

# Matrix Metalloproteinase 1 (*MMP1*) Is Associated with Early-Onset Lung Cancer

Wiebke Sauter,<sup>1,9</sup> Albert Rosenberger,<sup>2</sup> Lars Beckmann,<sup>3</sup> Silke Kropp,<sup>3</sup> Kirstin Mittelstrass,<sup>1</sup> Maria Timofeeva,<sup>4</sup> Gabi Wölke,<sup>1</sup> Angelika Steinwachs,<sup>1</sup> Daniela Scheiner,<sup>1</sup> Eckart Meese,<sup>5</sup> Gerhard Sybrecht,<sup>6</sup> Florian Kronenberg,<sup>7</sup> Hendrik Dienemann,<sup>8</sup> The LUCY-Consortium, Jenny Chang-Claude,<sup>3</sup> Thomas Illig,<sup>1</sup> Heinz-Erich Wichmann,<sup>1,9</sup> Heike Bickeböllner,<sup>2</sup> and Angela Risch<sup>4</sup>

<sup>1</sup>Institute of Epidemiology, GSF-National Research Center for Environment and Health, Neuherberg, Germany; <sup>2</sup>Department of Genetic Epidemiology, Georg-August University of Göttingen, Medical School, Göttingen, Germany; <sup>3</sup>Division of Clinical Epidemiology, German Cancer Research Centre (DKFZ), Heidelberg, Germany; <sup>4</sup>Division of Toxicology and Cancer Risk Factors, German Cancer Research Centre (DKFZ), Heidelberg, Germany; <sup>5</sup>Institute of Human Genetics, University of the Saarland, Saarbrücken, Germany; <sup>6</sup>Institute of Internal Medicine V, Medical School, University of the Saarland, Saarbrücken, Germany; <sup>7</sup>Division of Genetic Epidemiology, Department of Medical Genetics, Molecular, and Clinical Pharmacology, Innsbruck Medical University, Innsbruck, Austria; <sup>8</sup>Thoraxklinik, Heidelberg, Germany; <sup>9</sup>Department of Medical Informatics, Biometrics, and Epidemiology (IBE), University of Munich, Munich, Germany

## Abstract

Matrix metalloproteinases (MMP) play a key role in the breakdown of extracellular matrix and in inflammatory processes. MMP1 is the most highly expressed interstitial collagenase degrading fibrillar collagens. Over-expression of MMP1 has been shown in tumor tissues and has been suggested to be associated with tumor invasion and metastasis. Nine haplotype tagging and additional two intronic single nucleotide polymorphisms (SNP) of *MMP1* were genotyped in a case control sample, consisting of 635 lung cancer cases with onset of disease below 51 years of age and 1,300 age- and sex-matched cancer-free controls. Two regions of linkage disequilibrium (LD) of *MMP1* could be observed: a region of low LD comprising the 5' region including the promoter and a region of high LD starting from exon

1 to the end of the gene and including the 3' flanking region. Several SNPs were identified to be individually significantly associated with risk of early-onset lung cancer. The most significant effect was seen for rs1938901 ( $P = 0.0089$ ), rs193008 ( $P = 0.0108$ ), and rs996999 ( $P = 0.0459$ ). For rs996999, significance vanished after correction for multiple testing. For each of these SNPs, the major allele was associated with an increase in risk with an odds ratio between 1.2 and 1.3 (95% confidence interval, 1.0-1.5). The haplotype analysis supported these findings, especially for subgroups with high smoking intensity. In summary, we identified *MMP1* to be associated with an increased risk for lung cancer, which was modified by smoking. (Cancer Epidemiol Biomarkers Prev 2008;17(5):1127-35)

## Introduction

Lung cancer is the most common cancer with 1,350,000 new cases per year worldwide. Lung cancer mostly occurs in individuals of ages  $\geq 60$  years and constitutes a highly lethal disease with an average 5-year survival rate of 10% (1). The most important risk factor for lung cancer is smoking, with evidence of a strong dose-response relationship between smoking and lung cancer risk (2). In tobacco smoke, there are  $\sim 7,000$  substances including a multitude of carcinogens that induce a variety of DNA damages. Although smoking is considered to be the

predominant risk factor, only  $\sim 10\%$  of heavy smokers develop lung cancer. This suggests that genetic variation in sensitivity to carcinogen exposure may play an important role in the etiology of lung cancer (3).

Family aggregation and increased familial risk for lung cancer have been reported in several studies, providing indirect evidence that genetic factors contribute to susceptibility to lung cancer (4-6). Bailey-Wilson et al. (7) reported a region on chromosome 6q23-25 linked to familial lung cancer. The risk significantly increases with having an affected relative and further with earlier age of onset of the disease (4, 8-10).

Candidate susceptibility genes for lung cancer have been extensively studied, with most of the work focusing on mechanistically plausible variants in genes coding for enzymes involved in the activation and detoxification of carcinogens and repair of damage caused by tobacco smoke. Alterations in these pathways are hypothesized to affect an individual's processing of tobacco carcinogens and, therefore, the risk of developing lung cancer. Inflammation has been thought to play a role in carcinogenesis of a variety of cancers, including lung cancer, and may be due to infectious agents or other

Received 11/30/07; revised 2/9/08; accepted 2/26/08.

**Grant support:** National Genome Research Network and the Helmholtzgemeinschaft. Genotyping was done at the Genome Analysis Center of the GSF Research Center for Environment and Health. The KORA Surveys were financed by the GSF, which is funded by the German Federal Ministry of Education, Science, Research, and Technology and the State of Bavaria.

**Note:** Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

**Requests for reprints:** Wiebke Sauter, Institute of Epidemiology, GSF-National Research Center for Environment and Health, D-85764 Neuherberg, Germany. Phone: 49-89-3187-3527; Fax: 49-89-3187-4567. E-mail: wiebke.sauter@gsgf.de

Copyright © 2008 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-07-2840

environmental exposures such as tobacco smoke (11). Genes involved in the regulation of the inflammatory response were recently added to the lung cancer candidate genes in literature, and some of these focused genes belong to the family of matrix metalloproteinase (*MMP*).

*MMPs* comprise a structurally and functionally related family of zinc metalloproteinases degrading extracellular matrix and basement membrane barriers, and thus are thought to play a key role in angiogenesis, inflammatory processes, cancer development, cell proliferation, and apoptosis (12). The activity of *MMPs* is regulated at the level of transcription, activation, and inhibition by tissue inhibitors of metalloproteinases (*TIMP*; ref. 13). In this context, *MMPs* have been focused on as targets for therapeutic strategies. Because of their role in the degradation of the extracellular matrix leading to tumor invasion and metastasis, they may also serve as prognostic markers (14, 15).

Members of the *MMP* family have different substrate specificities and expression patterns. Among the *MMPs*, *MMP1* is the most highly expressed interstitial collagenase degrading fibrillar collagens, which are major constituents of the extracellular matrix.

*MMP1* is up-regulated in a wide variety of advanced cancers and has been suggested to be associated with tumor invasion and metastasis (16). A significant negative correlation between its expression and cancer survival has been found (17). Overexpression of *MMP1* has been reported in lung cancer cells (18, 19) and several association studies between DNA variants of *MMP1* and lung cancer have previously been conducted (20-22).

The level of *MMP1* expression can be influenced by single nucleotide polymorphisms (SNP), especially, but not exclusively, if they are located within the promoter region of the *MMP1* gene.

The aim of this case-control study of early-onset lung cancer was a detailed investigation of the genetic variability within the *MMP1* gene in relation to early-onset lung cancer using individual SNP and haplotype analysis and considering potential effect modifications by the amount of smoke exposure.

## Materials and Methods

**Study Population.** The present case-control study in Caucasians included 635 primary lung cancer patients being diagnosed before the age of 51 y and 1,300 cancer-free individuals matched for by sex and age. The control sample was collected within the population-based KORA study. Cases were selected either from the LUCY study or the Heidelberg lung cancer study.

The LUCY study (LUng Cancer in the Young) is a multicenter study with 31 participating clinics all over Germany. Only newly diagnosed patients with histologically or cytologically confirmed primary lung cancer entered the study. Detailed epidemiologic data on family history, tobacco and smoking exposure, education, and occupational exposure have been collected and blood samples were taken.

The Heidelberg lung cancer study is an ongoing hospital-based case-control study (23). The German Cancer Research Center (DKFZ) has recruited more than 1,000 lung cancer cases at and in collaboration with the

Thoraxklinik Heidelberg. Thereof, 163 lung cancer cases with onset of disease before age 51 y, recruited between 01/1997 and 12/2003, were included in the analysis. Data on occupational exposure, tobacco smoking, and educational status, and for a subgroup also data on family history of lung cancer, were assessed with a self-administered questionnaire.

The KORA study (Cooperative Health Research in the Augsburg Region) is a population-based epidemiologic survey of individuals living in or near the city of Augsburg, Southern Germany (24, 25). With the genotyped polymorphisms to date, a major population stratification between KORA (Southwest Germany) and two other cohorts from Northern Germany could not be detected in a genomic control approach (26).

Informed consent was obtained from all study participants and the studies were approved by the ethics committee of the Bayerische Landesärztekammer, the corresponding local ethics committees of the participating clinics, and the ethics committee of the University of Heidelberg.

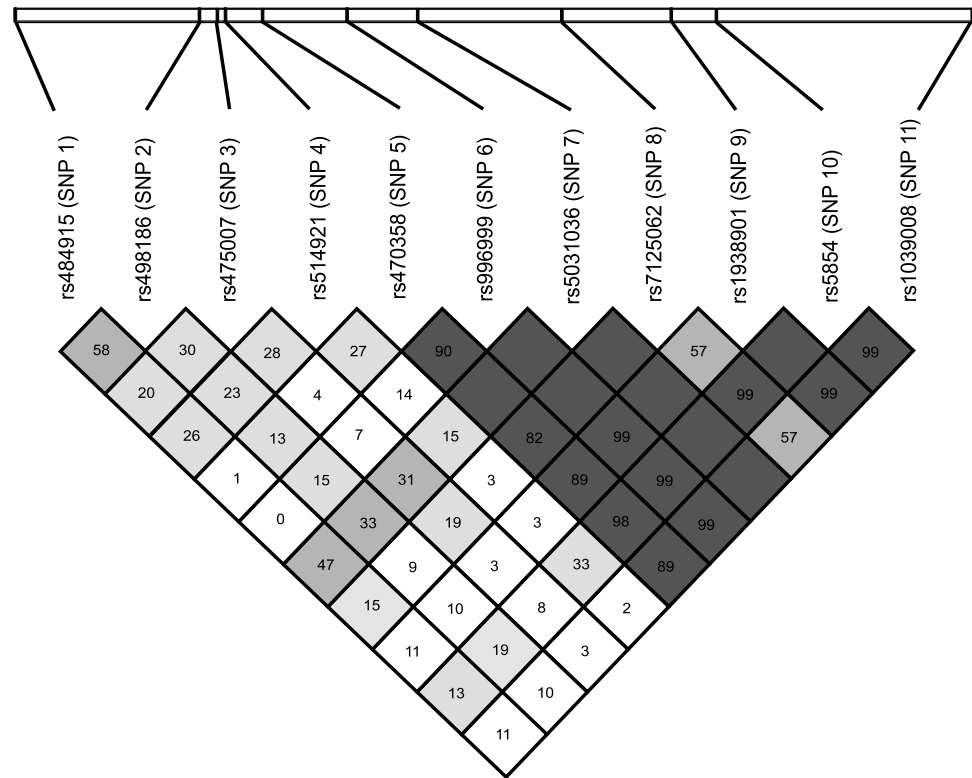
Concerning tobacco exposure, cases and controls were classified as never smokers, former smokers if they quit for at least 1 y at the time of diagnosis/interview, and current smokers. Individuals who smoked <1 pack-year were considered as never smokers (Table 1). Smoking status was quantified in pack-years and subjects were grouped into four categories: very light smokers (0-10 pack-years), light smokers (11-20 pack-years), moderate smokers (21-30 pack-years), and heavy

**Table 1. Demographic characteristics of cases and controls**

Characteristic	Cases (n = 635)	Controls (n = 1,300)
Age (y), mean ± SD		
Males	45.4 ± 4.1	45.2 ± 4.3
Females	44.7 ± 4.6	44.7 ± 4.7
Gender, n (%)		
Male	406 (64)	819 (63)
Female	229 (36)	481 (37)
Smoking status, n (%)		
Male		
Never smoker	12 (3)	243 (30)
Former smoker	69 (17)	249 (30)
Current smoker	322 (79)	325 (40)
Unknown	3 (1)	1 (0)
Female		
Never smoker	29 (13)	207 (43)
Former smoker	41 (18)	128 (27)
Current smoker	158 (69)	147 (30)
Unknown	1 (0)	—
Smoking categories (male + female), n (%)		
Weak smokers (<20 pack-years)	168 (26)	885 (68)
Strong smokers (≥20 pack-years)	452 (71)	385 (30)
Unknown	15 (2)	30 (2)
Histology, n (%)		
AC	222 (35)	—
SCLC	153 (24)	—
Other NSCLC	218 (34)	—
Other LC	42 (7)	—

NOTE: Categorical and quantitative variables are reported in absolute and relative frequencies. Metric variables are reported as mean ± SD. Other NSCLC, other non-small-cell lung cancers, excluding adenocarcinoma.

Abbreviations: AC, adenocarcinoma; SCLC, small-cell lung cancer; NSCLC, non-small-cell lung cancer; LC, lung cancer.



**Figure 1.** LD plot. Coefficients ( $D'$ ) of pairwise linkage disequilibrium between the 11 polymorphisms in the *MMP1* gene in the control population.

smokers ( $\geq 31$  pack-years). The subjects were also grouped into two sets defined by weak smokers (very light and light) and strong smokers (moderate and intensive).

Histologic subtypes of cancer were grouped into small-cell carcinoma, adenocarcinoma, other non-small-cell lung carcinoma, and other types of lung cancer (Table 1).

**Genotyping Methods and SNP Selection.** DNA was extracted from EDTA anticoagulated blood using a commercially available kit (Gentra) according to the manufacturer's protocol. For the Heidelberg lung cancer study, buffy coats from 5 mL of venous blood in EDTA were stored at  $-80^{\circ}\text{C}$ , and human genomic DNA was isolated with the QIAamp DNA blood midi kit according to the manufacturer's instructions (Qiagen GmbH).

Based on the data of HapMap phase II (release #20; ref. 27), a genomic region of 18.239 kb on chromosome 11q22-23 containing the entire *MMP1* gene plus 5,000 bp each upstream (3' untranslated region) and downstream (5' untranslated region) was selected. Nine haplotype tagging SNPs were identified using the following tagging criteria: pairwise tagging of the HapMap CEU Population with  $r^2 \geq 0.8$  and a minor allele frequency  $\geq 20\%$ . To increase SNP density, two intronic SNPs (rs5031036 and rs193890) were selected from the literature and public databases (Fig. 1).

SNPs were genotyped using iPLEX single base primer extension and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (Sequenom) as recently described (28). Genotyping was done by laboratory personnel blinded to case-control status. Standard genotyping quality control included 10%

duplicate samples and negative samples. Hardy-Weinberg equilibrium was tested in cases and controls separately using a log-likelihood ratio test. Pairwise linkage disequilibrium (LD) between the 11 SNPs in *MMP1* was measured by  $D'$  and  $r^2$ , and was calculated based on the genotypes of 1,300 control subjects.

**Statistical Analysis.** Single SNP association tests were calculated using conditional logistic regression. In the general model, two binary dummy variables, indicating heterozygous and homozygous predisposing alleles, were included into the logistic regression. Three models, dominant, recessive, and multiplicative, indicating the increase of risk for one additional predisposing allele, were fitted as well.

Cases and controls were frequency matched for gender and age at diagnosis or interview ( $\leq 40$ , 41-45, and 46-50 y). All analyses were done conditional on these age-gender strata and adjusted for smoking. We also fitted unconditional logistic models with age and sex as covariates and found the results almost identical to those of the conditional models (data not shown). These models were also carried out for cases of specific histologic subtypes. We carried out a Monte Carlo permutation, often considered as the gold standard for analysis (29), with 1,500 replications assigning the case-control status by permutation to evaluate whether single SNP results could be accountable to multiple testing while considering the underlying LD structure. Adjusted  $P$  values ( $P_{\text{adjust}}$ ) are derived as the  $P$ -quantiles of the empirical distribution of the 1,500 permutations where  $P$  is the unadjusted  $P$  value (30). The SNPs, chosen for this study, were carefully selected tagging SNPs. Thus, we applied two statistical methods using haplotype

information. First, we applied Mantel statistics based on haplotype sharing (31). In the vicinity of a predisposing mutation, haplotypes carrying a disease-relevant mutation (case haplotypes) should be more related than haplotypes not carrying such a mutation (random haplotypes). Therefore, case haplotypes are expected to share longer stretches of conserved DNA identical by descent around this mutation than random haplotypes. The pointwise test statistic correlates genetic similarity and phenotypic similarity across pairs of haplotypes to assess associations between a genetic marker  $x$  and phenotypes. The genetic similarity at  $x$  is measured as the shared length between haplotypes around the genetic marker as the number of markers surrounding  $x$  with the same alleles (i.e., identical by state). The phenotypic similarity for two haplotype copies is defined as the mean-corrected product, where the mean is estimated from the sample. Statistical significance was derived by Monte Carlo permutation with 10,000 replications. A step-down min- $P$  algorithm was applied to control the family-wise error rate. The algorithm is based on the permutations and thus takes into account the underlying LD structure (32).

The second haplotype-based approach was to test for differences in haplotype frequencies between cases and controls using the software Chaplin<sup>9</sup> (33, 34). Haplotype frequencies were derived by an extension of the standard Expectation-Maximization algorithm, called the Expectation-Conditional-Maximization algorithm. The approach allows for violations of Hardy-Weinberg equilibrium of haplotype frequencies in control and case samples that are typically required otherwise. Chaplin provides estimates of haplotype effects on disease and also tests whether such effects are significantly different from zero. We analyzed the multiplicative model of single haplotype effects and a general model. For both the haplotype frequency and the haplotype sharing analysis, the SNPs were separated into two LD sets. Due to the analysis of pairwise LD, the first set included SNP1 to SNP4 whereas the second set included SNP6 to SNP11 (Fig. 1). SNP5 was not included because of deviation from Hardy-Weinberg equilibrium in controls.

Single point and haplotype analyses were done for the whole data set and within groups defined by smoking status. The level of significance was set at  $\alpha = 5\%$  for each single test. All three statistical approaches used (i.e., single point analysis with conditional logistic regression, the Mantel statistics using haplotype sharing, and the haplotype frequency association test) account for multiple testing considering the correlation between the SNPs. The methods also provide  $P$  values for the null hypothesis that none of the SNPs is associated with the outcome (i.e., the family wise error rate). Pairwise LD was estimated using the software HaploView.<sup>10</sup> Single point association analysis was done using SAS 9.1. For the haplotype frequency association analysis, we used

Chaplin version 2.1. The software package Tomcat<sup>11</sup> was used for Mantel statistics using haplotype sharing. Haplotypes were estimated with the software fast-PHASE<sup>12</sup> (35).

## Results

### Subject Characteristics and Genotype Distribution.

There were no statistically significant differences in the distribution of sex and age between patients and controls. However, significantly more smokers were present among patients compared with controls. The most frequent histologic type of the neoplasm was adenocarcinoma (35%) and other non-small-cell lung carcinoma (34%; Table 1). Adenocarcinoma was diagnosed more frequently in women (47%) and other non-small-cell lung carcinoma was diagnosed more frequently in men (41%). Small cell carcinoma was diagnosed almost equally frequently in men and women (~25%; data not shown).

The genotyping success rates ranged from 98% to 100%. For 12 subjects, <9 of the 11 SNPs were genotyped successfully; the genotype distribution of SNP5 (rs470358) showed a slight deviation from Hardy-Weinberg equilibrium ( $P = 0.0229$ ) in controls, and this SNP was therefore excluded from haplotype analyses. Most polymorphisms chosen represent tagging SNPs and the minor allele frequencies ranged from 19% to 45%. The two additional intronic SNPs without tagging informations are the intronic SNP7 (rs5031036) and SNP9 (rs1938901) where the minor allele frequency in controls was 5% and 31%, respectively (Supplementary Table S3).

**Descriptive Analysis of the LD Structure of the Observed Region.** We investigated the LD structure within controls around *MMP1* covered by the genotyped SNPs (Fig. 1). The region was found to consist of two parts: a subregion of almost no LD covering the 5' flanking region [between SNP1 (rs484915) and SNP4 (rs51492)], with pairwise  $D' \leq 0.6$  and  $r^2 \leq 0.21$ , and a block of high LD between SNP5 (rs470358) in intron 1 and SNP11 (rs1939008) in the 3' untranslated region. All pairwise LDs exceeded  $D' \geq 0.99$  and  $r^2 \geq 0.22$  except SNP7 (rs5031036). Between the adjacent SNP8 (rs7125062) and SNP9 (rs1938901),  $D'$  was 0.57 ( $r^2 = 0.31$ ). The most pronounced multilocus LD was found in the 3' flanking region containing SNP9, SNP10, and SNP11 ( $\epsilon' = 0.62$ ).

**Tests for Association with Lung Cancer.** Table 2 presents the main results for single SNP association analysis with lung cancer using conditional logistic regression analysis. In the general model, SNP6 (rs996999), SNP9 (rs1938901), and SNP11 (rs193008) showed significant associations. The last two remained significant even after adjusting for multiple testing.

No significant association was observed for the four SNPs of the 5' flanking region (Supplementary Table S1). The lowest  $P$  value [ $P = 0.0027$ ; odds ratio (OR), 1.3; 95% confidence interval (95% CI), 1.1-1.5] was obtained for rs1938901 (SNP9, 7229 C>T, intron 8) assuming a multiplicative model of inheritance. For carriers of the C-allele (major allele), an increase in the risk for lung cancer was estimated with an OR of 1.5 (95% CI, 1.0-2.2;  $P = 0.0534$ ) for heterozygotes (CT) and an OR of 1.8 (95%

<sup>9</sup> <http://www.genetics.emory.edu/labs/epstein/software/chaplin>

<sup>10</sup> <http://www.broad.mit.edu/mpg/haploview>

<sup>11</sup> <http://www.dkfz.de/en/kepidemiologie/software/software.html>

<sup>12</sup> <http://stephenslab.uchicago.edu/software.html>

**Table 2. Conditional logistic regression analysis of lung cancer with SNPs in the *MMP1* gene**

Model	MAF (%)		All cases and controls		Strong smokers (≥20 pack-years)		Never and weak smokers (<20 pack-years)	
	Controls	Cases	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)
SNP6: rs996999 (1831 C>T)	<b>20</b>	<b>19</b>						
General			0.1654	1.5 (0.8-2.6)	0.195	1.6 (0.8-3.1)	0.5724	1.3 (0.5-3.5)
			0.0332*	1.8 (1.0-3.1)	0.0498*	1.9 (1.0-3.7)	0.3492	1.6 (0.6-4.0)
F-statistics			0.0459*	(P <sub>adjust</sub> = 0.1182)				
Dominant			0.0542	1.7 (1.0-2.9)	0.0751	1.8 (0.9-3.4)	0.4049	1.5 (0.6-3.7)
Recessive			0.0376*	1.3 (1.0-1.6)	0.0702	1.3 (1.0-1.7)	0.2883	1.2 (0.8-1.8)
Multiplicative	0.0152*	1.3 (1.0-1.5)	0.0314*	1.3 (1.0-1.6)	0.236	1.2 (0.9-1.7)		
SNP9: rs1938901 (7229 C>T)	<b>31</b>	<b>28</b>						
General			0.0534	1.5 (1.0-2.2)	0.0238*	1.79 (1.1-3.0)	0.7895	1.1 (0.6-2.0)
			0.0033 <sup>†</sup>	1.8 (1.2-2.7)	0.0016 <sup>†</sup>	2.23 (1.4-3.7)	0.4347	1.3 (0.7-2.3)
F-statistics			0.0089 <sup>†</sup>	(P <sub>adjust</sub> = 0.01821)				
Dominant			0.0096 <sup>†</sup>	1.7 (1.1-2.4)	0.0039 <sup>†</sup>	2.02 (1.3-3.3)	0.5702	1.2 (0.7-2.1)
Recessive			0.0156*	1.3 (1.1-1.6)	0.0208*	1.39 (1.1-1.8)	0.3251	1.2 (0.8-1.7)
Multiplicative	0.0027 <sup>†</sup>	1.3 (1.1-1.5)	0.0023 <sup>†</sup>	1.39 (1.1-1.7)	0.3174	1.1 (0.9-1.5)		
SNP11: rs193008 (12471 G>A)	<b>31</b>	<b>28</b>						
General			0.0603	1.5 (1.0-2.2)	0.0409*	1.7 (1.0-2.8)	0.6357	1.2 (0.6-2.1)
			0.0041 <sup>†</sup>	1.8 (1.2-2.6)	0.0029 <sup>†</sup>	2.14 (1.3-3.5)	0.3613	1.3 (0.7-2.4)
F-statistics			0.0108*	(P <sub>adjust</sub> = 0.02334)				
Dominant			0.0117*	1.6 (1.1-2.4)	0.0075 <sup>†</sup>	1.93 (1.2-3.1)	0.4602	1.6 (0.7-2.2)
Recessive			0.0173*	1.3 (1.0-1.6)	0.0208*	1.39 (1.1-1.8)	0.3554	1.2 (0.8-1.7)
Multiplicative	0.0032 <sup>†</sup>	1 (1.1-1.5)	0.0031 <sup>†</sup>	1.38 (1.1-1.7)	0.3011	1.2 (0.9-1.5)		

NOTE: Data are matched for age and gender. All ORs are adjusted for smoking.  $P_{\text{adjust}}$ ,  $P$  values adjusted for multiple testing. SNP position is counted in basepairs to/from start (AUG); risk allele is in boldface.  $F$ -statistics was used for the full model versus the model without any genetic effect. The table shows only the polymorphisms with significant association results; the entire results can be seen in Supplementary Table S3.

Abbreviation: MAF, minor allele frequency.

\* $P < 0.05$ .

<sup>†</sup> $P < 0.01$ .

CI, 1.2-2.7;  $P = 0.0033$ ) for homozygotes (CC). The nearby located polymorphism rs193008 (SNP11, 12471 G>A) showed an almost identically increased lung cancer risk for carriers of the G allele. No association ( $P = 0.9181$ ) was found for the intermediate polymorphism rs5854 (SNP10, 8020 G>A; Supplementary Table S1). A significant association with higher risk of lung cancer for carriers of the C allele was also observed for rs996999 in intron 4 (SNP6, 1831 C>T) with  $P = 0.0152$  and an OR of 1.3 (95% CI, 1.0-1.5) for each allele. For the same three SNPs, we further observed a stronger association with risk of lung cancer for those who smoked ≥20 pack-years whereas we

observed no association for those who smoked <20 pack-years (Table 2).

The associations of polymorphisms with respect to histologic subtypes of lung cancer are shown in Table 3. The estimated effects for the polymorphisms rs996999 (SNP6), rs1938901 (SNP9), and rs193008 (SNP11) with respect to adenocarcinoma were found to be larger than those for the other non-small-cell lung carcinoma and small-cell lung carcinoma. The lowest  $P$  value was found for rs193008 (SNP11), assuming an additive mode of inheritance. For heterozygotes, we estimated an OR of 1.8 (95% CI, 1.0-3.2;  $P = 0.0708$ ), and for homozygotes, an OR

**Table 3. Association of polymorphisms and different histologic subtypes of lung cancer for the general model**

Polymorphism	Risk allele	MAF (%)		Histologic subtype	One risk allele		Two risk alleles	
		Controls	Cases		OR (95% CI)	P	OR (95% CI)	P
SNP6: rs996999	C	<b>20</b>	<b>19</b>	AC	2.3 (0.9-5.6)	0.0809	2.6 (1.1-6.3)	0.0366*
				SCLC	1.9 (0.6-5.6)	0.2619	2.5 (0.9-7.4)	0.0852
				Other NSCLC	1.3 (0.6-2.8)	0.5157	1.4 (0.7-3.0)	0.3353
SNP9: rs1938901	C	<b>31</b>	<b>28</b>	AC	1.8 (1.0-3.3)	0.0607	2.2 (1.2-4.0)	0.0101*
				SCLC	1.5 (0.7-2.9)	0.2658	1.8 (0.9-3.4)	0.096
				Other NSCLC	1.6 (0.9-2.9)	0.1403	1.9 (1.0-3.4)	0.0367*
SNP11: rs193008	G	<b>31</b>	<b>28</b>	AC	1.8 (1.0-3.2)	0.0708	2.2 (1.2-4.0)	0.0100*
				SCLC	1.6 (0.8-3.2)	0.2033	1.9 (0.9-3.7)	0.0749
				Other NSCLC	1.5 (0.8-2.6)	0.2051	1.7 (1.0-3.0)	0.0716

NOTE: The table shows only the polymorphisms with significant association results. The entire results are shown in Supplementary Table S4.

\*SNP with significant association to histologic subtype.

**Table 4. Results from the Mantel statistics using haplotype sharing**

	<i>P</i> values for single SNPs						Global <i>P</i>
	SNP6	SNP7	SNP8	SNP9	SNP10	SNP11	
All cases and controls	0.222	0.222	0.222	0.146	0.176	0.222	0.146
Never and weak smoker (<20 pack-years)	0.888	0.868	0.888	0.939	0.904	0.921	0.868
Strong smoker (≥20 pack-years)	0.022*	0.022*	0.011*	0.007 <sup>†</sup>	0.010*	0.011*	0.007 <sup>†</sup>

NOTE: *P* values were derived by 10,000 Monte Carlo permutations. Global *P* value is derived by a step-down min-*P* algorithm to adjust for multiple testing.

\**P* < 0.05.

<sup>†</sup>*P* < 0.01.

of 2.2 (95% CI, 1.2-4.0; *P* = 0.01). For other non-small-cell lung carcinoma and small-cell lung carcinoma, significance was not achieved (*P* ~ 0.07). Due to the relative small number of small-cell lung carcinoma cases (*n* = 153; Table 1), the power for detecting association with small-cell lung carcinoma was rather low.

Table 4 presents the results for the Mantel statistics using haplotype sharing for the region of high LD (SNP6 to SNP11). No significant results were found in the data set consisting of all subjects and the subset of never and weak smokers (<20 pack-years). In the subset of strong smokers, however, significant results were found across the entire region. The lowest *P* value was found for SNP9 (*P* < 0.003). After correction for multiple testing, the entire region of high LD was found to be significantly associated with the risk of lung cancer (global *P* value *P* = 0.007). A haplotype sharing analysis of the 5' flanking region (SNP1-SNP4) did not yield significant results (data not shown).

We used the software Chaplin to test for differences in haplotype frequencies between cases and controls. Statistical tests were calculated in two separated SNP sets: first for SNP1 to SNP4, and second for SNP6 to SNP11. Testing for a multiplicative disease model, we did not find significant results for single haplotype effects in the respective data sets. However, for the haplotypes consisting of SNP6 to SNP11, globally significant results were found for strong smokers with *P* = 0.049 (Table 5). Four of the six estimated haplotypes showed positive effect estimates (i.e., indicating an increase in risk). Interestingly, the four haplotypes were characterized by the major alleles of SNP6, SNP9, and SNP11, which showed significant results in the single

point analysis. No significant results were found for the 5' flanking region (SNP1-SNP4).

## Discussion

In the study at hand, we investigated the association between genotypes and haplotypes of nine tagging SNPs and two intronic polymorphisms of the *MMP1* gene and the risk of early-onset lung cancer. Three polymorphisms were identified to be significantly associated with the risk of early-onset lung cancer development, whereas the common allele was found to represent the respective risk allele.

The human *MMP1* consists of 10 exons and is located on human chromosome 11q22.2-22.3. This gene is tightly linked to a cluster of eight *MMP* genes, including *MMP3*, *MMP7*, *MMP8*, *MMP10*, *MMP12*, *MMP13*, *MMP20*, and *MMP27*, and two pseudogenes (36). To date, most genetic epidemiologic studies on *MMP1* gene variation in relation to cancers have focused on rs1799750 (-1607 1G>2G) with conflicting results (20-22, 37-40). Zhu et al. (22) found an overall association of the 2G allele with elevated risk of lung cancer in Caucasians, which was not confirmed either by Su et al. (21) in a study of Caucasian patients or by Fang et al. (20) in a Chinese population. Remarkable is that several studies reported deviation from Hardy-Weinberg equilibrium in rs1799750 (20, 22, 37, 40).

Functional analyses have shown that the expression level of *MMP1* depends on the genetic variation within the promoter of the *MMP1* gene. This was mainly observed for the above-mentioned single base deletion at

**Table 5. Summary of haplotype-specific effects and global test for a multiplicative effect in the group of strong smokers**

Haplotype number	SNP6	SNP7	SNP8	SNP9	SNP10	SNP11	Exp( $\beta$ )	<i>P</i>
<i>MMP1</i> _01	0	0	0	0	1	0	0.94	0.952
<i>MMP1</i> _02	1	0	1	0	1	0	0.88	0.901
<i>MMP1</i> _03	1	0	1	1	0	1	1.16	0.880
<i>MMP1</i> _04	1	0	0	1	1	1	1.12	0.911
<i>MMP1</i> _05	1	0	1	1	1	1	1.49	0.692
<i>MMP1</i> _06	1	1	1	1	1	1	1.58	0.651
Allele 1	C	A	T	C	C	G		
Allele 0	T	G	C	T	T	A		
Robust score test								0.049*

NOTE: SNP coded as 0 for minor allele and 1 for major allele. Exp( $\beta$ ), haplotype effect estimate (positive values are indicating a risk increasing effect). Boldfaced data indicate risk increasing effect in single SNP association analyses.

\**P* < 0.05.

position  $-1607\ 1G>2G$  (rs1799750), where the nondeleted alleles (2G) are associated with higher risk for lung cancer and a higher transcription and protein expression (41). The  $-1607\ 2G$  allele is thought to form the core of a consensus DNA element recognized by the Ets transcription factor, which up-regulates *MMP1* transcription (42). Further investigations of association of  $-1607\ 1G>2G$  with allelic expression imbalance suggest that this polymorphism does not account for all differences in allelic expression observed (43). Transcription of a gene is more likely to be influenced by multiple polymorphisms, and these are hypothesized by some authors to be located in the promoter of that gene, which acts in concert to exert a haplotype effect (44, 45). Pearce et al. (46) investigated the promoter region in detail and found that the  $-1607\ 1G>2G$  deletion alone cannot fully segregate the various *MMP1* haplotypes that differ in promoter activity.

Although we did not analyze the "functional" *MMP1*  $-1607\ 1G>2G$  SNP, we observed that this polymorphism was not located in the associated region and not in the same LD block. Therefore, the modified expression levels may not be due to this polymorphism alone. These results suggest that an approach looking at the whole common genetic variation of the gene instead of an individual SNP-based approach should be used. Using such an approach, Sun et al. (47) recently found a lung cancer-associated haplotype structure in the *MMP1* gene in Chinese cases. The locus of the associated haplotypes is comparable with the results of the study at hand, suggesting a possible disease locus in the region from intron 6 to the 3' untranslated region.

MMPs are secreted zinc metalloproteinases that degrade collagens of the extracellular matrix and are important in tissue remodeling and repair during development and inflammation. In normal adult tissues, the levels of MMPs are usually low. However, their expression is elevated when the system faces disturbances, such as wound healing, repair, or remodeling processes, which occur in several pathologic conditions (15). There are several factors known to regulate, control, or affect the expression of *MMP1* in the tissue. The expression is transcriptionally regulated by growth factors, hormones, and cytokines (48-50) and the proteolytic activity is accurately controlled through activators and inhibitors of metalloproteinases. Additionally, environmental factors affect *MMP1* expression: Cigarette smoke induces mRNA and protein expression of *MMP1* in pulmonary epithelial cells and skin fibroblasts, whereas the expression of the *MMP1*-regulating tissue inhibitors (*TIMP1*) as well as the house keeping enzyme glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) is not affected (51-56). As a result, the collagenolytic activity of human airway cells increases (54).

High expression levels of *MMP1* are responsible for loosening cell adhesion. Carcinogenic substances easily penetrate the epidermis and infiltrate and damage cells. If cancer cells are generated, metastasis cells can spread more easily in the surrounding area (57, 58). Supporting the known effect of tobacco smoke on *MMP1* expression, we could observe a stronger association with risk of early-onset lung cancer for those who smoked  $\geq 20$  pack-years compared with those who smoked  $< 20$  pack-years.

Summarizing, it can be stated that the effect of *MMP* polymorphisms depends on the balance between MMPs and TIMPs, which is strongly influenced by tobacco exposure and hormone status and therefore seems to be a highly complex system. Recapitulating our results, we predict a disease locus either in the 3' flanking region or near the end of the gene (from intron 6 to the gene's end) especially increasing the risk for early-onset lung cancer in strong smokers. Due to the fact that the 3' untranslated region and the flanking sequences are responsible for stabilization of mRNA, it could be hypothesized that the major alleles causes a stable mRNA with normal protein expression. Exposure to tobacco smoke increases the *MMP1* level, resulting in increased collagenase activity, decreased cell adhesion, and easier penetration of particles and carcinogens. Carriers of the protective alleles may benefit from a less stable mRNA, developing lower *MMP1* levels and lower collagenase activity, even if they are exposed to tobacco smoke.

Lung cancer at young age is rare ( $< 10\%$  of all lung cancer cases), and this study is based on one of the largest groups of young patients worldwide. For the investigation of genetic associations, these are of specific value because they show a much stronger aggregation with familial lung cancer and other cancer entities in comparison with lung cancer cases at a higher age (9). Our study was population based and we had sufficient SNP coverage across the *MMP1* gene. It might be considered as a limitation that our study did not include the one frequently discussed promoter SNP (rs1799750) and that the haplotype analysis of the 5' flanking region is hampered by low LD.

However, to our knowledge, this is the first investigation in Caucasians on association of SNPs and haplotypes on the entire *MMP1*-gene. In conclusion, individuals carrying the major allele of rs996999, rs1938901, or rs193008 and/or belonging to a haplotype carrying the same alleles may harbor a DNA variant that affects lung cancer risk, and this association was strongly modified by the amount of tobacco smoke exposure. This relationship between smoking, *MMP1* polymorphisms, and lung cancer risk needs to be confirmed by other independent studies.

## Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

## Acknowledgments

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.

We thank all the participating hospitals (LUCY-Consortium): Aurich (Dr. Heidi Kleen); Bad Berka (Dr. med. R. Bonnet, Klinik für Pneumologie, Zentralklinik Bad Berka GmbH); Bonn (Prof. Ko, Dr. Geisen, Innere Medizin I, Johanniter Krankenhaus); Bonn (Dr. Stier, Medizinische Poliklinik, Universität Bonn); Bremen (Prof. Dr. D. Ukena, Dr. Penzl, Zentralkrankenhaus Bremen Ost, Pneumologische Klinik); Chemnitz (PD Dr. Schmidt, OA Dr. Jielge, Klinikum Chemnitz, Abteilung Innere Medizin); Coswig (Prof. Höffken, Dr. Schmidt, Fachkrankenhaus Coswig); Dieckholzen (Dr. Hamm, Kreiskrankenhaus Dieckholzen, Klinik für Pneumologie); Donaustauf (Prof. Pfeifer, Dr. v. Bülzingslößen, Fr. Schneider, Fachklinik für Atemwegserkran-

kungen); Essen (Prof. Teschler, Dr. Fischer, Ruhrlandklinik-Universitätsklinik, Abt. Pneumologie); Gauting (Prof. Häußinger, Prof. Thetter, Dr. Düll, Dr. Wagner, Pneumologische Klinik München-Gauting); Gera (CA MR Dr. Heil, OÄ Dr. Täuscher, OA Dr. Lange, II. Medizinische Klinik, Wald-Klinikum Gera); Göttingen (Prof. Trümper, Prof. Griesinger, Dr. Overbeck, Abteilung Onkologie, Hämatologie); Göttingen (Prof. Schöndube, Dr. Danner, Abteilung Thorax-, Herz-, und Gefäßchirurgie); Göttingen/Weende (Dr. med. Fleischer, Ev. Krankenhaus Göttingen-Weende e.V., Abteilung Allgemeinchirurgie); Greifenstein (Prof. Morr, Dr. M. Degen, Dr. Matter, Pneumologische Klinik, Waldhof Elgershausen); Greifswald (Prof. Ewert, Dr. Altesellmeier, Universitätsklinik Greifswald, Klinik für Innere Medizin B); Hannover (Prof. Schönhofer, Dr. Kohlaußen, Klinikum Hannover Oststadt, Medizinische KlinikII, Pneumologie); Heidelberg (Prof. Drings, Dr. Herrmann, Thoraxklinik-Heidelberg GmbH, Abt. Innere Medizin-Onkologie); Hildesheim (Prof. Kaiser, St. Bernward Krankenhaus, Medizinische Klinik II); Homburg (Prof. Sybrecht, OA Dr. Gröschel, Dr. Mack, Uniklinik des Saarlandes, Innere Medizin V); Immenhausen (Prof. Andreas, Dr. Rittmeyer, Fachklinik für Lungenerkrankungen); Köln (Priv. Doz. Dr. Stölben, Kliniken der Stadt Köln, Lungenklinik Krankenhaus Merheim); Köln (Prof. Wolf, Dr. Staratschek-Jox, Klinikum der Universität Köln, Klinik I für Innere Medizin); Leipzig (Prof. Gillisen, OA Dr. Cebulla, Städt. Klinikum St. Georg, Robert-Koch-Klinik); Leipzig (Dr. Kreyborg, Universitätsklinikum Leipzig, Medizinische Klinik I, Abteilung Pneumologie); Lengler (Prof. Crieé, Dr. Körber, Dr. Knaack, Ev. Krankenhaus Weende e.V., Standort Lengler, Abt. Pneumologie); München (Prof. Huber, Dr. Borgmeier, Klinikum der LMU-Innenstadt, Abt. Pneumologie); Neustadt a. Harz (Dr. Keppler, Dr. Schäfer, Evangelisches Fachkrankenhaus für Atemwegserkrankungen); Rotenburg (Prof. Schaberg, Dr. Struß, Diakoniekrankenhaus Rotenburg, Lungenklinik Unterstedt); St. Pölten-Österreich (OA Dr. M. Wiesholzer, Zentralklinikum St. Pölten, I. Medizinische Klinik). We gratefully acknowledge the KORA study group especially G. Fischer, H. Gallert, N. Klopp, C. Gieger and R. Holle Institute of Epidemiology, GSF-National Research Center for Environment and Health, Neuherberg, Germany and all individuals who participated as cases or controls in this study and the KORA Study Center and their co-workers for organizing and conducting the data collection.

## References

- Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005;55:74–108.
- Doll R, Peto R. Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and lifelong non-smokers. *J Epidemiol Community Health* 1978;32:303–13.
- Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers—a different disease. *Nat Rev Cancer* 2007;7:778–90.
- Yang P, Schwartz AG, McAllister AE, Aston CE, Swanson GM. Genetic analysis of families with nonsmoking lung cancer probands. *Genet Epidemiol* 1997;14:181–97.
- Amos CI, Caporaso NE, Weston A. Host factors in lung cancer risk: a review of interdisciplinary studies. *Cancer Epidemiol Biomarkers Prev* 1992;1:505–13.
- Etzell CJ, Amos CI, Spitz MR. Risk for smoking-related cancer among relatives of lung cancer patients. *Cancer Res* 2003;63:8531–5.
- Bailey-Wilson JE, Amos CI, Pinney SM, et al. A major lung cancer susceptibility locus maps to chromosome 6q23–25. *Am J Hum Genet* 2004;75:460–74.
- Bailey-Wilson JE, Sellers TA, Elston RC, Evens CC, Rothschild H. Evidence for a major gene effect in early-onset lung cancer. *J La State Med Soc* 1993;145:157–62.
- Kreuzer M, Kreienbrock L, Gerken M, et al. Risk factors for lung cancer in young adults. *Am J Epidemiol* 1998;147:1028–37.
- Schwartz AG, Yang P, Swanson GM. Familial risk of lung cancer among nonsmokers and their relatives. *Am J Epidemiol* 1996;144:554–62.
- Coussens LM, Werb Z. Inflammation and cancer. *Nature* 2002;420:860–7.
- Nagase H, Visse R, Murphy G. Structure and function of matrix metalloproteinases and TIMPs. *Cardiovasc Res* 2006;69:562–73.
- Visse R, Nagase H. Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res* 2003;92:827–39.
- Pardo A, Selman M. MMP-1: the elder of the family. *Int J Biochem Cell Biol* 2005;37:283–8.
- Brinckerhoff CE, Matrisian LM. Matrix metalloproteinases: a tail of a frog that became a prince. *Nat Rev Mol Cell Biol* 2002;3:207–14.
- Chambers AF, Matrisian LM. Changing views of the role of matrix metalloproteinases in metastasis. *J Natl Cancer Inst* 1997;89:1260–70.
- Murray GI, Duncan ME, O'Neil P, Melvin WT, Fothergill JE. Matrix metalloproteinase-1 is associated with poor prognosis in colorectal cancer. *Nat Med* 1996;2:461–2.
- Pritchard SC, Nicolson MC, Lloret C, et al. Expression of matrix metalloproteinases 1, 2, 9 and their tissue inhibitors in stage II non-small cell lung cancer: implications for MMP inhibition therapy. *Oncol Rep* 2001;8:421–4.
- Schutz A, Schneidenbach D, Aust G, Tannapfel A, Steinert M, Wittekind C. Differential expression and activity status of MMP-1, MMP-2 and MMP-9 in tumor and stromal cells of squamous cell carcinomas of the lung. *Tumour Biol* 2002;23:179–84.
- Fang S, Jin X, Wang R, et al. Polymorphisms in the MMP1 and MMP3 promoter and non-small cell lung carcinoma in North China. *Carcinogenesis* 2005;26:481–6.
- 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.
- Zhu Y, Spitz MR, Lei L, Mills GB, Wu X. A single nucleotide polymorphism in the matrix metalloproteinase-1 promoter enhances lung cancer susceptibility. *Cancer Res* 2001;61:7825–9.
- Dally H, Edler L, Jager B, et al. The CYP3A4\*1B allele increases risk for small cell lung cancer: effect of gender and smoking dose. *Pharmacogenetics* 2003;13:607–18.
- Holle R, Happich M, Lowel H, Wichmann HE. KORA—a research platform for population based health research. *Gesundheitswesen* 2005;67:S19–25.
- Wichmann HE, Gieger C, Illig T. MONICA/KORA Study Group KORA-gen—resource for population genetics, controls and broad spectrum of disease phenotypes. *Gesundheitswesen* 2005;67:26–30.
- Steffens M, Lamina C, Illig T, et al. SNP-based analysis of genetic substructure in the German population. *Hum Hered* 2006;62:20–9.
- de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. *Nat Genet* 2005;37:1217–23.
- Vollmert C, Windl O, Xiang W, et al. Significant association of a M129V independent polymorphism in the 5' UTR of the PRNP gene with sporadic Creutzfeldt-Jakob disease in a large German case-control study. *J Med Genet* 2006;43:e53.
- Conneely KN, Boehnke M. So many correlated tests, so little time! Rapid adjustment of *P* values for multiple correlated tests. *Am J Hum Genet* 2007;81:6.
- Westfall Ph H, Young SS. Resampling-based multiple testing: examples and methods for *P*-value adjustment. New York: Wiley & Sons; 1993.
- Beckmann L, Thomas DC, Fischer C, Chang-Claude J. Haplotype sharing analysis using mantel statistics. *Hum Hered* 2002;59:67–78.
- Obreiter M, Fischer C, Chang-Claude J, Beckmann L. SDMinP: a program to control the family wise error rate using step-down minP adjusted *P*-values. *Bioinformatics* 2005;21:3183–4.
- Epstein MP, Satten GA. Inference on haplotype effects in case-control studies using unphased genotype data. *Am J Hum Genet* 2003;73:1316–29.
- Satten GA, Epstein MP. Comparison of prospective and retrospective methods for haplotype inference in case-control studies. *Genet Epidemiol* 2004;27:192–201.
- Scheet P, Stephens M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *Am J Hum Genet* 2006;78:629–44.
- Puente XS, Sanchez LM, Overall CM, Lopez-Otin C. Human and mouse proteases: a comparative genomic approach. *Nat Rev Genet* 2003;4:544–58.
- Lei H, Hemminki K, Altieri A, et al. Promoter polymorphisms in matrix metalloproteinases and their inhibitors: few associations with breast cancer susceptibility and progression. *Breast Cancer Res Treat* 2007;103:61–9.



38. 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.
39. Przybylowska K, Zielinska J, Zadrozny M, et al. An association between the matrix metalloproteinase 1 promoter gene polymorphism and lymphnode metastasis in breast cancer. *J Exp Clin Cancer Res* 2004;23:121–5.
40. Jurajda M, Muzik J, Izakovicova HL, Vacha J. A newly identified single nucleotide polymorphism in the promoter of the matrix metalloproteinase-1 gene. *Mol Cell Probes* 2002;16:63–6.
41. 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.
42. 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.
43. Heighway J, Bowers NL, Smith S, Betticher DC, Koref MF. The use of allelic expression differences to ascertain functional polymorphisms acting in cis: analysis of MMP1 transcripts in normal lung tissue. *Ann Hum Genet* 2005;69:127–33.
44. Terry CF, Loukaci V, Green FR. Cooperative influence of genetic polymorphisms on interleukin 6 transcriptional regulation. *J Biol Chem* 2000;275:18138–44.
45. Frisdal E, Klerkx AH, Le Goff W, et al. Functional interaction between –629C/A, –971G/A and –1337C/T polymorphisms in the CETP gene is a major determinant of promoter activity and plasma CETP concentration in the REGRESS Study. *Hum Mol Genet* 2005;14:2607–18.
46. Pearce EG, Laxton RC, Pereira AC, Ye S. Haplotype effects on matrix metalloproteinase-1 gene promoter activity in cancer cells. *Mol Cancer Res* 2007;5:221–7.
47. Sun T, Gao Y, Tan W, et al. Haplotypes in matrix metalloproteinase gene cluster on chromosome 11q22 contribute to the risk of lung cancer development and progression. *Clin Cancer Res* 2006;12:7009–17.
48. Singer CF, Marbaix E, Kokorine I, et al. The matrix metalloproteinase-1 (MMP-1) expression in the human endometrium is inversely regulated by interleukin-1 $\alpha$  and sex steroids. *Ceska Gynekol* 2000;65:211–5.
49. Nakopoulou L, Giannopoulou I, Gakiopoulou H, Liapis H, Tzonou A, Davaris PS. Matrix metalloproteinase-1 and -3 in breast cancer: correlation with progesterone receptors and other clinicopathologic features. *Hum Pathol* 1999;30:436–42.
50. Chen B, Wen Y, Wang H, Polan ML. Differences in estrogen modulation of tissue inhibitor of matrix metalloproteinase-1 and matrix metalloproteinase-1 expression in cultured fibroblasts from continent and incontinent women. *Am J Obstet Gynecol* 2003;189:59–65.
51. Kim H, Liu X, Kohyama T, et al. Cigarette smoke stimulates MMP-1 production by human lung fibroblasts through the ERK1/2 pathway. *COPD* 2004;1:13–23.
52. Lahmann C, Bergemann J, Harrison G, Young AR. Matrix metalloproteinase-1 and skin ageing in smokers. *Lancet* 2001;357:935–6.
53. Mercer B, Brinckerhoff C, D'Armiento J. Activation of the MMP-1 promoter by cigarette smoke in human small airway epithelial cells requires ERK MAP kinase signaling: differential response of the 1G and 2G promoter sequences. *Proc Am Thorac Soc* 2006;3:477.
54. Mercer BA, Kolesnikova N, Sonett J, D'Armiento J. Extracellular regulated kinase/mitogen activated protein kinase is up-regulated in pulmonary emphysema and mediates matrix metalloproteinase-1 induction by cigarette smoke. *J Biol Chem* 2004;279:17690–6.
55. Phillips JA, Bonassar LJ. Matrix metalloproteinase activity synergizes with  $\alpha_2\beta_1$  integrins to enhance collagen remodeling. *Exp Cell Res* 2005;310:79–87.
56. Yin L, Morita A, Tsuji T. Alterations of extracellular matrix induced by tobacco smoke extract. *Arch Dermatol Res* 2000;292:188–94.
57. Brinckerhoff CE, Rutter JL, Benbow U. Interstitial collagenases as markers of tumor progression. *Clin Cancer Res* 2000;6:4823–30.
58. Egeblad M, Werb Z. New functions for the matrix metalloproteinases in cancer progression. *Nat Rev Cancer* 2002;2:161–74.