The \textit{E-cadherin} $-347G\rightarrow GA$ promoter polymorphism and its effect on transcriptional regulation

Yong Shin$^{1,5}$, Il-Jin Kim$^{1,5}$, Hio Chung Kang$^{1}$, Jae-Hyun Park$^{1}$, Hye-Rin Park$^{1}$, Hye-Won Park$^{1}$, Mi Ae Park$^{2}$, Jong Soo Lee$^{2}$, Kyong-Ah Yoon$^{1}$, Ja-Lok Ku$^{1}$ and Jae-Gahb Park$^{1,3}$

$^1$Korean Hereditary Tumor Registry, Cancer Research Center and Cancer Research Institute, Seoul National University and $^2$Research Institute and Hospital, National Cancer Center, Goyang, Gyeonggi, South Korea

$^3$To whom correspondence should be addressed at: National Cancer Center, 809, Maud-dong, Ilsan-gu, Goyang, Gyeonggi, South Korea

E-cadherin plays a critical role in epithelial cell-cell adhesion and maintenance of tissue architecture. Loss of E-cadherin expression in humans has been associated with cancer, and a number of cancer-related mutations have been identified. Here, we sought to investigate whether the $-347G\rightarrow GA$ single nucleotide polymorphism affects the transcriptional activity of the \textit{E-cadherin} gene. First, we measured the promoter activity of the $-347G\rightarrow GA$ polymorphism using a dual luciferase reporter assay and electrophoretic mobility shift assay (EMSA). The dual luciferase reporter assay showed that the GA allele decreased the transcriptional efficiency by 10-fold ($P < 0.001$) compared with the G allele. Similarly, EMSA revealed that the GA allele had a weak transcription factor binding strength compared with the G allele. We then examined the frequency of this polymorphism in familial gastric cancer (FGC) patients by denaturing high-performance liquid chromatography. We found that the \textit{E-cadherin} genotype ($-347G/G$ heterozygous or $G$ homozygous) was associated with FGC patients ($P < 0.05$) compared with the $G$ homozygous genotype. Taken together, these results suggest that the GA allele may cause weak transcription factor binding affinity and low transcriptional activity in E-cadherin expression.

Introduction

Gastric cancer is one of the most common cancers worldwide. Although the occurrence rate of gastric cancer has decreased in recent years, the incidence of the disease is still high in Asian countries such as Korea and Japan (1). However, relatively little is known regarding genetic susceptibility in the pathogenesis of gastric cancer (2). Mutations in the calcium-dependent cell adhesion molecule, E-cadherin, have been associated with the early development of gastric cancer (3). \textit{E-cadherin} germline mutations were first identified in New Zealand Maori families with early-onset diffuse gastric cancer; since then, the majority of \textit{E-cadherin} germline mutations have been reported in diffuse type gastric cancer (4–7). Recently, we reported a \textit{MET} germline mutation as well as \textit{E-cadherin} germline mutations in the diffuse type of familial gastric cancer (F GC) (6,8). E-cadherin is found predominantly in epithelial cells and plays a pivotal role in maintaining tissue integrity (9). A large number of reports have identified down-regulation of E-cadherin expression in human carcinomas, and E-cadherin function is lost during the development of most epithelial cancers. Indeed, it is thought that loss of E-cadherin function in cancer cells probably plays an important role in tumor development and progression (10). However, it is not yet understood how these losses of expression are governed. Just as nucleotide variations in the coding region of a gene can alter protein function, polymorphisms within the 5'-promoter region have been known to change the transcriptional efficiency of a variety of genes (11,12). Recently, two frequent polymorphisms in human cancers have been identified in the promoter region of the \textit{E-cadherin} gene. The first is a C$\rightarrow$A single nucleotide polymorphism (SNP) $-160$ nt from the transcriptional start site of the \textit{E-cadherin} gene promoter; transcription of the A allele is $68\%$ less efficient than that of the C allele (13). The second reported promoter variant is a G$\rightarrow$GA SNP $-347$ nt from the transcriptional start site of the \textit{E-cadherin} gene. The original report suggested that this polymorphism had no effect on transcriptional activity (14). In this study, we sought to better understand the mechanisms of altered \textit{E-cadherin} expression by investigating $-347 G\rightarrow GA$ polymorphism effects on \textit{E-cadherin} transcriptional activity.

Materials and methods

\textbf{DNA isolation from blood samples}

Blood samples of 28 cases from 27 FGC families without germline mutations in the \textit{E-cadherin} coding sequence (8) and 142 normal control individuals were collected from the Seoul National University Hospital, South Korea. Informed consent was obtained from all participants prior to testing. Twenty-seven Korean families affected with familial gastric cancer were investigated for genotyping of $-347G\rightarrow GA$ promoter polymorphism of \textit{E-cadherin} gene. Criteria for family inclusion were at least two first or second degree relatives affected with gastric cancer, at least one of whom was diagnosed with cancer before the age of 50 (8). Out of 27 probands (range 22–69 ages), 12 represented families suffering from diffuse types of gastric cancer, four represented families suffering from intestinal types and histological data for the type of the remaining 11 were not available. The classification of hereditary diffuse gastric cancer or hereditary intestinal gastric cancer was not possible in these families owing to the lack of histological information. The normal control population was randomly selected from 142 healthy Korean individuals. Peripheral blood lymphocytes were isolated from samples using Ficoll-Paque (Pharmacia Biotech, Uppsala, Sweden) according to the manufacturer’s instructions. Total genomic DNA was extracted with the Trizol reagent (Invitrogen, Carlsbad, CA) according to the manufacturer’s instructions.

\textbf{DNA analysis of the \textit{E-cadherin} promoter regions}

We screened each of the above samples for the $-347G\rightarrow GA$ \textit{E-cadherin} polymorphism using denaturing high-performance liquid chromatography (DHPLC) (WAVE$^\circledR$, Transgenic, Omaha, NE). DNA fragments containing

Abbreviations: DHPLC, denaturing high-performance liquid chromatography; EMSA, electrophoretic mobility shift assay; FGC, familial gastric cancer; PCR-RFLP, polymerase chain reaction-restriction fragment length polymorphism; SNP, single nucleotide polymorphism.

\text*{These authors contributed equally to this work.}
the promoter region of interest were amplified with the following primers: forward, 5'-CGCCCCAGCTTGTCTCTTCAC-3'; reverse, 5'-GGGACAGC-GAATCAGCA-3'. PCR amplification for DHPLC analysis was carried out in a volume of 25 µl containing 100 ng genomic DNA, 10 pmol of each primer, 0.25 mM each dNTP, 0.5 U of Taq polymerase and the provided reaction buffer (GeneCraft, Munster, Germany). Reactions were carried out in a programmable thermal cycler (MWG Biotech AG, Ebersberg, Germany) as follows: denaturation for 5 min at 94°C, followed by five cycles of 94°C for 30 s, 65°C for 30 s, 72°C for 1 min, then 30 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 1 min, followed by five cycles of 94°C for 30 s, 55°C for 30 s, 72°C for 1 min. 

**Results**

**Effect of the −347G→GA polymorphism on promoter activity**

To examine the effect of −347G→GA promoter region polymorphism on transcription of the E-cadherin gene, we measured promoter activity with a Dual Luciferase Reporter Assay System (Promega) and compared the activities of the −347GA and −347G alleles by transient transfection assay in CV-1, HeLa, SNU-719, AGS and KatoIII cells. As shown in Figure 1, significantly lower luciferase activities were generated by the pGL3-GA construct as compared with the pGL3-G construct. In CV-1 cells, the GA allele decreased the transcriptional efficiency by 10-fold (P = 0.000261) compared with the G allele. Similar results were obtained in HeLa, SNU-719, AGS and KatoIII cells (12-, 8-, 9-, 13-fold decrease, respectively).

**Effect of the −347G→GA polymorphism on the binding activity of nuclear factors**

To determine whether the −347G→GA polymorphism affects the binding activity of nuclear factors, synthetic −347G and −347GA oligonucleotides were incubated with HeLa cell nuclear extracts and subjected to EMSA. The −347GA oligonucleotide showed weak DNA-protein binding, whereas the −347GA oligonucleotide showed stronger DNA-protein binding (Figure 2). To verify the DNA-protein complex, competition assays were performed with specific and non-specific oligonucleotides (Figure 2). When a −347G oligonucleotide was used to compete with a −347GA oligonucleotide, it totally disrupted the −347GA oligonucleotide binding with nuclear protein. However, when a −347GA oligonucleotide was used to compete with a −347GA oligonucleotide, it was not as effective as a −347G oligonucleotide in disrupting oligonucleotide binding with nuclear protein.

**Allele frequencies in FGC samples and normal controls**

To determine whether there is a correlation between the promoter polymorphisms and FGC, we screened a 448 bp region (−529 to −82 from the transcriptional start site) of the E-cadherin promoter in 28 cases of FGC and 142 normal controls using PCR–RFLP or DHPLC (Figure 3). We identified the previously reported −160C→A and −347G→GA polymorphisms (13,14), and noted a positive association between the −347GA allele and FGC. Eleven (39.4%) of 28 FGC samples were heterozygous at this locus, as compared with 39 (27.5%) of 142 normal controls. Individuals with the E-cadherin genotype (−347GA heterozygous or GA homozygous) had an increased risk (P = 0.03059) for FGC (Table I). The distribution of genotypes was in Hardy-Weinberg equilibrium. In contrast, we did not identify a positive or negative association between the −160C→A polymorphism and FGC (Table I).

**Discussion**

In the present study, we focused on the effect of the −347G→GA E-cadherin polymorphism on transcriptional activity. Several major cis-acting elements have been identified within a short section of the proximal promoter of the E-cadherin gene, including two E boxes, a CAAT box, and a
GC-rich box SP1 binding site (14,18). The E-cadherin gene promoter thus exhibits a modular structure, suggesting that the strict control of epithelium-specific E-cadherin expression might result from interactions among the various regulatory elements (19). Our results demonstrate that the −347 SNP has a significant effect on transcriptional activity in transient transfection experiments. We performed transient transfections in CV-1, HeLa, SNU-719, AGS and KatoIII cells, because it was

---

**Fig. 1.** Dual luciferase reporter assay of the −347G→GA polymorphism. The human E-cadherin gene promoter (nucleotide −647 to +147 relative to the transcription initiation site) was cloned from homozygous (G) and heterozygous (GA) FGC patients. The fragment was inserted upstream of the luciferase reporter gene in plasmid pGL3 and transiently transfected into CV-1 (A), HeLa (B), SNU-719 (C), AGS (D) and KatoIII (E) cells. The luciferase activity of each construct was normalized against the activity of Renilla luciferase. Data are expressed as a percentage of the corrected luciferase activity of pGL3-control (bars indicate the means of three independent experiments).

**Fig. 2.** EMSA with HeLa nuclear extract using −347G and −347GA oligonucleotides. Binding activity of the −347G and −347GA oligonucleotides. The assay was performed in the presence (+) or absence (−) of HeLa nuclear extract. Competition assays were performed with unlabeled −347G or −347GA oligonucleotides. Each binding reaction contained 5 μg of HeLa nuclear extract and labeled −347G (lanes 2–6) or −347GA (lanes 8–12) oligonucleotides. Excess unlabeled −347G or −347GA oligonucleotides (10-, 50- and 100-fold) were included in the binding reactions as competitor (lanes 3–5 and 9–11, respectively). In addition, 100-fold excesses of unlabeled −347GA and −347G oligonucleotides were used to compete with −347G (lane 6) and −347GA (lane 12) oligonucleotides. Arrows indicate DNA–protein complexes. See online supplementary material for colour version of this figure.

---
reported previously that the $-347$ G—GA polymorphism of $E$-cadherin gene had no effect on transcriptional activity in CV-1 cells (14). In contrast to previous reported results, our study showed that in CV-1 cells, the GA allele of this polymorphism decreased the transcriptional efficiency by 10 fold ($P < 0.000261$) compared with the G allele (Figure 1). Similar results were obtained in HeLa, SNU-719, AGS and KatoIII cells. The molecular mechanism of this difference may relate to differences in the affinity of DNA-binding proteins to the two alleles of the $E$-cadherin promoter. We searched for putative transcriptional factors that might bind with the $-347$ SNP, using the Ds gene software package (Accelrys, San Diego, CA). We identified four putative transcription factors (Site_C2, ZESTE_CS, T-Ag-SV40.3, T-Ag-EP) with similarities to sequences near the $E$-cadherin $-347$G—GA promoter polymorphism. These putative transcription factors are not well characterized, and their in-depth study may be a target of future work. To investigate binding between the alleles and nuclear factors in general, we performed EMSA, which revealed that the $-347$GA allele had a weak transcriptional factor binding strength compared with the G allele. Therefore, it seems that

![Fig. 3. $E$-cadherin $-347$ G—GA polymorphism. (A) $-347$G (containing the G allele); (B) $-347$GA (containing the GA allele). The underline denotes the SNP site; arrow indicates DHPLC chromatogram and matched sequencing chromatogram. See online supplementary material for colour version of this figure.](https://academic.oup.com/carcin/article-abstract/25/6/895/2390681)

<table>
<thead>
<tr>
<th>Polymorphisms</th>
<th>Normal ($n = 142$)</th>
<th>FGC ($n = 28$)</th>
<th>OR* (95% CI*)</th>
<th>$\chi^2$</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-347$ G</td>
<td>103 (72.5%)</td>
<td>16 (57.1%)</td>
<td>1.815 (0.7-4.2)</td>
<td>6.97</td>
<td>0.03059</td>
</tr>
<tr>
<td>G/GA</td>
<td>39 (27.5%)</td>
<td>11 (39.3%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td>0 (0%)</td>
<td>1 (3.6%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$-160$ C</td>
<td>31 (21.8%)</td>
<td>6 (21.4%)</td>
<td>1.01 (0.3-2.7)</td>
<td>1.65</td>
<td>0.43726</td>
</tr>
<tr>
<td>C/A</td>
<td>110 (77.5%)</td>
<td>21 (75%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$-160$ A</td>
<td>1 (0.7%)</td>
<td>1 (3.6%)</td>
<td>5.2 (0.3-87.1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Odds ratio compared with homozygous individuals (GG or CC).

**95% CI; OR and 95% CIs were calculated by logistic regression with the GG or CC genotype as the reference group.**

Allele frequency.

In summary, we investigated the importance of the $-347$ G—GA polymorphism of $E$-cadherin gene. The GA allele was associated with significant suppression of $E$-cadherin transcription in CV-1, HeLa, SNU-719, AGS and KatoIII cells. Additionally, EMSA revealed that the GA allele had a weak transcriptional factor binding strength compared with the G allele. Therefore, it seems that
–347G→GA polymorphism may affect the expression of E-cadherin, possibly increasing the cancer risk.

Supplementary material
Supplementary material can be found at: http://www.carcin.oupjournals.org/

Acknowledgements
This work was supported by a research grant from the National Cancer Center, Korea, grant HMP-99-M-03-0001 from the Ministry of Health and Welfare, and the BK21 project for Medicine, Dentistry and Pharmacy.

References

Received August 20, 2003; revised November 16, 2003; accepted December 19, 2003