Definition of a Functional Single Nucleotide Polymorphism in the Cell Migration Inhibitory Gene MIIP That Affects the Risk of Breast Cancer

Fangfang Song1,2, Ping Ji3, Hong Zheng1,2, Fengju Song1,2, Yingmei Wang3,4, Xishan Hao1,2, Qingyi Wei5, Wei Zhang4, and Kexin Chen1,2

Abstract

The migration and invasion inhibitory protein MIIP is an inhibitor of cancer cell migration and invasion that inhibits breast tumorigenesis. In this case-control study, we evaluated the MIIP single nucleotide polymorphism (SNP) rs2295283 (codon 167, A>G, K>E) from 1,524 breast cancer patients and 1,592 age-matched controls for its association with breast cancer risk. SNP analysis included a validation set of 736 cases and 760 controls. Colony formation and cell migration assays were then conducted to functionally interrogate the genotype difference. When compared with the AA genotype, the combined AG + GG genotypes (167E) were associated with a significantly lower risk of breast cancer. In the test set, the protective effects of the AG + GG genotypes were more evident among participants with a family history of cancer. Further case series analysis revealed that the GG genotype was associated with reduced breast cancer susceptibility in cases of tumor size >2 cm and late clinical stage (II + III + IV). Colony formation assays showed that MIIP 167E (the G variant) was a more potent inhibitor of colony formation but not cell migration. These results suggest MIIP K167E as a functional genetic marker of breast cancer development and prognosis. Cancer Res; 70(3); 1024–32. ©2010 AACR.

Introduction

Most cancers are believed to result from gene-environment interactions. The ubiquitous industrial and biological pollutants seen in recent decades and the high energy intake and sedentary lifestyles of the modern society have put increasing pressure on the genetic makeup of cells, heightening the need to understand how genetic variants and mutations affect cancer development and prognosis.

Breast cancer is the most common cancer among women worldwide, and its incidence is increasing (1). Most recognized genetic mutations for breast cancer are found in the tumor suppressor genes (TSG) BRCA1 and BRCA2, which are responsible for familial breast cancer (2). Germ-line mutations of the p53 gene are also responsible for breast cancers associated with Li-Fraumeni syndrome (3). However, familial forms account for less than 10% of breast cancers in the general population. Most sporadic breast cancers, in contrast, are believed to be affected by polymorphisms, especially single-nucleotide polymorphisms (SNP), in concert with environmental exposures (4). Such SNPs have been reported for metabolism-related genes, oncogenes, and TSGs such as the p53 gene (5, 6).

The migration and invasion inhibitory protein (MIIP; also known as Ilp45) gene is a recently characterized putative TSG in glioma (7). MIIP is located on chromosome 1p36.22, which is commonly deleted in a wide spectrum of cancers; however, somatic mutations have not been found in this gene (7). The MIIP protein (~45 kDa) is coded by 10 exons of MIIP, and a number of SNPs are scattered throughout the gene, but the functional significance of these SNPs is not known.

The MIIP protein was first found to bind to a product of an oncogene, insulin-like growth factor–binding protein-2 (IGFBP-2), and to block the migration- and invasion-enhancing functions of IGFBP-2 (8). IGFBP-2, a recently recognized oncogene product that functions through both IGF-dependent and IGF-independent pathways (9, 10), is commonly upregulated in advanced stages of cancers, including breast cancers (11–15). Specifically in breast cancer, IGFBP-2 expression was found to increase in a stepwise manner from hyperplasia to atypical hyperplasia, to breast carcinoma in situ, and to invasive breast cancer (15), and from T1N0 to T1N1 breast carcinomas (12). One laboratory study showed that IGFBP-2, independent of the IGF receptor, induced expression of multiple genes that are functionally important.
for cell proliferation, adhesion, and migration in the Hs578T breast cancer cell line (10). Furthermore, an in vitro model showed the involvement of IGFBP-2 in antiestrogen resistance (16), and a xenograft model recently showed that IGFBP-2 inhibition led to decreased tumorigenesis, suggesting that IGFBP-2 may be a therapeutic target for breast cancer (17). Therefore, in its role as an endogenous IGFBP-2 inhibitor, MIIP is likely a key modifier gene for the oncogenic effect of IGFBP-2.

In this study, we used a population-based study design coupled with functional validation methods to investigate the role of MIIP genetic variants in breast cancer development and prognosis. We had first uncovered a nonsynonymous SNP of MIIP that was associated with various degrees of risk for breast cancer development and various prognoses. Because the two SNP isoforms encode two different amino acids, K or E, at codon 167, we then performed an in vitro tumorigenesis assay with these isoforms and found that codon 167E was a more potent colony formation inhibitor, which is consistent with the protective effect of the 167E variant genotypes (AG + GG or GG) in cancer risk and prognosis. Thus, the study herein provides new evidence that MIIP is a critical regulatory gene in the development of breast cancer.

Materials and Methods

Characteristics of participants and clinical protocol. All participants enrolled in this study were of Chinese Han ethnicity. The Ethics Committee of Tianjin Medical University Cancer Hospital approved the study protocol, and we obtained written informed consent from all patients and controls to participate in this study. At first, we consecutively recruited 1,524 patients with newly diagnosed and histologically confirmed breast cancer from the Breast Cancer Research Center in Tianjin Medical University Cancer Hospital between January 1, 2007 and February 1, 2008. During this same period, we collected blood samples from 1,592 healthy female volunteers (controls) from the nearby community, who were genetically unrelated to the patients and were frequency matched to the patients by age (±5 y). To validate the findings from the test set, an additional cohort of cases (736) and controls (760) recruited between June 2008 and January 2009 from the same hospital and community was used as a validation set. All the study participants met the following inclusion criteria: ages 18 to 65 y and no previous diagnosis of cancer. The response rates of the eligible patients and controls that we approached for recruitment were approximately 95% and 90%, respectively.

All participants completed a structured questionnaire about their socio-demographic characteristics, family history of cancer, reproductive factors, and lifestyle. From the patients, we also collected clinical information on tumor features and disease severity, including morphology, tumor size, lymph node metastasis, organ metastasis, tumor stage, and status of estrogen receptor and progesterone receptor. Each participant donated 20 mL of blood, which was collected into heparinized tubes and used for DNA extraction and the genotyping assays.

SNP selection. From the resequencing data of 45 Chinese Han individuals in the International HapMap Project SNP database, we selected three tagging SNPs (rs11588712, rs2295283, and rs2295289) with an $r^2$ threshold of 0.80, minor allele frequencies (MAF) of >0.05, and putative functional potentials (nonsynonymous SNPs). From the National Center for Biotechnology Information (NCBI) dbSNP database, we identified another three SNPs (rs11553925, rs35317667, and rs34874602) from individuals with mixed ethnic backgrounds. These SNPs, unavailable in the HapMap data, had a MAF of ≥0.05 across the whole genomic region with putative functional potentials.

SNP genotyping. Genomic DNA, extracted from whole blood by using the QIAGEN DNA Blood Mini Kit according to the manufacturer’s instructions, was stored at −20°C until analysis. Genotyping was done by using the MGB TaqMan probe assay (Applied Biosystems, Inc.). The concordance rate for genotypes was 100% in 10% of samples with duplicates. For SNP inapplicable for the TaqMan assay (specific and perfect probes and primers could not be designed), such as SNP rs11588712 (C>T) in our study, RFLP-PCR was used to identify the genotypes. Briefly, the PCR primers for SNP rs11588712 (5′-GACCTCATCGACCAGAGA-3′ and 5′-TGGCCTCAGCTTTATGTGTA-3′) produced a 550-bp DNA fragment, which was digested with XhoI (New England BioLabs, Inc.) overnight at 37°C. The digested product was separated on a 2.0% NuSieve 3:1 agarose gel (FMC BioProducts) stained with ethidium bromide, and photographed with Polaroid film. Theoretically, allele C lacks the XhoI restriction site and therefore produces a single 550-bp band, whereas allele T produces two bands of 274 and 276 bp, and the CT heterozygote produces three bands of 274, 276, and 530 bp.

In our present study, only the CC and CT genotypes were detected (Supplementary Fig. S1A), consistent with those reported in the HapMap database. Also, the assumed genotypes that we obtained were individually validated by direct sequencing of the PCR products (Supplementary Fig. S1B). More than 10% of the samples were randomly selected for replication, and results were 100% concordant.

MIIP mutagenesis and colony formation assay. The wild-type MIIP was cloned into pcDNA3.1, as described previously (8). To generate the construct with the MIIP SNP at the codon 167 site, one pair of primers was designed for point mutagenesis: primer 1, GGAGTCTGCAGTTCCTGA-GAGGAGCTGGCGCTC; primer 2, GAGGCGC-CAGCTCCTCTCAGGAAGTCGACAGTCC. PCR mutagenesis was done according to the protocol from Stratagene. The resulting construct with MIIP 167 SNP (pcDNA-MIIP 167E) was confirmed by sequencing.

For the colony formation assay, MCF7 cells were transfected with pcDNA3.1, pcDNA-MIIP, or pcDNA-MIIP 167E by Lipofectamine 2000, spread onto six-well plates in triplicate;
and selected with 200 ng/mL hygromycin B for 2 wk. The colonies were fixed with 4% formaldehyde, stained with 0.05% crystal violet, and counted as previously described (18). The MCF7 cell line was obtained from the American Type Culture Collection and had not passed in our laboratory for more than 10 times.

**Cell migration assay.** Cell migration assays were done by using transwell polycarbonate chambers (Becton Dickinson). Briefly, MCF7 cells were cotransfected with pCDNA-green fluorescent protein (GFP) and pCDNA-MIIP or pCDNA-MIIP 167E in a 1:3 ratio for 24 h, and then $2 \times 10^4$ transfected cells in serum-free DMEM were seeded in the upper well to allow the cells to migrate to the lower well with DMEM plus 20% fetal bovine serum. The chamber was incubated for 24 h at 37°C, and the migrated cells in the lower well were trypsinized and washed with 1× PBS. The GFP-positive migrated cells were counted by FACScan (Becton Dickinson). Assays were done in triplicate and repeated twice.

**Statistical analysis.** Differences in the distribution of demographic variables and known risk factors between breast cancer patients and controls were tested by the Student t test for continuous variables or by the $\chi^2$ test for categorical variables, as appropriate.

We also used the $\chi^2$ test to evaluate differences in allelic and genotypic frequencies of each SNP between patients and controls, as well as the Hardy-Weinberg equilibrium of each individual locus in controls. The multivariate logistic regression method was used to assess the association between breast cancer risk and polymorphisms in the MIIP gene. Odds ratios (OR) and 95% confidence intervals (95% CI) were calculated with adjustment for known risk factors for breast cancer, such as age, age at menarche, duration of breast-feeding, menopause status, oral contraception, family history of cancer, and physical activity. Also, potential gene-environment interaction at a multiplicative scale was assessed in the logistic regression analysis by comparing the changes in deviance between the models for main effects with or without the interaction terms. For breast cancer patients only, we also performed stratified case series analysis of the genotype data by clinical phenotypes.

Statistical analyses were done with the use of SAS software (version 9.0, SAS Institute) and a two-sided test, with $P < 0.05$ considered statistically significant.

**Results**

**Baseline characteristics of the study population.** The detailed baseline characteristics for the patients and controls in both the test set and the validation set are presented in Table 1. As expected, patients in the test set reported a greater number of known risk factors for breast cancer than did the cancer-free controls. For example, significantly larger proportions of patients than the controls had fewer birth numbers, less breast-feeding time and physical activity, were nulliparous and younger at menarche, and had history of oral contraceptive usage, benign breast disease, and family history of cancer. However, the two groups of the test set did not differ in the distribution based on age or menopausal status. Similar trends in the distributions of the aforementioned baseline variables were observed between cases and controls in the validation set, except for higher proportions of menopause (75% versus 55%), exercise (28% versus 10%), benign breast disease (16% versus 7%), and family history (25% versus 12%) in controls of the validation set than in those of the test set (Table 1), indicating some selection bias in controls of the validation set.

**Information on selected MIIP SNPs.** Supplementary Table S1 lists the six selected MIIP SNPs and genotyping data obtained from our pilot study of 188 control individuals. The observed MAFs of all three tagging SNPs (rs11588712, rs2295283, and rs2295289) from these controls were close to those reported in the HapMap database. However, the observed MAFs of the other three SNPs (rs11539325, rs35317667, and rs34874602), which are not available in the HapMap database, were markedly different from those in the NCBI dbSNP database, indicating that these SNPs may have some ethnic differences in their MAFs.

Because our initial genotyping efforts found that the minor variants of five of the six SNPs examined were either extremely rare or completely absent (observed only a few times from the heterozygotes for rs11553925, rs11588712, and rs2295289, and none for rs35317667 and rs34874602), these five SNPs were not included for further genotyping analysis due to statistical considerations. The remaining SNP, rs2295283, was subsequently genotyped and analyzed for its association with breast cancer risk in the test set. The two variants (A and G) of SNP rs2295283 result in two different amino acids (Lys versus Glu) at codon 167 (K and E) in the NH2-terminal region of the MIIP protein.

**Association of the selected MIIP SNP with breast cancer.** Genotype frequencies of the selected MIIP SNP rs2295283 among patients and controls are shown in Table 2. There was no significant difference in genotype distributions between the controls and those expected from the Hardy-Weinberg equilibrium ($P = 0.968$). Although the G-allele frequency was lower in patients (47.93%) than in controls (49.21%), the difference was not statistically significant ($P = 0.312$). With regard to genotypic distribution, the percentage of patients (71.98%) who carried the G allele (AG + GG genotype) was less than that of controls (74.18%), but again the difference did not reach statistical significance ($P = 0.166$).

The associations between MIIP SNP rs2295283 genotypes and breast cancer risk are presented in Table 2; all ORs and 95% CIs were calculated by using the common homozygous AA genotype as the reference group. The AG genotype was associated with a decreased risk of breast cancer compared with the AA genotype (OR, 0.75; 95% CI, 0.59–0.93; $P = 0.026$). With regard to genotypic distribution, the percentage of patients (71.98%) who carried the G allele (AG + GG genotype) was less than that of controls (74.18%), but again the difference did not reach statistical significance ($P = 0.166$).
genotype was not stronger than that of the AG genotype, possibly due to some unknown selection bias.

When the results were further stratified by family history of cancer (Table 2), we found that the protective effect of the G allele (the combined AG + GG genotype) was more pronounced in subjects who have a family history of cancer, contributing to a >50% risk reduction of breast cancer (OR, 0.42; 95% CI, 0.24–0.71; \( P = 0.001 \)). In those with a family history of cancer, the protective effect of the G allele was more marked because both the AG and GG genotypes provided a significant protection (Table 2).

The association of breast cancer risk with MIIP SNP rs2295283 was replicated in another independent case-control study (736 cases versus 760 controls). Genotyping results showed that this SNP was statistically associated with breast cancer risk in both the validation set and the final combined case-control set (Table 3). The AG + GG genotypes were associated with a decreased breast cancer risk (OR, 0.77; 95% CI, 0.60–0.98; \( P = 0.036 \)) compared with the AA genotype in the validation set. When we combined these two study sets, the presence of G allele was an independent protective factor for breast cancer, with adjusted ORs of 0.84 (95% CI, 0.72–1.00; \( P = 0.050 \)) and 0.81 (95% CI, 0.68–0.96; \( P = 0.016 \)) for genotypes AG and GG, respectively. Similar results were obtained with the combination of these two genotypes (AG + GG; Table 3).

### Table 1. Baseline characteristics of breast cancer patients and cancer-free controls from the test and the validation sets

<table>
<thead>
<tr>
<th>Variables</th>
<th>Test set, ( n (%) )</th>
<th>( P^* )</th>
<th>Validation set, ( n (%) )</th>
<th>( P^* )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients (( n = 1,524 ))</td>
<td>Controls (( n = 1,592 ))</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \leq 50 )</td>
<td>731 (47.97)</td>
<td>731 (45.92)</td>
<td>0.252</td>
<td>337 (45.79)</td>
</tr>
<tr>
<td>( &gt;50 )</td>
<td>793 (52.03)</td>
<td>861 (54.08)</td>
<td></td>
<td>399 (54.21)</td>
</tr>
<tr>
<td>Age at menarche (y)†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \leq 12 )</td>
<td>150 (9.85)</td>
<td>79 (4.97)</td>
<td>&lt;0.0001</td>
<td>75 (10.26)</td>
</tr>
<tr>
<td>( &gt;12 )</td>
<td>1,373 (90.15)</td>
<td>1,509 (95.03)</td>
<td></td>
<td>656 (89.74)</td>
</tr>
<tr>
<td>Pregnancy†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>37 (2.45)</td>
<td>22 (1.38)</td>
<td>0.030</td>
<td>22 (3.02)</td>
</tr>
<tr>
<td>Ever</td>
<td>1,474 (97.55)</td>
<td>1,567 (98.62)</td>
<td></td>
<td>706 (96.98)</td>
</tr>
<tr>
<td>No. of births‡</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \leq 2 )</td>
<td>1,287 (87.31)</td>
<td>1,166 (74.46)</td>
<td>&lt;0.0001</td>
<td>621 (87.96)</td>
</tr>
<tr>
<td>( &gt;2 )</td>
<td>187 (12.69)</td>
<td>400 (25.54)</td>
<td></td>
<td>85 (12.04)</td>
</tr>
<tr>
<td>Duration of breast-feeding (mo)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \leq 12 )</td>
<td>724 (47.51)</td>
<td>601 (37.75)</td>
<td>&lt;0.0001</td>
<td>329 (44.70)</td>
</tr>
<tr>
<td>( &gt;12 )</td>
<td>800 (52.49)</td>
<td>991 (62.25)</td>
<td></td>
<td>407 (55.30)</td>
</tr>
<tr>
<td>Menopause†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>723 (47.83)</td>
<td>706 (44.71)</td>
<td>0.103</td>
<td>345 (47.46)</td>
</tr>
<tr>
<td>Yes</td>
<td>795 (52.17)</td>
<td>873 (55.29)</td>
<td></td>
<td>382 (52.54)</td>
</tr>
<tr>
<td>Oral contraception use†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1,181 (81.67)</td>
<td>1,317 (84.48)</td>
<td>0.040</td>
<td>585 (84.29)</td>
</tr>
<tr>
<td>Ever</td>
<td>265 (18.33)</td>
<td>242 (15.52)</td>
<td></td>
<td>109 (15.71)</td>
</tr>
<tr>
<td>Exercise (times/wk)†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \leq 1 )</td>
<td>1,382 (96.37)</td>
<td>1,411 (89.87)</td>
<td>&lt;0.0001</td>
<td>690 (94.65)</td>
</tr>
<tr>
<td>( &gt;1 )</td>
<td>52 (3.63)</td>
<td>159 (10.13)</td>
<td></td>
<td>39 (5.35)</td>
</tr>
<tr>
<td>Benign breast disease†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1,113 (73.56)</td>
<td>1,460 (92.76)</td>
<td>&lt;0.0001</td>
<td>562 (76.88)</td>
</tr>
<tr>
<td>Ever</td>
<td>400 (26.44)</td>
<td>114 (7.24)</td>
<td></td>
<td>169 (23.12)</td>
</tr>
<tr>
<td>Family history of cancer†</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1,053 (69.14)</td>
<td>1,403 (88.41)</td>
<td>&lt;0.0001</td>
<td>517 (70.53)</td>
</tr>
<tr>
<td>Yes</td>
<td>470 (30.86)</td>
<td>184 (11.59)</td>
<td></td>
<td>216 (29.47)</td>
</tr>
</tbody>
</table>

*Two-sided \( \chi^2 \) test. \( P < 0.05 \) was considered statistically significant.
†Due to missing values, the number of patients was <1,524 and that of controls was <1,592 in the test set, and in the validation set, the number of patients was <736 and that of controls was <760.
‡First- and second-degree relatives.
Having shown that the MIIP codon 167 SNP was associated with breast cancer risk, we next sought to determine whether this SNP also affected breast cancer prognosis in the patient-only analysis. As shown in Table 4, we compared differences in genotypic distribution of MIIP SNPs stratified by each individual clinical phenotype in the test set with a larger sample size and calculated ORs by modeling the probability of the worse prognostic phenotype for breast cancer. Specifically, we

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Patients (n)</th>
<th>Controls (n)</th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>All participants</td>
<td>427 411</td>
<td>1.00</td>
<td>0.75 (0.59-0.93)</td>
<td>0.010</td>
</tr>
<tr>
<td>AG</td>
<td>795 733 364</td>
<td>0.86 (0.68-1.13)</td>
<td>0.322</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>1,181 1,097</td>
<td>0.78 (0.64-0.97)</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>AG + GG</td>
<td>1,097 1,181</td>
<td>0.78 (0.64-0.97)</td>
<td>0.026</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.** Frequency distribution of MIIP rs2295283 genotypes in the test set and their associations with risk of breast cancer represented by forest plot, adjusted by age, age at menarche, number of births, duration of breast-feeding, menopause, oral contraception, exercise, benign breast disease, and family history of cancer

**Distribution of MIIP genotypes by clinical phenotypes in the patient-only analysis.** Having shown that the MIIP codon 167 SNP was associated with breast cancer risk, we next sought to determine whether this SNP also affected breast cancer prognosis in the patient-only analysis. As shown in Table 4, we compared differences in genotypic distribution of MIIP SNPs stratified by each individual clinical phenotype in the test set with a larger sample size and calculated ORs by modeling the probability of the worse prognostic phenotype for breast cancer. Specifically, we

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Patients (n)</th>
<th>Controls (n)</th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test set</td>
<td>427 411</td>
<td>1.00</td>
<td>0.75 (0.59-0.93)</td>
<td>0.010</td>
</tr>
<tr>
<td>AG</td>
<td>733 795 364</td>
<td>0.86 (0.68-1.13)</td>
<td>0.322</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>1,181 1,097</td>
<td>0.78 (0.64-0.97)</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>AG + GG</td>
<td>1,097 1,181</td>
<td>0.78 (0.64-0.97)</td>
<td>0.026</td>
<td></td>
</tr>
<tr>
<td>Validation set</td>
<td>180 172</td>
<td>1.00</td>
<td>0.77 (0.59-1.00)</td>
<td>0.052</td>
</tr>
<tr>
<td>AG</td>
<td>372 374 182</td>
<td>0.76 (0.56-1.04)</td>
<td>0.084</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>572 332 182</td>
<td>0.77 (0.60-0.98)</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>AG + GG</td>
<td>1,653 1,769</td>
<td>0.83 (0.73-0.97)</td>
<td>0.018</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3.** Frequency distribution of MIIP rs2295283 genotypes in the validation and pooled sets and their associations with risk of breast cancer represented by forest plot, adjusted by age, age at menarche, number of births, duration of breast-feeding, menopause, oral contraception, exercise, benign breast disease, and family history of cancer
divided each phenotype into two groups, one with good prognosis and the other with poor prognosis. We observed that, when compared with AA genotype carriers, individuals carrying the GG genotype were less susceptible to developing larger tumors (>2 cm; OR, 0.67; 95% CI, 0.48–0.95; P = 0.023); these patients were also less likely to have late-stage disease (tumor-node-metastasis stage II + III + IV; OR, 0.63; 95% CI, 0.44–0.91; P = 0.012).

In all cases, there was no significant difference between the distributions of codon 167 SNP genotypes by mean age at diagnosis, tumor histologic types, occurrence of lymph node metastasis, or expression status of estrogen receptor or progesterone receptor, although a borderline significantly negative correlation existed between the GG genotype and invasive breast cancer (genotype GG versus genotype AA: OR, 0.46; 95% CI, 0.20–1.06; P = 0.068; Table 4). However, these results were not replicated in the validation set (data not shown), possibly due to the relatively smaller number of validation set, particularly for the stratification analysis, and relatively shorter follow-up of the patients in this replicative cohort.

**Functional evaluation of the two protein isoforms of MIIP codon 167 SNP.** The above studies showed that the G form or Glu (E) amino acid at codon 167 of the MIIP protein was associated with decreased risk of breast cancer, and the GG carriers tended to have smaller or lower-grade tumors. To determine whether the MIIP codon 167E is more potent in the inhibitory function than MIIP codon 167K,MIIP SNP and Risk of Breast Cancer

<table>
<thead>
<tr>
<th>Genotypes</th>
<th>Tumor features and disease severity</th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Morphology (n)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Noninvasive</td>
<td>Invasive</td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>14</td>
<td>413</td>
<td>1.00</td>
</tr>
<tr>
<td>AG</td>
<td>24</td>
<td>709</td>
<td>0.76 (0.34–1.69)</td>
</tr>
<tr>
<td>GG</td>
<td>18</td>
<td>346</td>
<td>0.46 (0.20–1.06)</td>
</tr>
<tr>
<td>Tumor size (n)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤2 cm</td>
<td>140</td>
<td>250</td>
<td>1.00</td>
</tr>
<tr>
<td>&gt;2 cm</td>
<td>221</td>
<td>434</td>
<td>1.02 (0.76–1.38)</td>
</tr>
<tr>
<td>Lymph node metastasis (n)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>255</td>
<td>170</td>
<td>1.00</td>
</tr>
<tr>
<td>Yes</td>
<td>442</td>
<td>287</td>
<td>1.04 (0.79–1.38)</td>
</tr>
<tr>
<td>Clinical stage (n)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 + I</td>
<td>102</td>
<td>288</td>
<td>1.00</td>
</tr>
<tr>
<td>II + III + IV</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>166</td>
<td>489</td>
<td>0.97 (0.69–1.35)</td>
</tr>
<tr>
<td>AG</td>
<td>108</td>
<td>227</td>
<td>0.63 (0.44–0.91)</td>
</tr>
<tr>
<td>GG</td>
<td>323</td>
<td>399</td>
<td>1.08 (0.82–1.43)</td>
</tr>
<tr>
<td>PR (n)</td>
<td>152</td>
<td>207</td>
<td>0.90 (0.67–1.28)</td>
</tr>
<tr>
<td>ER (n)</td>
<td>180</td>
<td>237</td>
<td>1.00</td>
</tr>
<tr>
<td>−</td>
<td>318</td>
<td>403</td>
<td>0.99 (0.75–1.30)</td>
</tr>
<tr>
<td>+</td>
<td>152</td>
<td>207</td>
<td>0.86 (0.62–1.18)</td>
</tr>
</tbody>
</table>

Abbreviations: ER, estrogen receptor; PR, progesterone receptor.
which would provide a functional explanation of our population-based study results, we cloned the two forms separately into an expression vector (Fig. 1A). We then transfected the two isoforms into the MCF7 breast cancer cell line, and both MIIP isoforms were expressed at equal levels (Fig. 1B). We then compared the effects of the two isoforms on colony formation in the presence of selection agent after 2 weeks and determined that the MIIP 167E inhibitory effect was higher and that this difference was statistically significant (Fig. 1C). Thus, the codon 167 SNP encodes two MIIP variants that display different tumor-suppressive activities.

Because MIIP was previously shown to inhibit cell migration (8), we also determined whether the two isoforms had different abilities in inhibiting breast cancer cell migration. Our results showed that the two isoforms had similar activity in inhibiting MCF7 cell migration (Fig. 1D).

Discussion

In this case-control study, we showed that the nonsynonymous codon 167 SNP of a recently identified TSG, MIIP, was not only associated with differential risk of breast cancer, particularly in individuals with a family history of cancer, but also associated with various tumor phenotypes in cancer patients. Specifically, codon 167E (G allele) was associated with a decreased risk of breast cancer in both the test set and the validation set. Consistently, patients carrying the codon 167E (G allele) of MIIP had tumors ≤2 cm and with a less advanced clinical stage. Thus, this study has uncovered the effects of a novel genetic variant that is important for breast cancer initiation and progression and has provided yet another piece of evidence that genetic makeup plays an important role in breast cancer development.

The MIIP gene was first characterized as an inhibitor of cancer cell migration and invasion (8). Our recent study showed that MIIP is a mitotic checkpoint protein and inhibits colony formation in vitro and tumor development and progression using a glioma mouse model,8 supporting that MIIP is a TSG in glioma. We next evaluated the role of MIIP as a suppressor in breast cancer by using an in vitro colony formation assay and found that MIIP potently suppressed colony formation in the MCF7 breast cancer cell line. The in vitro assays done for the current study allowed us to functionally compare the two isoforms of the MIIP codon 167 SNP. We hypothesized that MIIP codon 167E (G allele) had a more protective effect because it had a stronger inhibitory effect on tumorigenesis than did MIIP codon 167K, and the results from our in vitro colony formation assay supported this hypothesis. The differences found in tumor suppression activity between the two isoforms were also consistent with the association of codon 167E with smaller breast cancer tumors.

8 P. Ji, S.M. Smith, Y. Wang, et al. Inhibition of gliomagenesis and attenuation of mitotic transition by MIIP. Submitted for publication.
Thus, we have identified a mechanism through this functional study that explains the findings of our population-based studies. Although many nonsynonymous SNPs have been identified in various genes that have been associated with cancer risk, this type of functional study is particularly informative. Another example of such a successful effort was the functional study of the nonsynonymous codon 72 SNP of p53 TSG that changes the amino acid at codon 72, which has been shown to influence cancer risk, likely because of the functional differences between the two protein variants in DNA binding and transcriptional activation (19).

An interesting characteristic of the MIIP gene is that it reportedly contains eight nonsynonymous SNPs. It is intriguing that three of the common six SNPs we selected are not reported in the HapMap database, and our pilot study showed that the MAFs in Chinese subjects were markedly different from those obtained from other ethnic groups presented in the NCBI dbSNP database. This suggests that some of the nonsynonymous MIIP SNPs may have ethnic differences in their MAFs. An intriguing question is whether these SNPs may help explain the different incidences of breast cancer among the Chinese, white, and African American populations. Future case-control studies involving different ethnic groups should help answer these questions. Functional studies such as ours can provide mechanistic interpretation if population-based studies have revealed significant findings.

Although the present study was a considerably large study, it was a hospital-based study that has some inherent weaknesses. First, our study excluded nonsynonymous SNPs with low frequencies because they may also be involved in the etiology of breast cancer. Second, some of our patients were newly diagnosed between 2007 and 2008, and thus we did not have enough follow-up time for the survival analysis that otherwise would provide stronger evidence for the role of the MIIP SNP in breast cancer prognosis. Third, although our validation set consistently confirmed the overall association between the MIIP codon 167 SNP and breast cancer risk, our replication set was not large enough to replicate the results of the patient-only analysis. In addition, because of the known functional relationship between MIIP with IGFBP-2, in the future, it will be interesting to investigate the possible interactions between the MIIP SNP tested in the study and a recently reported SNP of IGFBP-2 that was shown to confer breast cancer risk (20).

In conclusion, we have identified the biological significance of a genetic variant in the coding region of the MIIP TSG in Chinese women that confers differential risk of developing breast cancer. This SNP is also associated with breast cancer prognosis. Our functional studies showed that the codon 167 SNP produces two MIIP isoforms that exert different inhibitory effects on tumorigenesis. Once this genetic variance is validated by larger studies with different ethnicities, it may be valuable for genetic counseling, especially for women with a family history of cancer, and may help oncologists in designing therapies with various degrees of aggressiveness.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

We thank Da Yang, Hongwei Han, and Lei Lei for their technical assistance, and Tamara Locke of the Department of Scientific Publications at M.D. Anderson Cancer Center for the editing of the manuscript.

Grant Support

National Natural Science Foundation of China (no. 30872172) and the Institutional Incentive Fund from The University of Texas M.D. Anderson Cancer Center. The sequencing core at the M.D. Anderson Cancer Center was supported by Cancer Center Core Grant from the National Cancer Institute (grant CA-16672). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received 10/12/09; revised 11/13/09; accepted 12/1/09; published OnlineFirst 1/26/10.

References

5. Freisinger F, Domchek SM. Clinical implications of low-penetrance TSG that changes the amino acid at codon 72, which has been shown to influence cancer risk, likely because of the functional differences between the two protein variants in DNA binding and transcriptional activation (19).


Definition of a Functional Single Nucleotide Polymorphism in the Cell Migration Inhibitory Gene MIIP That Affects the Risk of Breast Cancer

Fangfang Song, Ping Ji, Hong Zheng, et al.

Cancer Res  Published OnlineFirst January 26, 2010.