Advances in Brief

A Single Nucleotide Polymorphism in the E-cadherin Gene Promoter Alters Transcriptional Activities

Long-Cheng Li, Richard M. Chui, Masahiro Sasaki, Koichi Nakajima, Geetha Perinchery, Harry C. Au, Dana Nojima, Peter Carroll, and Rajvir Dahiya

Department of Urology, Veterans Affairs Medical Center, and University of California San Francisco, San Francisco, California 94121

Abstract

E-cadherin plays a critical role in many aspects of cell adhesion, epithelial development, and the establishment and maintenance of epithelial polarity. The loss of the adhesive function of E-cadherin is a critical step in the promotion of epithelial cells to a more malignant phenotype. We identified a C/A single nucleotide polymorphism at −160 from the transcriptional start site of the E-cadherin gene promoter. Transient transfection experiments showed that the A allele of this polymorphism decreased the transcriptional efficiency by 68% compared with the C allele (P < 0.001). Electrophoretic mobility shift and footprinting assays revealed that the C allele had a stronger transcriptional factor binding strength than the A allele. These results indicate that the −160 C/A polymorphism has a direct effect on E-cadherin gene transcriptional regulation. This allelic variation may be a potential genetic marker that can help identify those individuals at higher risk for invasive/metastatic diseases.

Introduction

E-cadherin, one of the classic cadherins, plays a major role in the establishment and maintenance of intercellular adhesion, cell polarity, and tissue architecture (1). Epithelia are essential and abundant tissues in most eukaryotic organisms; >90% of the malignant human tumors are derived from epithelia. Abnormalities in the expression and cellular localization of E-cadherin are frequently associated with high tumor grade, infiltrative growth, and lymph node metastasis in a variety of human malignancies (2–6). Compelling experimental evidence indicates that E-cadherin serves as a potent invasion suppressor gene (7, 8). In addition, a tumor suppressor effect of E-cadherin has been suggested in human cancer (9, 10). Dysfunction of E-cadherin has also been associated with a number of nonmalignant diseases such as ulcerative and Crohn’s colitis, Langerhans’ cell histiocytosis, endometriosis, and autosomal dominant polycystic kidney disease (11–13).

Genetic factors contribute to virtually every human disease, conferring susceptibility or resistance, or influencing interaction with environmental factors (14). The most common type of human genetic variation is SNP, which occurs about once in every 1000 bases of the 3 billion bases in the human genome; some of those occurring in the promoter have been shown to produce profound effects on the transcription of its gene (15). We hypothesize that polymorphism in the promoter region of the E-cadherin gene is responsible for interindividual variation in the production of E-cadherin and in turn leads to individual susceptibility to invasive/metastatic carcinoma and other epithelial dysfunctions. We therefore screened the proximal promoter of the E-cadherin gene in search of common genetic variants with distinct effects on the transcriptional activity of the gene. Here we report a C/A SNP at position −160 relative to the transcription start site in the E-cadherin promoter that alters transcription factor binding and promoter strength.

Materials and Methods

DNA Analysis. For sequencing of the promoter region of the E-cadherin gene, a 454-bp fragment of the proximal promoter spanning from position −277 to +177 was amplified by PCR using primer E-cad S1 and E-cad S2 (Table 1). The PCR products were sequenced on an ABI sequencer with Dye Terminators (Applied Biosystems) using both upstream and downstream primers. For RFLP analysis, DNA was amplified using primer E-cad 5’ and E-cad 3’ (Table 1). PCR products were digested with either Hpyhl or AfIII. The digestion reactions were fractionated on a 4% agarose gel. The C allele created an Hpyhl site, and the A allele created an AfIII site.

Generation of Luciferase-Reporter Constructs and Transfection. DNA was amplified with primer E-cad Koml and E-cad BglIII, each of them had a Koml and BglII site introduced to the 5’ end, respectively. The C allele and the A allele were amplified from human PC3 and LNCaP prostate cancer cell line DNA samples, respectively. The PCR products were digested with Koml and BglII and then cloned into promoterless pGL3 Enhancer vector (Promega). The vector containing either C alleles or A alleles were designated as pGL3-C or pGL3-A, respectively. Plasmid DNA was obtained by transforming the vector into JM 109 cells and subsequent large-scale plasmid preparation using Qiafilter Plasmid Maxi kit (Qiagen). Reporter constructs were sequenced prior to use in reporter assays.

The human prostate cancer cell line DU145 was plated into a 24-well culture plate at a density of 5 × 104 cells/well and grown overnight to 60% confluence. In each experiment, three different luciferase reporter plasmids were transfected: (a) pGL3-C; (b) pGL3-A; and (c) pGL3-Control (Promega), which contains SV40 promoter and enhancer sequences. The DNA mixture for transfection was composed of test plasmid (0.75 μg) and pSV-β-galactosidase Control vector (Promega; 0.04 μg) that served as internal control to normalize activities of luciferase. Transfection was carried out using the calcium phosphate method. Luciferase activity was measured with a luminometer (Model TD-20/20; Promega), and the β-galactosidase activity was measured in a plate reader. To correct for transfection efficiency, light units from the luciferase assay were divided by the absorbance reading from the β-galactosidase assay. The corrected E-cadherin promoter-driven luciferase activity is expressed as a percentage of the pGL3-Control SV40 promoter-driven luciferase activity that served as the positive control in every transfection experiment. Luciferase activity was expressed as relative luciferase units. The promoterless pGL3-basic vector (Promega) lacking promoter and enhancer was used as a negative control in each of the transfection experiments. Statistics were performed using Student’s unpaired two-tailed t test.

EMSA. Complementary oligonucleotide pairs corresponding to human E-cadherin gene promoter sequence (from −175 to −147) were synthesized (UCSF Biomolecular Resource Center). Each of the oligonucleotide pairs was annealed and purified on a 6% polyacrylamide gel. The oligonucleotides were labeled with [γ-32P]ATP. A 1-μl (50,000 cpm) sample of 32P-labeled probe was incubated with 5 μg of HeLa nuclear extract (New England Biolabs) for 20 min at room temperature. Protein-DNA binding specificity was tested by
Allele-specific Binding of Nuclear Proteins to the −160 Polymorphic Site. To understand the mechanism by which the polymorphic alleles produced different promoter strengths, two synthetic double-stranded DNA probes, designated EC and EA (−175 to −147) were subjected to EMSA, EC corresponds the C allele, and EA corresponds the A allele. The sequences of the oligonucleotides used as probes and competitor in the EMSA analysis are shown in Table 1. Two DNA-protein complexes as indicated in Fig. 3 were observed. Complex I was evident in Lanes 2, 7, and 8 but was almost invisible in the corresponding binding reactions with probe EA. Complex II, the major DNA-protein binding complex, was evident in both binding reactions with probe EC and EA (Lanes 2, 7, 10, and 15) but was more abundant with probe EC (Fig. 3).

To verify the specificity of DNA-protein complexes, competition assays using specific and nonspecific oligonucleotides were performed (Fig. 3). The binding was inhibited competitively by adding specific competitive oligonucleotides (Lanes 3–6 and 11–14) but not by a nonspecific competitor (Lanes 7 and 15). When the oligonucleotide EC was used as a competitor with probe EA, it completely disrupted complexes I and II (Lane 16). However, when oligonucleotide EA was used as a competitor with probe EC, the disruption of DNA-protein complexes was not as effective as oligonucleotides EC with probe EA (Lane 16) or with probe EC (Fig. 3, Lanes 3–6).

To further define the binding site of the potential transcription factor suggested by EMSA analysis, the region surrounding the polymorphic site was also examined by DNase I footprinting analysis. A 284-bp DNA fragment containing the human E-cadherin gene promoter sequence between −234 and +48 was used as a template in DNase I protection assays with HeLa nuclear extract (Fig. 4). A footprint was clearly visible from −164 to −157 with the Ecad-C probe. The appearance of the protected regions on the DNA template was dependent on the concentration of nuclear protein in each DNase I digestion and became more visible in the presence of an increasing amount of proteins (Fig. 4, compare Lane 2 with Lane 5). The specific

A SNP IN E-CADHERIN PROMOTER

Table 1 Sequences of oligonucleotides synthesized for PCR, EMSA probe, and competitors

<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th>Human E-cadherin promoter</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR primers</td>
<td>−160 CCAGGCTAGGGGTCACCCGGGTCTCATGCGA−147</td>
</tr>
<tr>
<td>E-cad S1</td>
<td>−277 TTCTGATCCAGGTCTTTAGTGACG−254</td>
</tr>
<tr>
<td>E-cad S2</td>
<td>+177 GTCGCTGCACTGACGCAGACGACG−154</td>
</tr>
<tr>
<td>Ecad-KpnI</td>
<td>−365 ATCCGGAACCTTGGGTAAGAGGATGACG−345</td>
</tr>
<tr>
<td>Ecad-BglII</td>
<td>+4 ATGCACTTTGGAATGACGCAGACGACG−30</td>
</tr>
<tr>
<td>Ecad 5′</td>
<td>−277 TCCAGGCTTTAGTGACG−251</td>
</tr>
<tr>
<td>Ecad 3′</td>
<td>−246 GGCCACAGCCAATGACG−99</td>
</tr>
<tr>
<td>EMSA oligonucleotides</td>
<td></td>
</tr>
<tr>
<td>EC</td>
<td>5′ CCAGGCTAGGGGTCACCCGGGTCTCATGCGA3′</td>
</tr>
<tr>
<td>EA</td>
<td>5′ CCAGGCTAGGGGTCACCCGGGTCTCATGCGA3′</td>
</tr>
<tr>
<td>EN</td>
<td>5′ AGCTTCGAGGGGTCACCCGGGTCTCATGCGA3′</td>
</tr>
</tbody>
</table>

Results

Identification of a Polymorphism Site at −160 of E-cadherin Promoter. We screened a 454-bp region (between −277 and +177) of the E-cadherin proximal promoter and part of exon 1 by sequencing seven human prostate cell line DNA samples and found a C/A polymorphism site at position −160 relative to the transcriptional start site (dbSNP accession number, ss18684; Fig. 1). No additional polymorphisms or other genetic variations were detected.

Effects of the −160 Polymorphism on Promoter Activity. To examine the potential effects of the −160 C/A polymorphism on E-cadherin gene transcription, a 413-bp promoter of E-cadherin gene (−365 to +48) carrying either the C or A allele was inserted upstream of the luciferase gene in the pGL3 promoterless enhancer plasmid vector. The activity of E-cadherin C/A promoter-luciferase reporter gene constructs were assessed in transient transfection assays in DU145 human prostate cancer cells. Triplicate experiments were performed using DNA from different plasmid preparations. As shown in Fig. 2, significantly lower luciferase activities were observed for the pGL-A construct as compared with the pGL-C construct (a 68% decrease; P < 0.001).

![Fig. 1. E-cadherin −160 C/A polymorphism. A, sequencing chromogram. *, the polymorphic site. B, RFLP patterns: CA, individual heterozygous; AA, individual homozygous for A allele; CC, homozygous for C allele; u, uncut PCR product; a, PCR products cut with ApIII; h, PCR products cut with HpaII; M, 100-bp DNA marker.](https://cancerres.aacrjournals.org)
bases protected were identified by alignment with modified Gilbert-Maxim G+A sequencing reaction run in parallel and are indicated by the sequences shown at the side of the gel in Fig. 4. No protection of this region could be detected with probe Ecad-A (Fig. 4).

Discussion

In the present study, we have screened the proximal promoter region of the human E-cadherin gene for sequence variants and identified a common SNP at position −160 from the transcriptional start site. Previous studies have shown that a fragment spanning −399 to +31 relative to the transcription start site of the human E-cadherin gene possesses basal promoter activity (16, 17); thus, we focused our screening on this region. Several major cis-acting elements have been identified within a short section of the proximal promoter. Among these are the two E boxes, a CAAT box, and a SP1 binding site (17). Our results demonstrate that the polymorphism at position −160 has a significant effect on transcriptional activity in transient transfection studies. The molecular mechanism of this difference may well be explained as the difference in affinity of the DNA-binding protein(s) to the two allelic forms of the E-cadherin promoters. Our footprinting data clearly show that this region from −164 to −157 is protected by nuclear protein(s) from DNase I digestion. At least two proteins are involved in forming DNA-protein complexes, as evidenced by the EMSA assays in which two specific DNA-protein complexes were observed. We have searched the transcriptional factor database using TRANSFAC,4 and there are no known transcriptional factor binding sites that have homology to the sequence around the C/A polymorphism site of the human E-cadherin gene promoter. Most likely, this polymorphic site is a binding site for unknown transcription factors that are required for the E-cadherin promoter to function at an adequate level. The decreased transcriptional activity observed in DU145 cells transfected with plasmid carrying the A allele may be explained as the result of structure differences between the A and the C alleles, which hinders access of DNA by transcription factors. However, the change of a cytosine to an adenine in the DNA structure does not abandon the binding completely. A loose binding may still occur, as evidenced by the EMSA assays in which a probe containing the C allele produced more abundant DNA-protein complexes than a probe containing the A allele. In addition, when an oligonucleotide containing the C allele was used to compete with a probe containing the A allele, it totally disrupted probe binding with nuclear protein. However, when an oligonucleotide containing the A allele was used to compete with a probe containing the C allele, it was not as effective as an oligonucleotide containing the C allele in disrupting probe binding with nuclear proteins. It is clear that the sequence containing the C allele will have strong binding activities with transcriptional factors, leading to high transcriptional activities. This difference suggests that binding affinity of protein with DNA may be the basis for the observed differences in transcriptional activity of the two alleles.

In summary, the −160 C/A polymorphism, located within the regulatory region of E-cadherin promoter, influences E-cadherin transcription by altering transcription factor binding. This SNP could have significant effects on the susceptibility or vulnerability to develop carcinoma and subsequently invasiveness and metastasis of carcinoma. This hypothesis is currently being tested in common human carcinomas.

Fig. 2. Transient transfection assay to measure promoter activity of C and A alleles in DU145 cells. A, schematic of the human E-cadherin promoter. B, the human E-cadherin gene promoter corresponding to positions −365 to +48 relative to the transcription initiation site (+1) was cloned from PC3 and LNCaP prostate cancer cell lines upstream of the luciferase reporter gene in plasmid pGL3 enhancer in the 5’ to 3’ orientation. Each allele luciferase reporter gene construct was transiently transfected into DU145 prostate cancer cells. Data were normalized to β-galactosidase activity and are expressed as a percentage of the corrected luciferase activity of pGL3-control (means of three independent experiments; bars, SE).

Fig. 3. EMSA with HeLa cell extracts and oligonucleotide probes containing C and A alleles of human E-cadherin gene promoter. Each binding reaction contained 5 μg of nuclear protein and labeled EC (Lanes 2–8) or EA (Lanes 10–16) oligonucleotide probe. Excess unlabeled EC or EA oligonucleotides (10-, 20-, and 50-fold) were included in the binding reactions as competitor in Lanes 3–6 and Lanes 11–14. In addition, 50-fold excess of unlabeled EA and EC oligonucleotides were used to compete with probes EC (Lane 8) and EA (Lane 16). Fifty-fold excess of nonspecific competitor EN was used in Lane 7 and Lane 15. Arrowheads, specific retarded complexes, I and II. The free probe is shown at the bottom of each lane.

4 Internet address: http://transfac.gbf.de/TRANSFAC/programs.html.

875
Fig. 4. The effects of the C/A polymorphism on footprints in DNase I footprinting assays. The polymorphic site from the C allele (positions –164 to –157) is protected from digestion by DNase I in the presence of HeLa nuclear extracts. The position of the protected regions is indicated with a box, and the actual bases are shown at the size of the gel.

Lane 1, Ecad-C template digested in the absence of nuclear proteins; Lanes 2–5, Ecad-C template digested in the presence of 15, 20, and 25 μg of nuclear protein, respectively; Lane 6, Ecad-A template digested in the absence of HeLa nuclear extracts. The position of the protected regions is indicated with a box, and the actual bases are shown at the size of the gel.

Lane 6, Ecad-A template digested in the absence of nuclear proteins; Lanes 7–10, Ecad-A template digested in the presence of 15, 20, and 25 μg of nuclear protein, respectively. Lane M, G+A ladder.

References
A Single Nucleotide Polymorphism in the *E-cadherin* Gene Promoter Alters Transcriptional Activities

Long-Cheng Li, Richard M. Chui, Masahiro Sasaki, et al.

*Cancer Res* 2000;60:873-876.

Updated version  Access the most recent version of this article at:  
http://cancerres.aacrjournals.org/content/60/4/873

Cited articles  This article cites 16 articles, 7 of which you can access for free at:  
http://cancerres.aacrjournals.org/content/60/4/873.full.html#ref-list-1

Citing articles  This article has been cited by 21 HighWire-hosted articles. Access the articles at:  
/content/60/4/873.full.html#related-urls

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.