

Association of *Matrix Metalloproteinase-8* Gene Variation with Breast Cancer Prognosis

Julie Decock,¹ Ji-Rong Long,⁵ Ross C. Laxton,^{2,3} Xiao-Ou Shu,⁵ Conrad Hodgkinson,² Wouter Hendrickx,¹ Eve G. Pearce,³ Yu-Tang Gao,⁶ Andresa C. Pereira,⁴ Robert Paridaens,¹ Wei Zheng,⁵ and Shu Ye^{2,3}

¹Laboratory for Experimental Oncology, Department of General Medical Oncology, University Hospital Gasthuisberg, Leuven, Belgium;

²Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine, University of London, London, United Kingdom; ³Human Genetics Division, University of Southampton School of Medicine, Southampton, United Kingdom;

⁴Faculdade de Odontologia de São José dos Campos—Universidade Estadual Paulista, Departamento de Biociências e Diagnóstico

Bucal, Brazil; ⁵Department of Medicine, Vanderbilt Epidemiology Centre, Vanderbilt–Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee; and ⁶Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China

Abstract

Animal and cell studies indicate an inhibitory effect of matrix metalloproteinase-8 (MMP8) on tumorigenesis and metastasis. We investigated whether *MMP8* gene variation was associated with breast cancer metastasis and prognosis in humans. We first studied nine tagging single nucleotide polymorphisms (SNP) in the *MMP8* gene in 140 clinically and pathologically well-characterized breast cancer patients. Four of the SNPs were found to be associated with lymph node metastasis, the most pronounced being a promoter SNP (rs11225395) with its minor allele (*T*) associating with reduced susceptibility to lymph node metastasis ($P = 0.02$). This SNP was further evaluated for association with cancer relapse and survival among a cohort of ~1,100 breast cancer patients who had been followed for cancer recurrence and mortality for a median of 7.1 years. The *T* allele was associated with reduced cancer relapse and greater survival, particularly among patients with earlier stage cancer. Among patients of tumor-node-metastasis stage 0 to II, the adjusted hazard ratio of disease-free survival was 0.7 [95% confidence interval (95% CI), 0.5–0.9] for patients carrying *T* allele compared with those homozygous for the *C* allele ($P = 0.02$). *In vitro* experiments showed that the *T* allele had higher promoter activity than the *C* allele in breast cancer cells. Electrophoretic mobility shift assays showed binding of nuclear proteins to the DNA sequence at the SNP site of the *T* allele but not that of the *C* allele. The data suggest that *MMP8* gene variation may influence breast cancer prognosis and support the notion that *MMP8* has an inhibitory effect on cancer metastasis. [Cancer Res 2007;67(21):10214–21]

Introduction

Breast cancer is the most common malignancy and a leading cause of cancer mortality among women in the United States,

Europe, and many other parts of the world (1–3). The vast majority of these deaths result from metastasis. Currently, prognostic stratification of patients is carried out according to the tumor-node-metastasis (TNM) staging system, which classifies breast cancers based on the extent of the tumor, spread to lymph nodes, and metastasis (4).

A better understanding of the mechanisms underlying cancer cell spreading and the identification of factors that influence these processes could potentially help identify high-risk patients and develop novel therapeutic measures. Recent studies have shown that germ line genetic variations contribute to interindividual variability in breast cancer susceptibility and might influence the likelihood of metastasis (5, 6). Such germ line genetic variations are exemplified by *BRCA1* and *BRCA2* gene mutations (5, 6). However, the genes identified thus far explain only ~20% of breast cancer cases, suggesting that much of the variability in breast cancer susceptibility and progression is related to variations in many unidentified genes which likely have only moderate individual effects (5, 7).

Matrix metalloproteinases (MMP) can degrade extracellular matrix constituents and have long been believed to play an important role in cancer cell spreading, which entails destruction of the extracellular matrix. Recent studies show that MMPs also have a number of other properties that can impact on tumorigenesis and metastasis. For instance, MMPs can process and activate a number of chemokines and growth factors and thereby affect cell proliferation, cell differentiation, apoptosis, and angiogenesis (8).

MMP8 (also known as neutrophil collagenase and collagenase-2) was originally thought to be produced exclusively by neutrophils, but has subsequently been found to be expressed also by a variety of other cell types (9). Expression of *MMP8* in squamous cell carcinomas and in breast cancer cells has also been shown (10–13). Evidence from recent studies indicates that *MMP8* has an antitumor property. For instance, it has been shown that mice lacking an active *MMP8* gene have increased susceptibility to skin tumor compared with *MMP8* wild-type mice (14). Moreover, studies have indicated that *MMP8* expression has an effect on the metastatic behavior of breast cancer cells, such that increased *MMP8* expression in breast cancer cells inhibits their ability to spread (11, 12).

A recent report indicates that in humans, single nucleotide polymorphisms (SNP) in the *MMP8* gene have an influence on *MMP8* expression (15). Because there is evidence of *MMP8* having an inhibitory effect on the metastatic ability of breast cancer cells as mentioned above, we hypothesized that breast cancer patients

Note: Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

J. Decock, J.-R. Long, and R.C. Laxton contributed equally to this work.

Requests for reprints: Shu Ye, William Harvey Research Institute, John Vane Science Building, Charthouse Square, London EC1M 6BQ, United Kingdom. Phone: 44-207-882-3425; Fax: 44-207-882-3408; E-mail: s.ye@qmul.ac.uk and Wei Zheng, Vanderbilt Epidemiology Centre, 2525 West End Avenue, 8th floor, Nashville, TN 37203. Phone: 615-936-0682; Fax: 615-936-8241; E-mail: wei.zheng@vanderbilt.edu.

©2007 American Association for Cancer Research.

doi:10.1158/0008-5472.CAN-07-1683

of certain *MMP8* genotypes are predisposed to cancer spreading. To test this hypothesis, we studied nine tagging SNPs in the *MMP8* gene in a group of breast cancer patients. Four of the SNPs studied were found to be associated with lymph node involvement, the most significant association being with an SNP (rs11225395) in the promoter of the gene. This SNP was further analyzed in a large cohort of breast cancer patients who have been recruited to a population-based epidemiology study, the Shanghai Breast Cancer Study, to evaluate the association of this SNP with breast cancer relapse and survival. *In vitro* experiments were undertaken to examine whether this SNP has an effect on *MMP8* promoter activity in breast cancer cells.

Subjects and Methods

Subjects

The Leuven Breast Cancer Study. A group of 140 consecutive patients with primary operable breast cancer, newly diagnosed at the Multidisciplinary Breast Centre of the University Hospital of Leuven between 2005 and 2006, were recruited to this study. Patient characteristics were extracted from clinical files, tumor characteristics and lymph node status were retrieved from the pathology reports, and all data were eventually collected in one central database. None of the patients received neoadjuvant treatment or had any history of cancer, bilateral cancer, or Paget's disease of the nipple. The study was conducted in the frame of the European Union Framework 6 Cancerdegradome Project, and its design was approved by the Medical Ethical Committee of the University Hospital Leuven.

A 10-mL peripheral blood sample for DNA analysis was collected from each patient. Tumor tissues for histologic examinations were snap-frozen in liquid nitrogen within 1 h after surgery. Typing of primary tumors was done according to the WHO classification, and the Ellis and Elston system was used for histologic grading. Pathologic examination of lymph nodes (three sections per node) was done with H&E staining. Sentinel lymph nodes classified as negative were additionally stained for epithelial markers.

The Shanghai Breast Cancer Study. Included in this study are 1,455 breast cancer patients who were recruited as part of the Shanghai Breast Cancer Study, a population-based case-control study conducted among Chinese women in the Shanghai urban area (16, 17). These patients were permanent residents of Shanghai who were diagnosed with an incident breast cancer between August 1996 and March 1998 and were between the ages of 25 and 64 years. Details of the study design and study population have been described elsewhere (16). The overall response rate was 91%. A peripheral blood sample (10 mL from each woman) was obtained from 1,193 patients, 82% of the 1,455 study participants. The blood samples were processed within 6 h of collection and stored at -70°C until the relevant bioassays were conducted. Medical charts were reviewed to abstract information on cancer diagnosis, TNM stage, and cancer treatment. Pathologic slides for all cases were reviewed independently by two senior pathologists to confirm cancer diagnosis.

Patients were followed until July 2005 for cancer recurrence and mortality using a combination of two active follow-up surveys and record linkage to death certificates kept by the Vital Statistics Unit of the Shanghai Municipal Centre for Disease Control and Prevention. The median follow-up time for the cohort was 7.1 years (18). During the follow-up period, 313 deaths were identified: 283 from breast cancer, 27 from other diseases, and 3 from unclear causes. The study was approved by the institutional review boards of all participating institutes. Informed consent was obtained from each participant.

DNA Extraction from Peripheral Blood Samples

DNA extraction from the peripheral blood samples of the Leuven Breast Cancer Study was carried out using a salting-out method (19). In the Shanghai Breast Cancer Study, genomic DNA was extracted from buffy coats (WBC) using Puregene DNA Purification Kits (Gentra Systems) following the manufacturer's protocol.

SNP Selection

Tagging SNPs were selected from the HapMap database.⁷ As in January 2006, the database contained 35 common SNPs (minor allele frequency >0.05) at the *MMP8* locus, including 3.5 kb upstream and 3.5 kb downstream of the gene. The Tagger algorithm incorporated in the HapMap Web site was used to identify a set of tagging SNPs to cover the 35 common SNPs at this locus with minimum pairwise correlation coefficients (r^2) of 0.8 and was configured to include the promoter SNPs rs11225395 and rs1320632, which had been suggested to influence *MMP8* expression (15) and the nonsynonymous coding SNP rs1940475. The analysis identified nine tagging SNPs: rs10895353, rs7943404, rs11225395, rs1320632, rs1940475, rs1892886, rs17099436, rs2508383, and rs1276284 (Supplementary Fig. S1A).

Genotyping

The Leuven Breast Cancer Study subjects were genotyped for the nine SNPs described above using the TaqMan method. Primers and probes were designed and synthesized by Applied Biosystems (ABI Assays-by-design), with Applied Biosystems VIC (VIC) and 6-carboxy-fluorescein (FAM) labeled probes, respectively, to detect the two different alleles of each SNP. The assays were done on an ABI Prism 7900 Sequence Detector with the Allelic Discrimination Sequence Detector Software from Applied Biosystems. The Shanghai Breast Cancer Study subjects were genotyped for SNP rs11225395 using the TaqMan method with Applied Biosystems Assay-on-Demand (Assay ID C_1366493_10) primers and probes, and genotype was successfully determined for 1,100 cases. The laboratory staff was blind to the identity of the subjects. Quality control samples were included in genotyping assays. In the Leuven Breast Cancer Study, each 96-well plate contained blank (water) controls and duplicate DNA controls. In the Shanghai Breast Cancer Study, each 96-well plate contained one well of water, two wells of CEPH 1347-02 DNA, two wells of blinded quality control DNA, and two wells of unblinded quality control DNA samples, where the blinded and unblinded quality control samples were taken from the second tube of the samples included in the study, and the genotype consistency rate for the 56 study quality control samples was 95.5%. In addition, DNA of Chinese samples that were used in HapMap ($n = 45$) and Perlegen ($n = 24$) were purchased from Coriell Cell Repositories⁸ and genotyped for this SNP. The consistence rate of this SNP was 95.2% compared with the data from HapMap⁷ and Perlegen.⁹

Construction of *MMP8* Promoter-Luciferase Reporter Gene Plasmids

For each of the two alleles of the rs11225395 SNP, which arises from a *C* to *T* substitution at position -799 bp (relative to the transcriptional start site) in the *MMP8* gene promoter, a 968-bp DNA fragment from position -872 to $+96$ bp was generated by PCR in which DNA from an individual homozygous for that allele was used as template. Each of the two amplicons was first cloned into the pCR-Blunt II-TOPO vector (Invitrogen) and then subcloned into the pGL3-basic vector (Promega) with the *MMP8* gene promoter being placed upstream of the firefly luciferase reporter gene. All constructs were sequenced to verify that the cloned *MMP8* gene promoter was in the desired orientation and free from misincorporation of nucleotide during the PCR.

Cell Culture

The MDA-MB-231 breast cancer cell line was purchased from the European Collection of Animal Cell Cultures (ECACC) and cultivated following instructions provided by ECACC. In brief, the cells were cultured in Leibovitz's L15 medium supplemented with 2 mmol/L L-glutamine, 1.0 mmol/L sodium pyruvate, 0.1 mmol/L nonessential amino acids, 10% heat-inactivated fetal bovine serum, 100 $\mu\text{g}/\text{mL}$ streptomycin, and 100 units/mL penicillin. The cells were incubated at 37°C with 100% air.

⁷ <http://www.hapmap.org>

⁸ <http://locus.umdnj.edu/ccr/>

⁹ <http://genome.perlegen.com>

Transient Transfection and Luciferase Reporter Assay

MDA-MB-231 breast cancer cells were transfected with the reporter gene constructs described above. Transfection was done with the use of Effectene (Qiagen) in duplicate for each of the constructs. The pRL-TK (Promega) plasmid, which contains a *Renilla* luciferase gene, was co-transfected into the cells to serve as a reference for transfection efficiency. The cells were then lysed, and the activities of firefly luciferase and *Renilla* luciferase were measured using a dual-luciferase assay kit (Promega). *MMP8* promoter activity was measured as a ratio of firefly luciferase activity to *Renilla* luciferase activity, and the means (and SEs) from three independent experiments are presented.

Electrophoretic Mobility Shift Assay

Biotin-labeled, double-stranded oligonucleotide probes (*C*: 5'-CCAT-GCAGAGCCATAGTAGCTCC-3' and *T*: 5'-CCATGCAGAGCTTATAG-GTAGCTCC-3') corresponding to the sequence from nucleotide position

–810 to –787 in the *MMP8* gene promoter, with either a *C* or *T* at the rs11225395 SNP site (position –799), were synthesized. The probes were incubated with nuclear protein extracts from MDA-MB-231 breast cancer cells, in the presence or absence of competitors, i.e., unlabeled probe *C* (referred to as competitor *C*), unlabeled probe *T* (referred to as competitor *T*), or a nonspecific sequence (referred to as nonspecific competitor). Protein-DNA complexes were resolved by PAGE and detected using a LightShift Chemiluminescent EMSA kit (Pierce Biotechnology). Three independent experiments were carried out.

Statistical Analyses

Allele and genotype frequencies were calculated by the gene counting method. χ^2 analyses were done to test genotypic and allelic association with clinical and pathologic parameters. SNP pair-wise linkage disequilibrium coefficient, haplotype frequencies, and haplotypic effects on lymph node metastasis were determined by using the THESIAS program, which

Table 1. *MMP8* SNPs in relation to lymph node metastasis in the Leuven Breast Cancer Study

Genotype	Number (%)		P^*	P^\dagger	Allele	Number (frequency)		OR (95% CI)	P^\ddagger	
	LN+	LN–				LN+	LN–			
rs10895353										
AA	56 (73.7)	49 (81.7)	0.23	0.17	<i>A</i>	132 (0.87)	108 (0.90)	1.4 (0.6, 2.9)	0.42	
AG	20 (26.3)	10 (16.7)								<i>G</i>
GG	–	1 (1.7)								
rs7943404										
TT	26 (32.9)	26 (42.6)	0.27	0.12	<i>T</i>	82 (0.52)	76 (0.62)	1.5 (1.0, 2.5)	0.08	
CT	30 (32.9)	24 (39.3)								<i>C</i>
CC	23 (39.1)	11 (18.0)								
rs11225395										
CC	36 (46.2)	20 (32.8)	0.09	0.04	<i>C</i>	103 (0.66)	64 (0.52)	0.6 (0.4, 0.9)	0.02	
CT	31 (39.7)	24 (39.3)								<i>T</i>
TT	11 (14.1)	17 (27.9)								
rs1320632										
AA	69 (87.3)	51 (83.6)	0.53	0.53	<i>A</i>	148 (0.94)	112 (0.92)	0.8 (0.3, 1.9)	0.55	
AG	10 (12.7)	10 (16.4)								<i>G</i>
rs1940475										
AA	35 (44.3)	17 (27.9)	0.12	0.50	<i>A</i>	97 (0.61)	59 (0.48)	0.6 (0.4, 1.0)	0.03	
AG	27 (34.2)	25 (41.0)								<i>G</i>
GG	17 (21.5)	19 (31.1)								
rs1892886										
TT	51 (67.1)	28 (45.9)	0.04	0.03	<i>T</i>	123 (0.81)	85 (0.70)	0.5 (0.3, 1.0)	0.03	
AT	21 (27.6)	29 (47.5)								<i>A</i>
AA	4 (5.3)	4 (6.6)								
rs17099436										
TT	73 (92.4)	55 (90.2)	0.64	0.64	<i>T</i>	152 (0.96)	116 (0.95)	0.8 (0.2, 2.4)	0.65	
AT	6 (7.6)	6 (9.8)								<i>A</i>
rs2508383										
CC	55 (70.5)	41 (69.5)	0.73	0.90	<i>C</i>	131 (0.84)	97 (0.82)	0.9 (0.5, 1.7)	0.70	
CT	21 (26.9)	15 (25.4)								<i>T</i>
TT	2 (2.6)	3 (5.1)								
rs1276284										
GG	47 (59.6)	22 (36.7)	0.02	0.04	<i>G</i>	116 (0.73)	73 (0.61)	0.6 (0.3, 0.9)	0.03	
AG	22 (27.8)	29 (48.3)								<i>A</i>
AA	10 (12.7)	9 (15.0)								

NOTE: LN+, patients with lymph node metastasis; LN–, patients without lymph node metastasis.

Abbreviation: OR, odds ratio.

* P value for genotypic χ^2 test.

† P value for trend test.

‡ P value for allelic χ^2 test.

Table 2. Major *MMP8* haplotypes in relation to lymph node metastasis in the Leuven Breast Cancer Study

Haplotype identification	Haplotype composition*	Haplotype frequency		Odds ratio (95% CI) for lymph node metastasis	P
		LN+	LN-		
Hap 1	ACCAATTCG	0.45	0.36	1.00 (reference)	
Hap 2	ATTAGATCA	0.18	0.30	0.52 (0.29–0.96)	0.04
Hap 3	GTCAATTCG	0.13	0.09	1.29 (0.52–3.22)	0.56
Hap 4	ATTGGTTTA	0.06	0.08	0.75 (0.29–1.96)	0.57

NOTE: LN+, patients with lymph node metastasis; LN-, patients without lymph node involvement.

*Haplotypes deriving from the following SNPs in order: rs10895353; rs7943404, rs11225395, rs1320632, rs1940475, rs1892886, rs17099436, rs2508383 and rs1276284.

implements a stochastic EM (Expectation-Maximization) algorithm (20). Cancer recurrence/metastasis and death due to breast cancer were the study outcomes for the analysis of disease-free survival, and death from any cause was the study outcome for the analysis of overall survival. Survival time was calculated as the time from cancer diagnosis until death (for overall survival) or the occurrence of study outcomes (cancer recurrence/metastasis and death due to breast cancer). Observations were censored at the date of last contact or noncancer death (for disease-free survival). The Kaplan-Meier method was used to estimate the survival rate, and differences in survival across genotype groups were examined using the log-rank test. The Cox proportional hazard model was employed to compute hazard ratios (HR). All statistical analyses were done using SAS version 9.1 (SAS Institute), and all *P*'s were for two-sided tests.

Results

The Leuven Breast Cancer Study. The median age at diagnosis of breast cancer in this patient group was 56 (range, 33–86). Of the 140 patients, 87 were postmenopausal, and 53 were premenopausal. The tumor size was small (≤ 2 cm) in 68 patients and large (> 2 cm) in 72 patients. The tumors from 14 patients were well differentiated, the tumors from 52 patients were moderately differentiated, and the tumors from 72 patients were poorly differentiated. Of the 140 patients, 61 had no lymph node involvement, and 79 were lymph node positive. Using the American Joint Committee on Cancer TNM staging system, 39 patients were classified as stage I, 74 were classified as stage II, and the remaining 27 were classified as stage III.

The patients were genotyped for a panel of nine SNPs in the *MMP8* gene (Supplementary Fig. S1A, SNP selection is described in Materials and Methods). No association was detected between the SNPs and tumor size, histologic grade, or tumor subtype. However, there was an association between lymph node metastasis with four SNPs studied, i.e., rs11225395, rs1940475, rs1892886, and rs1276284, with the minor alleles (*T* allele of SNP rs11225395, *G* allele of SNP rs1940475, *A* allele of SNP rs1892886, and *A* allele of SNP rs1276284) having lower frequencies in patients with lymph node metastasis than in patients without lymph node involvement ($P = 0.02$, $P = 0.03$, $P = 0.03$, and $P = 0.03$, respectively, Table 1).

Haplotype analysis showed that there were four major haplotypes (each having a frequency $> 5\%$) deriving from the nine SNPs studied (Table 2). Compared with the most common haplotype ACCAATTCG (Hap 1), the ATTAGATCA haplotype (Hap 2) was associated with a lower odds ratio [0.52; 95% confidence interval (95% CI), 0.29–0.96] for lymph node metastasis ($P = 0.04$, Table 2). Compared with the most common haplotype (Hap 1), the GTCAATTCG haplotype (Hap 3) had a higher odds ratio (1.29; 95% CI, 0.52–3.22), whereas the ATTGGTTTA haplotype (Hap 4) had a lower odds ratio (0.75; 95% CI, 0.29–1.96); however, the differences were not statistically significant (Table 2).

The Shanghai Breast Cancer Study. The subjects of the Shanghai Breast Cancer Study were genotyped for the *MMP8* rs11225395 SNP which (*a*) was per se shown to be associated with lymph node metastasis in the Leuven study described above; (*b*)

Table 3. *MMP8* SNP rs11225395 in relation to TNM stage and breast cancer relapse in the Shanghai Breast Cancer Study

Genotype	TNM stage				P	Cancer relapse in all patients			Cancer relapse in early stage (0–II) patients		
	0–I, n (%)	IIa, n (%)	IIb, n (%)	III–IV, n (%)		No relapse, n (%)	Relapse, n (%)	P	No relapse, n (%)	Relapse, n (%)	P
CC	100 (36.0)	145 (37.3)	94 (39.3)	47 (38.8)		303 (73.9)	107 (26.1)		256 (76.4)	79 (23.6)	
CT + TT	178 (64.0)	244 (62.7)	145 (60.7)	74 (61.2)	0.87*	514 (78.1)	144 (21.9)	0.11*	451 (82.2)	98 (17.9)	0.04*
CT	143 (51.4)	195 (50.1)	108 (45.2)	54 (44.6)	0.69 [†]	408 (79.2)	107 (10.8)	0.13 [†]	357 (82.3)	77 (17.7)	0.12 [†]
TT	35 (12.6)	49 (12.6)	37 (15.5)	20 (16.5)		106 (74.1)	37 (25.9)		94 (81.7)	21 (18.3)	

**P* value for the comparison of CT + TT versus CC.

[†]*P* value for trend test comparing CC, CT, and TT.

Table 4. *MMP8* SNP rs11225395 in relation to survival among breast cancer patients in the Shanghai Breast Cancer Study

Population	Genotype	Overall survival				Disease-free survival			
		Cases	Events	HR (95% CI)	<i>P</i>	Cases	Events	HR (95% CI)	<i>P</i>
All women	CC	415	98	1.0 (reference)		410	127	1.0 (reference)	
	CT + TT	681	139	0.9 (0.7, 1.1)	0.23*	658	172	0.8 (0.7, 1.0)	0.09*
	CT	530	109	0.9 (0.7, 1.1)	0.25 [†]	515	131	0.8 (0.6, 1.0)	0.24 [†]
	TT	151	30	0.8 (0.5, 1.2)		143	41	0.9 (0.6, 1.3)	
TNM 0-II	CC	339	70	1.0 (reference)		335	93	1.0 (reference)	
	CT + TT	567	83	0.7 (0.5, 1.0)	0.02*	549	115	0.7 (0.5, 0.9)	0.02*
	CT	446	69	0.7 (0.5, 1.0)	0.01 [†]	434	91	0.7 (0.5, 1.0)	0.03 [†]
	TT	121	14	0.5 (0.3, 1.0)		115	24	0.7 (0.5, 1.1)	
TNM III-IV	CC	47	19	1.0 (reference)		46	24	1.0 (reference)	
	CT + TT	74	40	1.5 (0.9, 2.6)	0.14*	70	42	1.3 (0.8, 2.2)	0.31*
	CT	54	30	1.6 (0.9, 2.8)	0.26 [†]	52	30	1.2 (0.7, 2.1)	0.22 [†]
	TT	20	10	1.4 (0.6, 3.0)		18	12	1.5 (0.8, 3.1)	

**P* value for the comparison of CT + TT versus CC.

[†]*P* value for trend test comparing CC, CT, and TT.

tagged the four major *MMP8* haplotypes (with the *T* allele tagging Hap 2 and Hap 4, which were associated with reduced metastasis, whereas the *C* allele tagging Hap 1 and Hap 3 associating with high metastasis risk, in the Leuven Study; Table 2), (c) was found to be in

linkage disequilibrium with the three other SNPs (rs1940475, rs1892886, and rs1276284, Supplementary Fig. S1B) that were associated with lymph node metastasis in the Leuven study, and (d) had previously been suggested to have an effect on *MMP8*

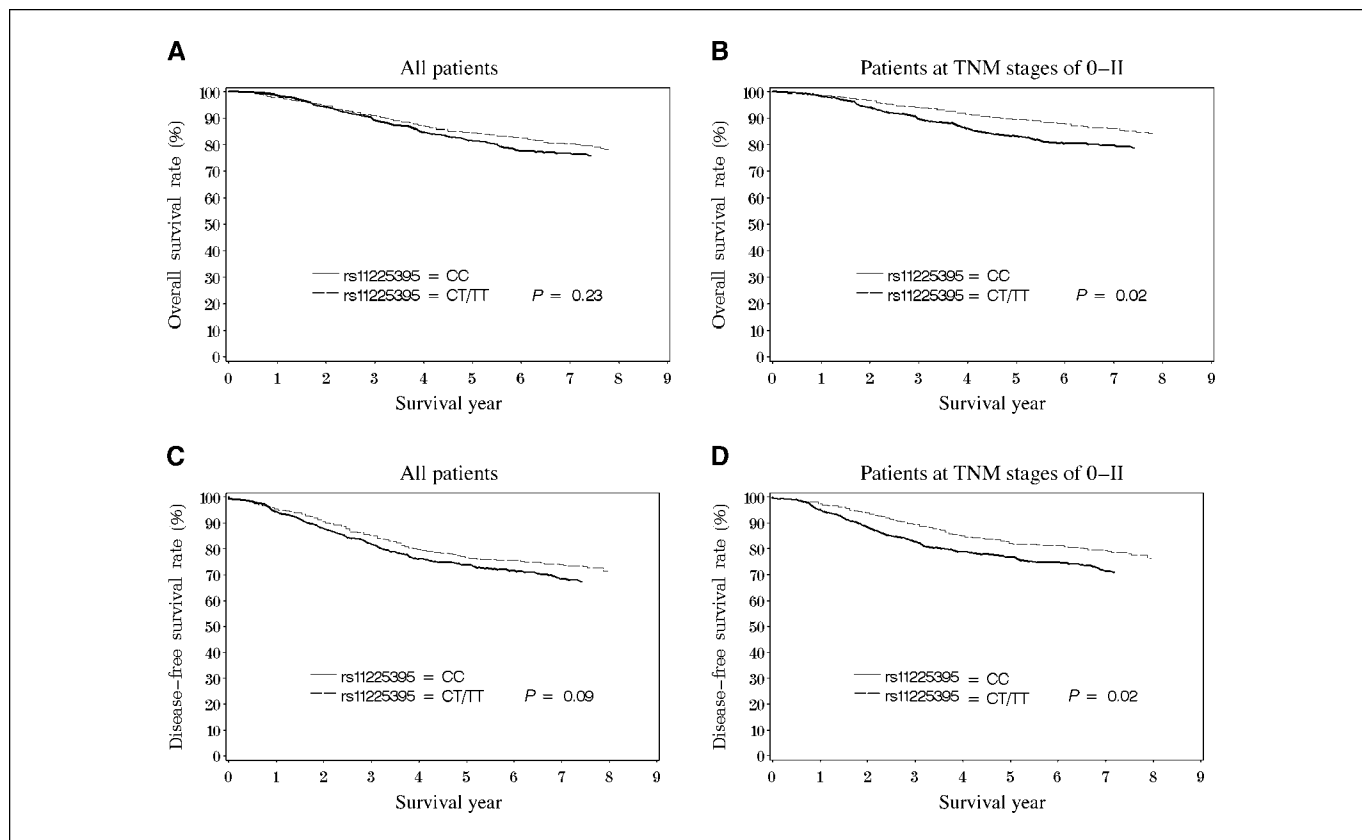


Figure 1. Overall and disease-free survival of breast cancer patients after diagnosis stratified by genotypes of SNP rs11225395. A, overall survival among all women; (B), overall survival among women at TNM stages of 0 to II; (C), disease-free survival among all women; (D), disease-free survival among women at TNM stages of 0 to II.

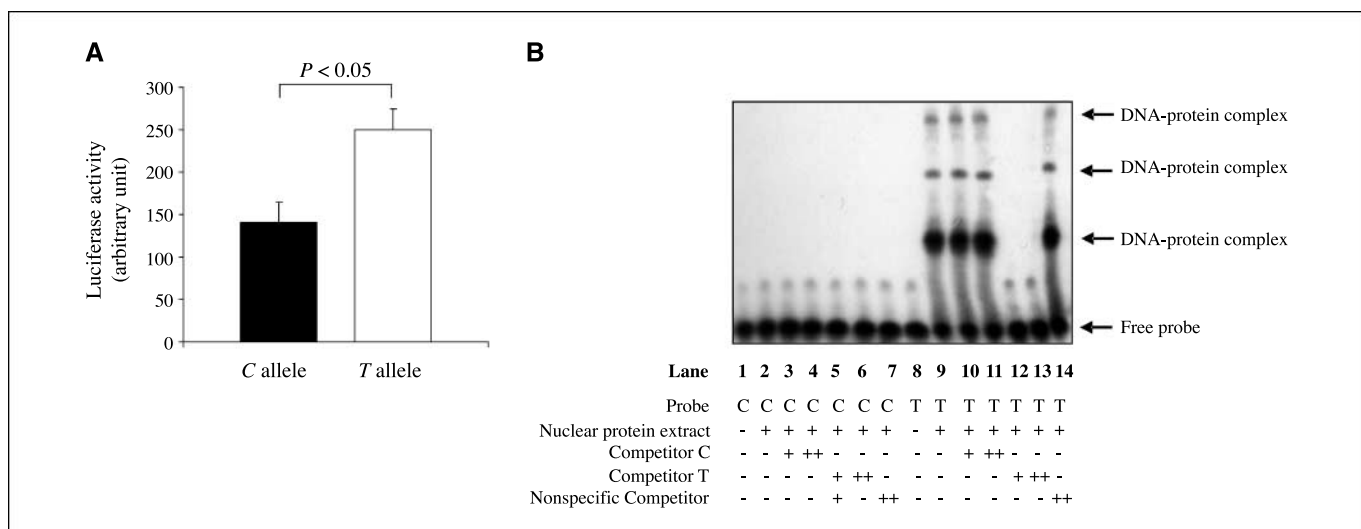


Figure 2. A, *MMP8* gene promoter activity of the T and C alleles of the -799C>T SNP (rs11225395). Columns, relative *MMP8* promoter activity of the T and C alleles in MDA-MB-231 breast cancer cells, measured by dual-luciferase assays as described in Materials and Methods. Data shown are mean (\pm SD) from three independent experiments. B, representative results of electrophoretic mobility shift assays. Nuclear protein extracts from MDA-MB-231 breast cancer cells were incubated with biotin-labeled probes corresponding to the C allele (lanes 1–7) or the T allele (lanes 8–14) of the -799C>T SNP (rs11225395) in the absence or presence of competitors. Lanes 1 and 8, no nuclear protein extract; lane 2 and 9, no competitor; lanes 3 and 10, competitor C in 20-fold molar excess (+); lanes 4 and 11, competitor C in 50-fold molar excess (++); lanes 5 and 12, competitor T in 20-fold molar excess (+); lanes 6 and 13, competitor T in 50-fold molar excess (++); lanes 7 and 14, nonspecific competitor in 50-fold molar excess (++).

expression (15). The genotype distribution of the rs11225395 SNP in the Shanghai breast cancer patients stratified by TMN stage and cancer relapse status is presented in Table 3. More patients with an early-stage cancer carried the minor allele (T) of the SNP than those with a late-stage cancer, although the difference was not statistically significant. Consistent with this finding, patients who carried the T allele tended to have a lower rate of relapse than those homozygous for the C allele, and this association was more pronounced in patients with an early-stage cancer (TMN stage 0–II; $P = 0.04$, Table 3).

Patients carrying the T allele had a higher survival rate, compared with those homozygous for the C allele (Table 4). This association was more apparent in patients with an early-stage cancer (TNM stage 0–II). The adjusted HR for T allele carriers were 0.7 (95% CI, 0.5–1.0; $P = 0.02$) for overall survival and 0.7 (95% CI, 0.5–0.9; $P = 0.02$) for disease-free (cancer recurrence/metastasis-free) survival. Additional adjustment for other prognostic factors for breast cancer, including TNM stage and cancer treatments, did not change the pattern of the association. The Kaplan-Meier survival curves presented in Fig. 1 show the tendency of increased overall survival and disease-free survival in patients with the T allele during the follow-up period.

Allele-specific differences in *MMP8* gene promoter activity. The rs11225395 SNP arises from a C to T nucleotide substitution at position -799 bp (relative to the transcriptional start site) in the *MMP8* promoter. We carried out reporter assays in MDA-MB-231 breast cancer cells to investigate whether this SNP has an effect on *MMP8* promoter activity in breast cancer cells. The experiments showed that the T allele had approximately 1.8-fold higher promoter activity than the C allele ($P < 0.05$, Fig. 2).

To investigate whether the sequence at the rs11225395 SNP site interacted with nuclear proteins of breast cancer cells and, if so, whether this differed between the C and T alleles, we did electrophoretic mobility shift assays in which oligonucleotide probes

corresponding to the C or T allele were incubated with nuclear protein extracts from MDA-MB-231 breast cancer cells. Although no DNA-protein complexes were detected using the probe corresponding to the C allele (Fig. 2B, lane 2), three DNA-protein complexes were detected using the T allele probe (Fig. 2B, lane 9). The bands of the DNA-protein complexes diminished when large amounts of an unlabeled T allele oligonucleotide were included in the assay to serve as a competitor for the T allele probe (Fig. 2B, lanes 12 and 13). In contrast, the above DNA-protein complexes bands were unaffected by the presence of large amounts of an unlabeled C allele oligonucleotide (Fig. 2B, lanes 10 and 11) or large amounts of a nonspecific oligonucleotide competitor (Fig. 2B, lane 14).

Discussion

In this study, we first found an association of SNPs (rs11225395, rs1940475, rs1892886, and rs1276284) in the *MMP8* gene with lymph node metastasis in a group of breast cancer patients. In particular, the T allele of the promoter SNP rs11225395, which had previously been suggested to have an influence on *MMP8* expression, was found to be associated with reduced risk of lymph node metastasis. Subsequently, in a second investigation conducted in a large group of Chinese breast cancer patients, we found that the T allele of this SNP was associated with reduced cancer relapse and increased survival. *In vitro* experiments showed that the T allele had increased promoter activity in MDA-MB-231 breast cancer cells, suggesting that the SNP could potentially exert a direct functional effect. Further experiments showed that the DNA sequence at the SNP site of the T allele interacted with nuclear proteins from MDA-MB-231 breast cancer cells, and that this DNA-protein interaction was not detected for the C allele. Taken together, these results are consistent with the notion that *MMP8* has an inhibitory effect on breast cancer metastasis as indicated by the animal and cell studies described previously (11, 12).

Lymph nodes are often the first sites to which breast cancer cells spread (4). Studies have shown that lymph node involvement at primary diagnosis of breast cancer patients predicts an unfavorable outcome after first recurrence, independent of the site of relapse and disease-free interval (21). This suggests that primary lymph node involvement is an indicator of breast cancers that have an aggressive metastatic behavior, rather than merely a time-dependent phenomenon (21). In this study, we found an association between lymph node metastasis and MMP8 SNPs, especially SNP rs11225395. Furthermore, we observed that the minor allele of this functional SNP was associated with reduced cancer relapse and improved survival after breast cancer diagnosis in an independent study conducted in Chinese women.

The association of the *T* allele of the rs11225395 SNP with better prognosis of breast cancer in the Chinese women was more apparent in those with an early-stage cancer (TNM stage 0–II) where spread of cancer cells was restricted to lymph nodes and there was no involvement of other tissues or organs. This may indicate that MMP8 has a greater protective effect against lymph node metastasis, as compared with distant metastasis to other organs. In agreement, animal studies showed that reducing MMP8 levels by the ribozyme knockdown technique dramatically increased lymph node metastasis of breast cancer cells orthotopically implanted in mice, but has a smaller effect on metastasis to the lung; thus, it was thought that a reduction in MMP8 level predisposes breast cancer cells to travel, survive, and grow better in the lymphatic system than by the hematogenous route to other organs (11).

The functional analyses of the rs11225395 SNP indicates that it has an effect on *MMP8* promoter activity in breast cancer cells, with the *T* allele having a greater strength in driving gene expression. The electrophoretic mobility shift assays reveal that the *T* allele, but not the *C* allele, interacts with nuclear protein(s) from breast cancer cells, with three DNA-protein complexes detected, which may have derived from different nuclear proteins or different fractions of the same protein. Because the *T* allele has higher promoter activity than the *C* allele, it is possible that the nuclear protein(s) selectively binding to the *T* allele is a transcriptional enhancer and accounts for the higher promoter activity of the *T* allele. An *in silico* scan of the DNA sequence at the rs11225395 SNP site did not reveal a good match with consensus cis-elements catalogued in transcription factor binding site databases. Because not all transcription factors and their binding sites have been identified to date and, thus, the current databases do not have a complete coverage, it is possible that the nuclear protein(s) binding to the *T* allele is a transcription factor not contained in the databases.

In the Shanghai study, we focused on the rs11225395 SNP, because among the nine SNPs examined in the Leuven study, the rs11225395 SNP was the most significantly associated with lymph node metastasis, and because it tagged the four major *MMP8* haplotypes, i.e., the *T* allele tagged Hap 2 and Hap 4, which were associated with reduced metastasis, whereas the *C* allele tagged Hap 1 and Hap 3, which had higher metastasis risk, in the Leuven study. In addition, this SNP was in linkage disequilibrium with the other three SNPs (i.e., rs1940475, rs1892886, and rs1276284) that were associated with metastasis in the Leuven study and had previously been suggested to have an effect on *MMP8* expression (15).

The minor alleles of SNPs rs11225395, rs1940475, rs1892886, and rs1276284 were associated with lower risk of lymph node

involvement. It is possible that some of these SNPs are only genetic markers that mark the effects of certain functional SNPs through linkage disequilibrium. The functional analysis of SNP rs11225395 indicates that it has an effect on *MMP8* promoter activity, which is consistent with our results of the association between the *T* allele with reduced risk of lymph node metastasis and cancer relapse and improved survival. It remains unknown whether SNPs rs1940475, rs1892886, and rs1276284 also have direct functional effects. The rs1940475 SNP results in a glutamate to lysine substitution at amino acid residue 87 located in the propeptide of *MMP8* and might potentially affect pro-*MMP8* activation. The rs1892886 SNP is located in intron 4, and the rs1276284 SNP is in the 3' flanking region of the *MMP8* gene. The locations of these two latter SNPs suggest that they are less likely to be functional, although this cannot be precluded. In addition, it is possible that these SNPs (rs11225395, rs1940475, rs1892886, and rs1276284) might mark the effect of other SNPs that were not examined in this study.

SNPs in several other *MMP* genes have been associated with susceptibility and/or spreading of cancers, the most extensively studied being the *MMP1* –1607 1G>2G polymorphism. The 2G allele of this polymorphism has higher *MMP1* expression levels than the 1G allele (22, 23) and has been reported to be associated with increased susceptibility and/or metastasis of melanoma, breast cancer, ovarian cancer, lung cancer, and colorectal cancer (22–31). The present study suggests that SNPs in the *MMP8* gene are also associated with breast cancer metastasis, and that in contrast to the 2G allele of the *MMP1* –1607 1G>2G polymorphism, the high-expression *T* allele of the *MMP8* –799C>T SNP (rs11225395) is associated with lower susceptibility to metastasis and better survival in breast cancer patients. It would be interesting to investigate whether *MMP8* variation is also related to susceptibility and/or metastasis of other cancers.

A limitation of this study is the relatively small sample size in the Leuven study. However, the promising association between SNP rs11225395 with breast cancer metastasis was replicated in an independent study with a larger sample size. The strong study methodology in the Shanghai study, including a population-based design, high response rate, and high follow-up rate, lends credibility to the results reported in this paper. A potential concern for the Shanghai study, however, is that the information regarding cancer relapse and cause of death was collected based on self-report or death certificate, which might have compromised the accuracy of these outcome data. Potential errors in outcome assessment, however, are likely to be random, which tend to attenuate the association observed. Although the results from our study need to be replicated in other studies, our data suggest that genetic variation in the *MMP8* gene can influence breast cancer prognosis and support the notion that *MMP8* has an inhibitory effect on breast cancer metastasis.

Acknowledgments

Received 5/8/2007; revised 8/31/2007; accepted 9/7/2007.

Grant support: European Union Framework Programme 6 Project LSHC-CT-2003-503297 (the CancerDegradome) and research grants (R01CA64227 and R01CA090899) from National Cancer Institute. J. Decock and A.C. Pereira were supported by grants from the "Vlaamse Liga tegen Kanker" and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) in Brazil (PDEE 2730/05-7), respectively.

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.

References

1. Bray F, Sankila R, Ferlay J, et al. Estimates of cancer incidence and mortality in Europe in 1995. *Eur J Cancer* 2002;38:99-166.
2. Micheli A, Mugno E, Krogh V, et al. Cancer prevalence in European registry areas. *Ann Oncol* 2002;13:840-65.
3. Edwards BK, Brown ML, Wingo PA, et al. Annual report to the nation on the status of cancer, 1975-2002, featuring population-based trends in cancer treatment. *J Natl Cancer Inst* 2005;97:1407-27.
4. Singletary SE, Connolly JL. Breast cancer staging: working with the sixth edition of the AJCC cancer staging manual. *CA Cancer J Clin* 2006;56:37-47.
5. Houlston RS, Peto J. The search for low-penetrance cancer susceptibility alleles. *Oncogene* 2004;23:6471-6.
6. Foulkes WD, Rosenblatt J, Chappuis PO. The contribution of inherited factors to the clinicopathological features and behavior of breast cancer. *J Mammary Gland Biol Neoplasia* 2001;6:453-65.
7. Thompson D, Easton D. The genetic epidemiology of breast cancer genes. *J Mammary Gland Biol Neoplasia* 2004;9:221-36.
8. Egeblad M, Werb Z. New functions for the matrix metalloproteinases in cancer progression. *Nat Rev Cancer* 2002;2:161-74.
9. Van LP, Libert C. Matrix metalloproteinase-8: cleavage can be decisive. *Cytokine Growth Factor Rev* 2006;17:217-23.
10. Moilanen M, Pirila E, Grenman R, et al. Expression and regulation of collagenase-2 (MMP-8) in head and neck squamous cell carcinomas. *J Pathol* 2002;197:72-81.
11. Montel V, Kleeman J, Agarwal D, et al. Altered metastatic behavior of human breast cancer cells after experimental manipulation of matrix metalloproteinase 8 gene expression. *Cancer Res* 2004;64:1687-94.
12. Agarwal D, Goodison S, Nicholson B, et al. Expression of matrix metalloproteinase 8 (MMP-8) and tyrosinase-related protein-1 (TYRP-1) correlates with the absence of metastasis in an isogenic human breast cancer model. *Differentiation* 2003;71:114-25.
13. Duffy MJ, Blaser J, Duggan C, et al. Assay of matrix metalloproteinases types 8 and 9 by ELISA in human breast cancer. *Br J Cancer* 1995;71:1025-8.
14. Balbin M, Fueyo A, Tester AM, et al. Loss of collagenase-2 confers increased skin tumor susceptibility to male mice. *Nat Genet* 2003;35:252-7.
15. Wang H, Parry S, Macones G, et al. Functionally significant SNP MMP8 promoter haplotypes and pre-term premature rupture of membranes (PPROM). *Hum Mol Genet* 2004;13:2659-69.
16. Gao YT, Shu XO, Dai Q, et al. Association of menstrual and reproductive factors with breast cancer risk: results from the Shanghai Breast Cancer Study. *Int J Cancer* 2000;87:295-300.
17. Shu XO, Gao YT, Cai Q, et al. Genetic polymorphisms in the TGF- β 1 gene and breast cancer survival: a report from the Shanghai Breast Cancer Study. *Cancer Res* 2004;64:836-9.
18. Long JR, Kataoka N, Shu XO, et al. Genetic polymorphisms of the CYP19A1 gene and breast cancer survival. *Cancer Epidemiol Biomarkers Prev* 2006;15:2115-22.
19. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;16:1215-23.
20. Tregouet DA, Escolano S, Tiret L, et al. A new algorithm for haplotype-based association analysis: the stochastic-EM algorithm. *Ann Hum Genet* 2004;68:165-77.
21. Rack B, Janni W, Gerber B, et al. Patients with recurrent breast cancer: does the primary axillary lymph node status predict more aggressive tumor progression? *Breast Cancer Res Treat* 2003;82:83-92.
22. Rutter JL, Mitchell TI, Buttice G, et al. A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter creates an Ets binding site and augments transcription. *Cancer Res* 1998;58:5321-5.
23. Kanamori Y, Matsushima M, Minaguchi T, et al. Correlation between expression of the matrix metalloproteinase-1 gene in ovarian cancers and an insertion/deletion polymorphism in its promoter region. *Cancer Res* 1999;59:4225-7.
24. Ye S, Dhillon S, Turner SJ, et al. Invasiveness of cutaneous malignant melanoma is influenced by matrix metalloproteinase 1 gene polymorphism. *Cancer Res* 2001;61:1296-8.
25. Przybylowska K, Kluczna A, Zadrozny M, et al. Polymorphisms of the promoter regions of matrix metalloproteinases genes MMP-1 and MMP-9 in breast cancer. *Breast Cancer Res Treat* 2006;95:65-72.
26. Wenham RM, Calingaert B, Ali S, et al. Matrix metalloproteinase-1 gene promoter polymorphism and risk of ovarian cancer. *J Soc Gynecol Investig* 2003;10:381-7.
27. Zhu Y, Spitz MR, Lei L, et al. A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter enhances lung cancer susceptibility. *Cancer Res* 2001;61:7825-9.
28. Su L, Zhou W, Park S, et al. Matrix metalloproteinase-1 promoter polymorphism and lung cancer risk. *Cancer Epidemiol Biomarkers Prev* 2005;14:567-70.
29. Ghilardi G, Biondi ML, Mangoni J, et al. Matrix metalloproteinase-1 promoter polymorphism 1G/2G is correlated with colorectal cancer invasiveness. *Clin Cancer Res* 2001;7:2344-6.
30. Hinoda Y, Okayama N, Takano N, et al. Association of functional polymorphisms of matrix metalloproteinase (MMP)-1 and MMP-3 genes with colorectal cancer. *Int J Cancer* 2002;102:526-9.
31. Zinzindohoue F, Lecomte T, Ferraz JM, et al. Prognostic significance of MMP-1 and MMP-3 functional promoter polymorphisms in colorectal cancer. *Clin Cancer Res* 2005;11:594-9.