3)			396 (42.1)		
IV	120 (16.0)			97 (10.3)		

TNM = tumor node metastasis.
* Two-sided χ^2 test for the frequency distributions of selected variables.
† Information of depth of invasion was not available for 2 patients in the validation set.

washed with Phosphate Buffered Saline and lysed with 1× passive lysis buffer. Luciferase activity was determined with a dual luciferase report assay system (Promega) following the manufacturer's protocol.

Electrophoretic Mobility Shift Assay (EMSA)

SGC-7901 cells ($2–4 \times 10^6$) were collected to prepare nuclear extracts by the NE-PER kit (Pierce; Rockford, IL). Biotin-labeled oligonucleotides probes were annealed at 50 fmol. Reactions were applied onto 5% polyacrylamide gels, and then transferred to nylon membranes. Biotin-labeled DNA was detected by the LightShift Chemiluminescent EMSA Kit (Pierce).

Statistical Analysis

In this study, statistical analyses were performed by SAS software (version 9.13; SAS Institute Inc, Cary, NC). All P values were 2-sided with $P < 0.05$. Goodness-of-fit χ^2 test was used to test for each SNP in control subjects to meet the Hardy–Weinberg equilibrium (HWE). Differences in the distributions of demographic characteristics, selected variables, and frequencies of genotypes between cases and controls were detected by Student t test or χ^2 test. Unconditional univariate and multivariate logistic regression analyses were done to estimate adjusted odds ratios (ORs) and 95% confidence intervals (95% CIs), which were used to evaluate the importance of genetic variants in different clinical features of GC.

RESULTS

Expression of miR-196a in Gastric Tumor and Adjacent Normal Tissues

MiR-196a expression levels were detected in 66 new diagnostic gastric tumor and adjacent normal tissue samples, which were chosen to keep the same genetic background. As shown in Figure 1, the average miR-196a level in tumor tissues was significantly higher than that in adjacent normal tissues (53.89 ± 3.991 vs 22.68 ± 2.568, $P < 0.001$).

Identification of SNPs in *MIR196A2* Gene Promoter Region

With regard to higher levels of miR-196a in GC tissues, we attempted to investigate the regulatory mechanism of miR-196a-2 expression. After screening the promoter region

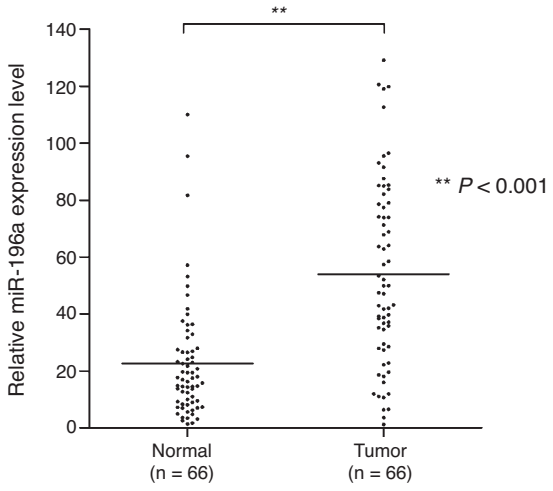

FIGURE 1. MiR-196a expression in 66 gastric cancer patients. MiR-196a levels in tumor tissues (T) were significantly higher than those in corresponding normal tissues (N) ($P < 0.001$).

of *MIR196A2* gene with SNP database (http://www.ncbi.nlm.nih.gov/snp), 2 SNPs within *MIR196A2* gene reveal their presence in Chinese population with MAF >5%, named rs12304647 A>C and rs35010275 G>C (Figure 2A).

Effects of SNPs in *MIR196A2* Promoter Region on Gene Transcriptional Regulation

To identify whether these 2 variations have influence on the transcription activity of *MIR196A2*, we constructed full-length luciferase reporter plasmid and its variants by using pGL3-basic vector, named pGL3–1000 wt, 1000mt1, 1000mt2, and 1000mt3, which were transfected into SGC-7901 and normal GES-1. As shown in Figure 2B, comparing with pGL3-1000 wt and pGL3-1000mt1, the transcription activity of pGL3-1000mt2 and pGL3-1000mt3 were significantly lower in SGC-7901 and GES-1, which suggested these SNPs in *MIR196A2* promoter region took effect on its transcription. To further confirm the individual functions of rs12304647 and rs35010275, 3 plasmids containing shorter *MIR196A* promoter regions (pGL3-700 wt, 700mt, and 500) were transfected in SGC-7901 cell line. All of these reporter plasmid assays displayed higher luciferase activities than pGL3-basic vector. pGL3-1000 wt and pGL3-1000mt1 showed significant increases in promoter activity than pGL3-700 wt and pGL3-700mt1. However, opposite trends displayed in pGL3-1000mt2 and pGL3-1000mt3 (Figure 2C). These results suggested that different allele of rs35010275 might possess different function in miR-196a expression.

Based on the above results, we further investigated potential regulatory role of rs35010275 G>C. EMSA was performed with probes carrying the SNP major (G) or minor (C) allele.

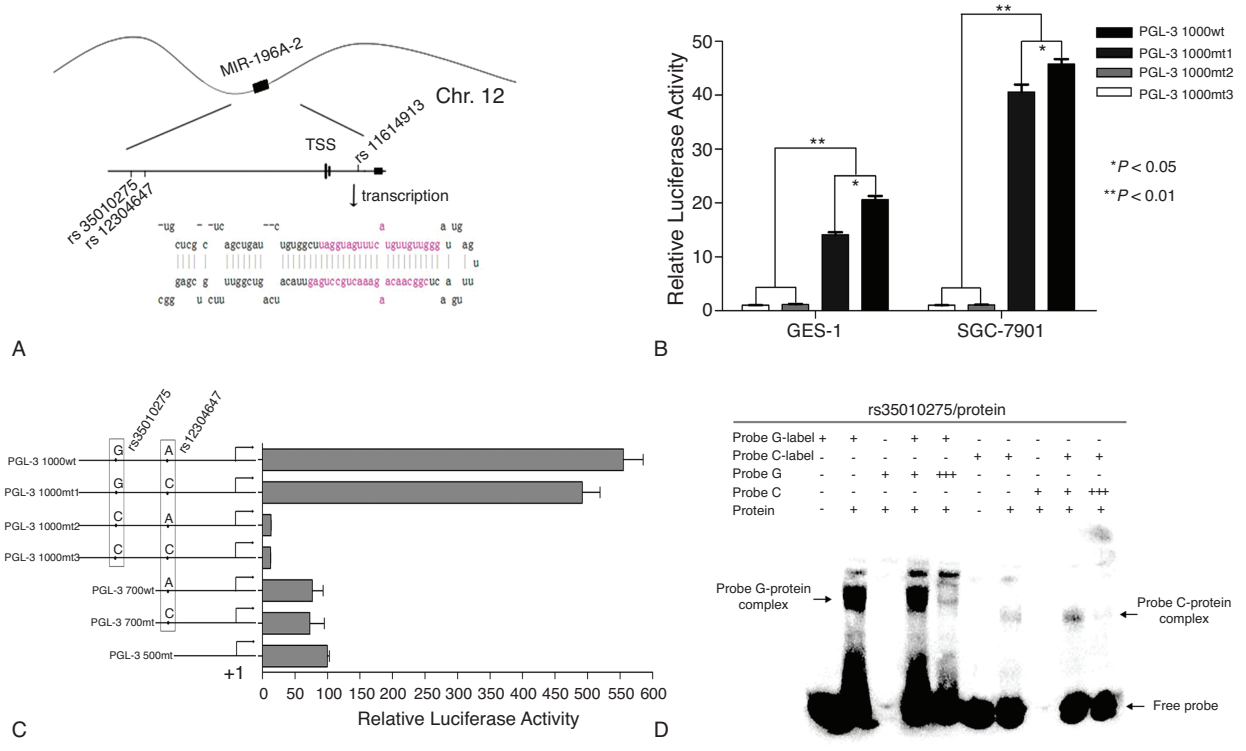

FIGURE 2. SNPs on promoter region regulated miR-196a-2 expression. (A) Schematic representation of *MIR196A* containing rs35010275 G>C and rs12304647 A>C, which were in upstream of mature miR-196a-2. (B) In both normal and cancerous cell lines, plasmid with rs35010275 G and rs12304647 A displayed a higher luciferase activity than other plasmids. Meanwhile, both plasmids with rs35010275 C allele possessed lower luciferase activities. (C) Luciferase reporter assay showed that rs35010275 G allele had significant higher promoter activity, while C allele of which had lower promoter activity. The graphs showed the value of luciferase activity from different plasmids containing different *MIR196A2* promoter haplotypes. (D) EMSA assays revealed G and C alleles recruited different proteins or complexes with different molecular weight. EMSA = electrophoretic mobility shift assay, SNPs = single-nucleotide polymorphisms.

Proteins or protein complexes binding to G-probe have larger molecular weight than those bind to C-probe, suggesting that they might be different (Figure 2D). Through competition assays between G and C-probes, it was obvious that both of the protein complexes displayed less intense bands as the decreasing concentration of both probes. These results suggested that these different protein complexes might contain sequence-specific DNA-binding domain that could recognize different alleles of rs35010275.

Characteristics of GC Patients and Controls

The frequency distributions of cases and controls in 2 stages are summarized in Table 1. The cases and controls appeared to be adequately matched on age ($P = 0.424$ and 0.343) and sex ($P = 0.406$ and 0.546). In testing set, 26.8%, 21.9%, 35.3%, and 16.0% of patients were in stage I, II, III, and IV disease, respectively, whereas, 27.8%, 19.8%, 42.1%, and 10.3% of the patients in validation set were diagnosed in stage I, II, III, and IV, respectively.

Association of rs35010275 G>C Polymorphism With MiR-196a-2 Levels in GC Tissues

Next, we performed genotyping analysis for rs35010275 in above 66 patients' tissues from which we successfully obtained both genomic DNA and RNA. Data were showed as following: GG, 36 cases (54.6%); GC, 24 cases (36.3%); and CC, 6 cases (9.1%). As presented in Figure 3, patients carrying GG genotype had higher miR-196a-2 expression levels than those with GC or CC genotype in both cancer tissues ($P = 0.032$ for GG vs GC and $P = 0.001$ for GG vs CC) and adjacent normal tissues ($P = 0.048$ for GG vs GC and $P = 0.049$ for GG vs CC).

Overall Effects of *MIR196A2* Promoter rs35010275 G>C on Risk of GC in 2-Stage Case–Control Studies

We further performed a 2-stage case–control study to evaluate the overall associations between rs35010275 G>C polymorphism and GC risk. Genotype frequencies in the combined controls was conformed to HWE ($\chi^2 = 0.885$, $P = 0.347$). As presented in Table 2, we found that GC/CC genotypes were associated with a significantly decreased risk of GC compared with GG genotype in both stage of the study (adjusted OR = 0.82, 95% CI = 0.67–0.99 in testing set and adjusted OR = 0.82, 95% CI = 0.68–0.97 in validation set). Moreover, with the expansion of samples in combined analysis, more significant protective effects displayed in GC, CC, and GC/CC genotypes (adjusted OR = 0.82, 95% CI = 0.72–0.95 for GC; adjusted OR = 0.76, 95% CI = 0.59–0.99 for CC; adjusted OR = 0.81, 95% CI = 0.71–0.93 for GC/CC) compared to GG genotype. Simultaneously, difference between the subjects carrying C allele and G allele was observed in allele comparison with adjusted OR = 0.85, 95% CI = 0.73–1.00 in testing set, adjusted OR = 0.85, 95% CI = 0.74–0.98 in validation set, and adjusted OR = 0.85, 95% CI = 0.77–0.94 in combined set.

Associations Between rs35010275 G>C Polymorphism and Clinical Features of GC

Next, we evaluated the effects of rs35010275 G>C on the progression of GC by comparing the rs35010275 G>C polymorphism with different clinical features. As shown in Table 3, with polymorphism in genetic dominant model of combined analysis, GC/CC genotypes have a statistical significantly decreased risk of GC compared with wild homozygous GG genotype in patients with cardiac cancer (adjusted OR = 0.71, 95% CI = 0.59–0.85), diffuse type (0.85, 0.73–1.00), intestinal type (0.77, 0.64–0.92), T3 depth invasion (0.74, 0.63–0.87),

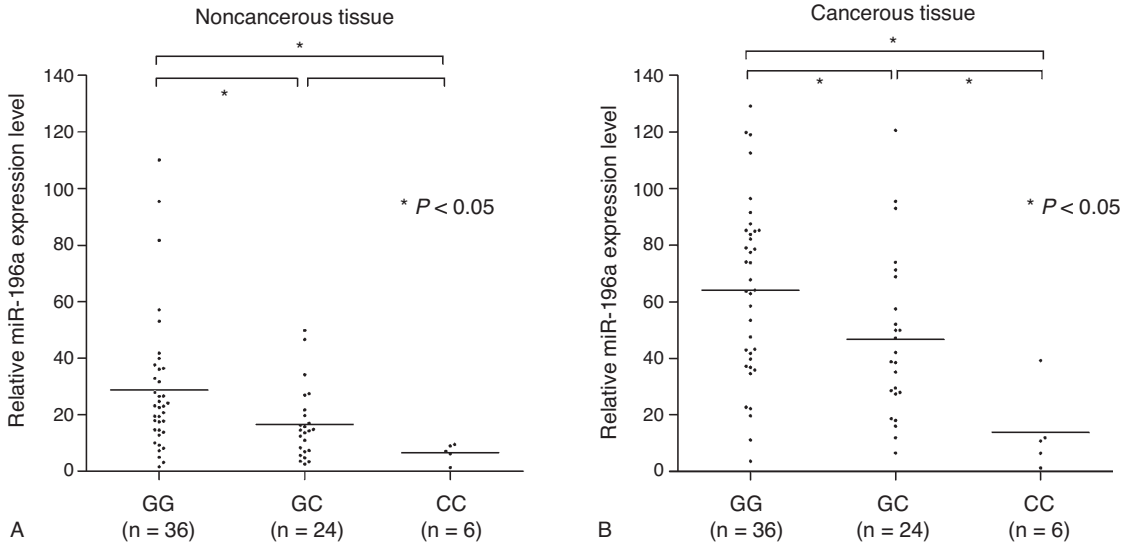

FIGURE 3. Associations between rs35010275 G>C polymorphism and miR-196a levels in 66 gastric cancer patients. (a) MiR-196a levels in noncancerous tissues according to different genotypes. Patients with GG genotype also showed significantly higher expression of miR-196a than those with GC and CC genotyeps ($P = 0.048$ and 0.049, respectively). (B) MiR-196a-2 levels in cancerous tissues according to different genotypes. Patients carrying GG genotype showed significantly higher miR-196a levels relative to those with GC and CC genotype ($P = 0.032$ and 0.001, respectively). Horizontal line is for the mean value of miR-196a levels in cancerous and noncancerous tissues, respectively.

TABLE 2. Genotype frequencies and overall effects of *MIR196A2* rs35010275 G>C polymorphism on gastric cancer risk

Genotypes	Testing set			Validation set			Combined set		
	Cases/Controls (753/854)	P^*	Adjusted OR (95% CI)†	Cases/Controls (940/1061)	P^*	Adjusted OR (95% CI)†	Cases/Controls (1693/1915)	P^*	Adjusted OR (95% CI)†
GG	436/452		1.00 (ref.)	532/548		1.00 (ref.)	968/1000		1.00 (ref.)
GC	264/331	0.069	0.82 (0.67-1.02)	348/432	**0.043**	**0.83 (0.69-0.99)**	612/763	**0.006**	**0.82 (0.72-0.95)**
CC	53/71	0.179	0.77 (0.53-1.13)	60/81	0.119	0.75 (0.53-1.08)	113/152	**0.040**	**0.76 (0.59-0.99)**
GC/CC	317/402	**0.042**	**0.82 (0.67-0.99)**	408/513	**0.023**	**0.82 (0.68-0.97)**	725/915	**0.002**	**0.81 (0.71-0.93)**
G allele	1136/1235		1.00 (ref.)	1412/1528		1.00 (ref.)	2548/2763		1.00 (ref.)
C allele	370/473	**0.046**	**0.85 (0.73-1.00)**	468/594	**0.023**	**0.85 (0.74-0.98)**	838/1067	**0.002**	**0.85 (0.77-0.94)**
P_{trend}		**0.049**			**0.027**			**0.003**	

CI = confidence interval, OR = odds ratio.
* Two-sided χ^2 test for genotype frequency distributions between cases and controls.
† Adjusted for age and sex in logistic regression model.

positive lymph node metastasis (0.73, 0.62–0.85), negative distant metastasis (0.82, 0.72–0.94), and advanced stage (0.72, 0.62–0.85).

DISCUSSION

In the present study, miR-196a acted as an oncogenic miRNA, which expressed significantly higher in cancerous tissues than in noncancerous tissues. Moreover, a notable SNP (rs35010275 G>C) in our study displayed potential transcript regulating function on *MIR196A2* and was associated with the susceptibility of GC in the Chinese population. These findings suggested that *MIR196A2* rs35010275 G>C might be regarded as a potential biomarker for GC prediction.

It has been reported that the miR-196a may display different effects as an oncogenic miRNA in the pathology of

TABLE 3. Stratified analyses of *MIR196A2* rs35010275 G>C polymorphism associated with gastric cancer progression by genetic dominant model

Variables	Testing set			Validation set			Combined set		
	GG/(GC/CC)	P^*	Adjusted OR (95% CI)†	GG/(GC/CC)	P^*	Adjusted OR (95% CI)†	GG/(GC/CC)	P^*	Adjusted OR (95% CI)†
Controls									
Total	452/402		1.00 (ref.)	548/513		1.00 (ref.)	1000/915		1.00 (ref.)
Cases									
Tumor sites									
Cardia	184/111	**0.005**	**0.68 (0.52-0.89)**	188/130	**0.015**	**0.73 (0.56-0.94)**	372/241	**<0.001**	**0.71 (0.59-0.85)**
Non-cardia	252/206	0.400	0.91 (0.72-1.14)	344/278	0.136	0.86 (0.70-1.05)	596/484	0.096	0.88 (0.76-1.02)
Histological types									
Diffuse	250/187	0.148	0.84 (0.67-1.06)	299/242	0.157	0.86 (0.70-1.06)	549/429	**0.043**	**0.85 (0.73-1.00)**
Intestinal	186/130	0.065	0.78 (0.60-1.02)	233/166	**0.020**	**0.76 (0.60-0.96)**	419/296	**0.003**	**0.77 (0.64-0.92)**
Depth of invasion‡									
T1	72/58	0.560	0.90 (0.62-1.30)	80/68	0.685	0.93 (0.66-1.31)	152/126	0.482	0.91 (0.71-1.18)
T2	69/61	0.992	1.00 (0.69-1.44)	107/93	0.696	0.94 (0.70-1.27)	176/154	0.782	0.97 (0.77-1.22)
T3	231/150	**0.010**	**0.72 (0.56-0.92)**	319/226	**0.009**	**0.76 (0.62-0.93)**	550/376	**<0.001**	**0.74 (0.63-0.87)**
T4	64/48	0.369	0.83 (0.56-1.24)	26/19	0.635	0.87 (0.48-1.56)	90/67	0.291	0.84 (0.60-1.16)
Lymph node metastasis									
N0	161/136	0.684	0.95 (0.73-1.23)	195/179	0.829	0.97 (0.77-1.23)	356/315	0.664	0.96 (0.81-1.15)
N1/N2/N3	275/181	**0.009**	**0.74 (0.58-0.93)**	337/229	**0.002**	**0.72 (0.59-0.89)**	612/410	**<0.001**	**0.73 (0.62-0.85)**
Distant metastasis									
M0	382/273	**0.034**	**0.80 (0.65-0.98)**	493/389	0.053	0.84 (0.70-1.00)	875/662	**0.004**	**0.82 (0.72-0.94)**
M1	54/44	0.635	0.90 (0.59-1.38)	39/19	**0.027**	**0.53 (0.30-0.93)**	93/63	0.079	0.74 (0.53-1.04)
TNM stages									
I+II	203/164	0.429	0.90 (0.71-1.16)	237/210	0.546	0.93 (0.75-1.17)	440/374	0.330	0.92 (0.78-1.09)
III+IV	233/153	**0.013**	**0.73 (0.57-0.94)**	295/198	**0.003**	**0.72 (0.58-0.89)**	528/351	**<0.001**	**0.72 (0.62-0.85)**

CI = confidence interval, OR = odds ratio, TNM = tumor node metastasis.
* Two-sided χ^2 test for the frequency distributions of selected variables between cases and controls.
† Adjusted for age and sex in logistic regression model.
‡ Information was not available for 2 patients in validation set and combined set.

GC, after quantitatively assessing the levels of miR-196a in patients.[16,30] Our study also confirmed that miR-196a unduly existed in GC tissues, which were consistent with the results reported by Sun et al.[40] Besides, the genotyping of the 66 patients also manifested that patients carrying rs35010275 C allele were correlated with lower expression level of miR-196a in both tumor and normal tissues. These evidences suggested that the increase of miR-196a expression in rs35010275 GG genotype was a crucial event in gastric carcinogenesis; while some other transcriptional or post-transcriptional mechanism might exist in normal tissues to protect the abnormal expression of miR-196a.

It was proposed that various genomic and epigenomic mechanisms were involved in dysregulation of miRNA genes in cancers. The study by Croce[41] had concretely demonstrated these different causes such as deletions, amplifications, mutations, epigenetic silencing, and even the aberrant expression of transcription factors targeting miRNAs. To date, compared with thorough disclosure of miR-196a functions, the culprits of abnormal expression of miR-196a have been involved seldom. Derived from 2 different loci, nominated *MIR196A1* and *MIR196A2* gene, detailed mechanism of miR-196a expression may be diverse and complicated. With regard to affecting the mature process of miR-196a-2 transcription,[23,42,43] rs11614913 SNP was one exploration raising the concerns of researchers.[20,42,44,45] Recent studies have identified this polymorphism as a risk factor for several diseases, including hepatocellular carcinoma,[20] leukemia,[46] lung cancer,[22,23] breast cancer,[44,47] colorectal cancer,[32] and pancreatic adenocarcinoma.[48,49] Apart from *MIR196A1*, *MIR196A2* possesses its own promoter in structure, in which the variations might also alter miR-196a expression and then influence the susceptibility of GC. In this study, we identified 2 SNPs (rs12304647 A>C and rs35010275 G>C) in *MIR196A2* promoter region of 1000 bp away from the transcription starting site. Taken functional analyses together, these observations supported the idea that *MIR196A2* rs35010275 G>C resulted in different recruits of transcription factors or complexes to *MIR196A2* gene promoter region, which modulated the miR-196a expression. To our knowledge, this is the first study providing direct evidence that *MIR196A2* promoter polymorphism may influence individuals' susceptibility to GC through affecting miRNA biogenesis.

MiRNAs potentially display their distinct effects in various biologic processes via influencing their target genes. As dysregulation of miR-196a may contribute to tumor detachment, migration, invasion, and proliferation[13,16,32,40,50,51] through regulating its fundamental target genes, it is biologically plausible that rs35010275 G>C influence the development of GC. Our 2-stage epidemiological study indicated that rs35010275 C allele displayed protective effect on GC with T3 depth of invasion, positive lymph node metastasis, negative distant metastasis, and advanced TNM stage. Therefore, *MIR196A2* may participate in the invasion and migration pathways of gastric carcinogenesis, and rs35010275 C allele may act as a repressor in hazardous biological processes. In previous studies, the members of homeobox Gene (HOX) family[15,50,52] and annexin A1[13,53] have been reported as target genes of miR-196a in vivo. Schimanski et al[32] identified that through regulating HOX family genes, miR-196a could activate the v-akt murine thymoma viral oncogene (AKT) signaling pathway by increasing phosphorylation of AKT. These results suggested that *MIR196A2* rs35010275 could trigger the expression of miR-196a and finally induce the AKT pathway to influence susceptibility and processes of gastric carcinogenesis.

Considering rs35010275 G>C as an molecular biomarker to predict individuals' susceptibility to GC, our study also showed that the protective effects of rs35010275 C allele were more predominant in cardia patients (adjust OR = 0.71, 95% CI = 0.59–0.85, $P < 0.001$) and stages III and VI patients (adjust OR = 0.72, 95% CI = 0.62–0.85, $P < 0.001$), suggesting that there exists associations of rs35010275 G>C with tumor site and advanced TNM stage. Further investigation to validate this association is necessary, and it would be interesting to clarify whether the carcinogenesis mechanism in cancer with low TNM stage might be distinct from that in cancer with advanced TNM stage. Besides, *MIR196A2* rs35010275 G>C might majorly take its contribution on precluding the onset of cancer in cardia.

Several limitations exist in our study. First, small sample size may limit the statistical power of our study, especially for subgroup analyses. Second, our study was retrospective hospital-based case–control studies; the inherent selection bias and information bias were unavoidable. Under this circumstance, we applied a rigorous epidemiological design in selecting study subjects and conducted statistical adjustment for known risk factors to minimize the potential biases. Finally, our study lacks some information on GC risk factors, such as *Helicobacter pylori* infection status, diet habit, tobacco smoking, and alcohol consumption status. Therefore, our results need to be validated in population-based studies with larger sample size and more detailed information.

In conclusion, our study revealed that miR-196a was dysregulated in GC. Moreover, a functional SNP rs35010275 G>C in the promoter region of *MIR196A2* gene was significantly associated with miR-196a-2 expression. Further insights into functional and clinical investigation of *MIR196A2* gene may contribute to the diagnosis and prognosis of GC.

ACKNOWLEDGMENT

The authors would like to appreciate the technical help of Biolight Tech Company, Nanjing, China.

REFERENCES

1. Danaei G, Vander Hoorn S, Lopez AD, et al. Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors. *Lancet.* 2005;366:1784–1793.
2. Ferlay J, Shin HR, Bray F, et al. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer.* 2010;127:2893–2917.
3. Ajani JA. Optimizing docetaxel chemotherapy in patients with cancer of the gastric and gastroesophageal junction: evolution of the docetaxel, cisplatin, and 5-fluorouracil regimen. *Cancer.* 2008;113:945–955.
4. Yoong J, Michael M, Leong T. Targeted therapies for gastric cancer: current status. *Drugs.* 2011;71:1367–1384.
5. Field K, Michael M, Leong T. Locally advanced and metastatic gastric cancer: current management and new treatment developments. *Drugs.* 2008;68:299–317.
6. Matsusaka K, Kaneda A, Nagae G, et al. Classification of Epstein-Barr virus-positive gastric cancers by definition of DNA methylation epigenotypes. *Cancer Res.* 2011;71:7187–7197.
7. Rusiecki JA, Al-Nabhani M, Tarantini L, et al. Global DNA methylation and tumor suppressor gene promoter methylation and gastric cancer risk in an Omani Arab population. *Epigenomics.* 2011;3:417–429.

8. Xu L, Li X, Chu ES, et al. Epigenetic inactivation of BCL6B, a novel functional tumour suppressor for gastric cancer, is associated with poor survival. *Gut.* 2012;61:977–985.
9. Li X, Zhang Y, Ding J, et al. Survival prediction of gastric cancer by a seven-microRNA signature. *Gut.* 2010;59:579–585.
10. Gao C, Zhang Z, Liu W, et al. Reduced microRNA-218 expression is associated with high nuclear factor kappa B activation in gastric cancer. *Cancer.* 2010;116:41–49.
11. Craig VJ, Cogliatti SB, Rehrauer H, et al. Epigenetic silencing of microRNA-203 dysregulates ABL1 expression and drives Helicobacter-associated gastric lymphomagenesis. *Cancer Res.* 2011;71:3616–3624.
12. Ambros V. The functions of animal microRNAs. *Nature*. 2004;431:350–355.
13. Luthra R, Singh RR, Luthra MG, et al. MicroRNA-196a targets annexin A1: a microRNA-mediated mechanism of annexin A1 downregulation in cancers. *Oncogene.* 2008;27:6667–6678.
14. Mansfield JH, Harfe BD, Nissen R, et al. MicroRNA-responsive ''sensor'' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. *Nat Genet.* 2004;36:1079–1083.
15. Niinuma T, Suzuki H, Nojima M, et al. Upregulation of miR-196a and HOTAIR drive malignant character in gastrointestinal stromal tumors. *Cancer Res.* 2012;72:1126–1136.
16. Tsai KW, Liao YL, Wu CW, et al. Aberrant expression of miR-196a in gastric cancers and correlation with recurrence. *Genes, Chromosomes Cancer.* 2012;51:394–401.
17. Lung RW, Wang X, Tong JH, et al. A single nucleotide polymorphism in microRNA-146a is associated with the risk for nasopharyngeal carcinoma. *Mol Carcinog.* 2013;52:E28–E38.
18. Smith RA, Jedlinski DJ, Gabrovska PN, et al. A genetic variant located in miR-423 is associated with reduced breast cancer risk. *Cancer Genomics Proteomics.* 2012;9:115–118.
19. Zhang M, Jin M, Yu Y, et al. Associations of miRNA polymorphisms and female physiological characteristics with breast cancer risk in Chinese population. *Eur J Cancer Care.* 2012;21:274–280.
20. Li XD, Li ZG, Song XX, et al. A variant in microRNA-196a2 is associated with susceptibility to hepatocellular carcinoma in Chinese patients with cirrhosis. *Pathology.* 2010;42:669–673.
21. Xu Y, Liu L, Liu J, et al. A potentially functional polymorphism in the promoter region of miR-34b/c is associated with an increased risk for primary hepatocellular carcinoma. *Int J Cancer.* 2011;128:412–417.
22. Hu Z, Chen J, Tian T, et al. Genetic variants of miRNA sequences and non-small cell lung cancer survival. *J Clin Investig.* 2008;118:2600–2608.
23. Tian T, Shu Y, Chen J, et al. A functional genetic variant in microRNA-196a2 is associated with increased susceptibility of lung cancer in Chinese. *Cancer Epidemiol, Biomarkers Prevention.* 2009;18:1183–1187.
24. Christensen BC, Avissar-Whiting M, Ouellet LG, et al. Mature microRNA sequence polymorphism in MIR196A2 is associated with risk and prognosis of head and neck cancer. *Clin Cancer Res.* 2010;16:3713–3720.
25. Mu YP, Su XL. Polymorphism in pre-miR-30c contributes to gastric cancer risk in a Chinese population. *Med Oncol.* 2012;29:1723–1732.
26. Kogo R, Mimori K, Tanaka F, et al. Clinical significance of miR-146a in gastric cancer cases. *Clin Cancer Res.* 2011;17:4277–4284.
27. Wang S, Tao G, Wu D, et al. A functional polymorphism in MIR196A2 is associated with risk and prognosis of gastric cancer. *Mol Carcinog.* 2013;52(Suppl 1):E87–E95.
28. Okubo M, Tahara T, Shibata T, et al. Association between common genetic variants in pre-microRNAs and gastric cancer risk in Japanese population. *Helicobacter.* 2010;15:524–531.
29. Peng S, Kuang Z, Sheng C, et al. Association of microRNA-196a-2 gene polymorphism with gastric cancer risk in a Chinese population. *Dig Dis Sci.* 2010;55:2288–2293.
30. Inoue T, Iinuma H, Ogawa E, et al. Clinicopathological and prognostic significance of microRNA-107 and its relationship to DICER1 mRNA expression in gastric cancer. *Oncol Rep.* 2012;27:1759–1764.
31. Ahmed FE, Ahmed NC, Vos PW, et al. Diagnostic microRNA markers to screen for sporadic human colon cancer in stool: I. Proof of principle. *Cancer Genomics Proteomics.* 2013;10:93–113.
32. Schimanski CC, Frerichs K, Rahman F, et al. High miR-196a levels promote the oncogenic phenotype of colorectal cancer cells. *World J Gastroenterol.* 2009;15:2089–2096.
33. Liu G, Gramling S, Munoz D, et al. Two novel BRM insertion promoter sequence variants are associated with loss of BRM expression and lung cancer risk. *Oncogene.* 2011;30:3295–3304.
34. Hao B, Miao X, Li Y, et al. A novel T-77C polymorphism in DNA repair gene XRCC1 contributes to diminished promoter activity and increased risk of non-small cell lung cancer. *Oncogene.* 2006;25:3613–3620.
35. Wang M, Yuan L, Wu D, et al. A novel XPF −357A>C polymorphism predicts risk and recurrence of bladder cancer. *Oncogene.* 2010;29:1920–1928.
36. Ozsolak F, Poling LL, Wang Z, et al. Chromatin structure analyses identify miRNA promoters. *Genes Dev.* 2008;22:3172–3183.
37. Wang M, Bai J, Tan Y, et al. Genetic variant in PSCA predicts survival of diffuse-type gastric cancer in a Chinese population. *Int J Cancer.* 2011;129:1207–1213.
38. 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 Pathologica Microbiologica Scandinavica.* 1965;64:31–49.
39. Green FL, Page DL, Fleming ID, et al. Stomach of the digestive system. *AJCC Cancer Staging Handbook*. 6th ed. New York: Springer Press; 2002:111–118.
40. Sun M, Liu XH, Li JH, et al. MiR-196a is upregulated in gastric cancer and promotes cell proliferation by downregulating p27(kip1). *Mol Cancer Therap.* 2012;11:842–852.
41. Croce CM. Causes and consequences of microRNA dysregulation in cancer. *Nat Rev Genet.* 2009;10:704–714.
42. Xu J, Hu Z, Xu Z, et al. Functional variant in microRNA-196a2 contributes to the susceptibility of congenital heart disease in a Chinese population. *Hum Mutat.* 2009;30:1231–1236.
43. Hu Z, Liang J, Wang Z, et al. Common genetic variants in pre-microRNAs were associated with increased risk of breast cancer in Chinese women. *Hum Mutat.* 2009;30:79–84.
44. Hoffman AE, Zheng T, Yi C, et al. MicroRNA miR-196a-2 and breast cancer: a genetic and epigenetic association study and functional analysis. *Cancer Res.* 2009;69:5970–5977.
45. Catucci I, Yang R, Verderio P, et al. Evaluation of SNPs in miR-146a, miR196a2 and miR-499 as low-penetrance alleles in German and Italian familial breast cancer cases. *Hum Mutat.* 2010;31:E1052–E1057.
46. Coskun E, von der Heide EK, Schlee C, et al. The role of microRNA-196a and microRNA-196b as ERG regulators in acute myeloid leukemia and acute T-lymphoblastic leukemia. *Leuk Res.* 2011;35:208–213.

47. Hui AB, Shi W, Boutros PC, et al. Robust global micro-RNA profiling with formalin-fixed paraffin-embedded breast cancer tissues. *Lab Investig.* 2009;89:597–606.

48. Kong X, Du Y, Wang G, et al. Detection of differentially expressed microRNAs in serum of pancreatic ductal adenocarcinoma patients: miR-196a could be a potential marker for poor prognosis. *Dig Dis Sci.* 2011;56:602–609.

49. Szafranska-Schwarzbach AE, Adai AT, Lee LS, et al. Development of a miRNA-based diagnostic assay for pancreatic ductal adenocarcinoma. *Expert Rev Mol Diagn.* 2011;11:249–257.

50. Braig S, Mueller DW, Rothhammer T, et al. MicroRNA miR-196a is a central regulator of HOX-B7 and BMP4 expression in malignant melanoma. *Cell Mol Life Sci.* 2010;67:3535–3548.

51. Kim YJ, Bae SW, Yu SS, et al. MiR-196a regulates proliferation and osteogenic differentiation in mesenchymal stem cells derived from human adipose tissue. *J Bone Miner Res.* 2009;24:816–825.

52. Mueller DW, Bosserhoff AK. MicroRNA miR-196a controls melanoma-associated genes by regulating HOX-C8 expression. *Int J Cancer.* 2011;129:1064–1074.

53. Kan T, Meltzer SJ. MicroRNAs in Barrett's esophagus and esophageal adenocarcinoma. *Curr Opin Pharmacol.* 2009;9:727–732.