# Variant genotypes and haplotypes of the epidermal growth factor gene promoter are associated with a decreased risk of gastric cancer in a high-risk Chinese population

**Guangfu Jin,[1,2] Ruifen Miao,[1] Yimei Deng,[1] Zhibin Hu,[1] Yan Zhou,[3] Yongfei Tan,[3] Jianming Wang,[4] Zhaolai Hua,[4] Weiliang Ding,[3] Lina Wang,[1] Wensen Chen,[1] Jing Shen,[1] Xinru Wang,[2] Yaochu Xu[1] and Hongbing Shen[1,2,5]**

[1]Department of Epidemiology and Biostatistics, Cancer Research Center of Nanjing Medical University, Nanjing 210029;[2] Laboratory of Toxicology, Nanjing Medical University School of Public Health, Nanjing 210029;[3] Yixing People's Hospital, Yixing City 214200, Jiangsu Province; and [4]Yangzhong Cancer Institute, Yangzhong City 212200, Jiangsu Province, China

(Received January 5, 2007/Revised January 14, 2007/Accepted February 15, 2007/Online publication April 6, 2007)

**Epidermal growth factor (EGF), a ligand of the EGF receptor, plays a critical role in the development of gastric cancer. Genetic variants in its promoter region may influence transcription activity and contribute to gastric cancer predisposition. To test this hypothesis, we genotyped three *EGF* promoter polymorphisms (G61A, G-1380A, and A-1744G) in a case–control study of 675 gastric cancer cases and 704 cancer-free controls. We found that the variant genotypes of *EGF* 61GA/AA were associated with a significantly decreased risk of gastric cancer (OR = 0.77, 95% CI = 0.61–0.95), when compared with wild-type homozygote 61GG. In the combined analysis with all three loci of *EGF*, subjects carrying one or more variant loci had a significantly decreased risk of gastric cancer in a dose–response manner (adjusted OR = 0.58, 95% CI = 0.42–0.80 for subjects carrying one variant locus and OR = 0.46, 95% CI = 0.32–0.66 for those carrying two to three variant loci, respectively; trend test: $\chi^2 = 16.14$, $P < 0.001$). Compared with the most common haplotype GGA, haplotypes AGA, GGG and GAA (each containing one variant allele) were associated with 33%, 29% and 34% significantly decreased risk of gastric cancer (adjusted OR = 0.67, 95% CI = 0.55–0.82 for AGA; OR = 0.71, 95% CI = 0.57–0.88 for GGG and OR = 0.66, 95% CI = 0.52–0.84 for GAA, respectively). Our findings indicate that variant genotypes and haplotypes of *EGF* promoter might play a role in gastric carcinogenesis. (*Cancer Sci* 2007; 98: 864–868)**

Gastric cancer is the second leading cause of cancer-related mortality worldwide, accounting for ~700 000 deaths annually.[1] Almost 40% of the gastric cancer cases occur in China with a remarkable geographic variation.[2] Epidemiological studies suggest that some environmental exposures (e.g. salty diet, tobacco smoking and *Helicobacter pylori* infection) are important for the development of gastric cancer.[3,4] However, accumulating evidence indicates that host factors and genetic alterations may also play an important role in gastric carcinogenesis through gene–environment interactions.[5]

The *EGF* gene encodes a ligand for the EGFR, a receptor of tyrosine kinase. When binding with EGFR, EGF can activate multiple signaling pathways, regulating cell proliferation and differentiation.[6–9] Studies showed that EGF and EGFR were highly expressed in gastric cancer,[10,11] cooperating with *H. pylori* and inflammatory cytokines in gastric carcinogenesis.[12,13]

Shahbazi *et al*. analyzed the *EGF* gene region from position –1350 to 164 and identified a G to A substitution at position 61 in the 5′ untranslated region, where the presence of the variant 61 A allele leads to a decreased *in vitro* EGF production in peripheral blood mononuclear cells.[14] Therefore, it was hypothesized that this promoter variant might be associated with risk of gastric cancer. In a hospital-based case-control study in Japan (200 cases and 230 controls), Hanai *et al*. reported that *EGF* G61A (rs4444903) was involved not only in the occurrence but also in the progression of gastric cancer.[15] However, this result was not supported by a later study in Japan (202 cases and 454 controls), although the main effect of *EGF* G61A was in the same direction.[16] Because the single locus may not represent the functional region of the gene promoter, it is biologically possible that other promoter variants may be also involved in gastric cancer susceptibility through a haplotype effect. Therefore, we used the public SNP database (http://www.ncbi.nlm.nih.gov/) to select SNPs in the promoter region of *EGF*. Apart from G61A, we chose two SNPs, G-1380A (rs11568835) and A-1744G (rs3756261), with a minor allele frequency >0.05 in a Chinese population.

To evaluate the effects of these three *EGF* promoter SNPs and haplotypes in gastric cancer susceptibility, we carried out genotyping analyses for SNPs of G61A, G-1380A and A-1744G in 675 gastric cancer cases and 704 cancer-free controls frequency-matched to the cases on age and sex.

## Materials and Methods

**Study subjects.** A total of 675 incident gastric cancer patients were consecutively recruited from the Yang-Zhong and Yi-Xing counties, two areas with high gastric cancer mortality, in Jiangsu Province, China, between January 2003 and July 2005. All of the cases were of local ethnic Han Chinese residents. Residents with histopathologically diagnosed adenocarcinoma in a stable medical condition as determined by their physician and who were willing to participate in the study and provide blood samples were included in the study. Subjects were not restricted by age or sex. The response rate of the cases was 89.4% (675/755). The eligible controls were cancer-free individuals who had lived in the study areas for at least 5 years without a self-reported history of any kind of cancer and were selected from a name list of the residents in each selected village, according to the frequencies of age (±5) and gender groups of the cases (frequency-matching). A total of 704 controls completed interviews and donated blood samples and the response rate was 84.8% (704/830).

[5]To whom correspondence should be addressed. E-mail: hbshen@njmu.edu.cn

Abbreviations: CHO, Chinese hamster ovary; CI, confidence interval; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; ELISA, enzyme-linked immunosorbent assay; GBM, glioblastoma multiforme; HNF1, hepatocyte nuclear factor 1; LD, linkage disequilibrium; OR, odds ratio; PCR, polymerase chain reaction; PIRA, primer-introduced restriction analysis; RFLP, restriction fragment length polymorphism; RTK, receptor of tyrosine kinase; SNP, single nucleotide polymorphism.



doi: 10.1111/j.1349-7006.2007.00463.x

After informed consent was obtained, each subject was personally interviewed face-to-face by trained interviewers using a standard questionnaire to obtain information on demographic data (e.g. age and gender) and related factors, including tobacco and alcohol use. After the interview, an approximately 5-mL venous blood sample was collected from each subject. Individuals that smoked one cigarette per day for over one year were defined as smokers, and those that consumed three or more alcohol drinks a week for over 6 months were considered alcohol drinkers. IgG antibodies to *H. pylori* infection was detected by an ELISA according to the manufacturer's instructions (Anti-*H. pylori* enzyme immunoassay, Bell Biotech Inc. Beijing, China). The study was approved by the Institutional Review Board of Nanjing Medical University.

**Genotyping.** Genomic DNA was extracted from a leukocyte pellet by proteinase K digestion and was followed by phenol-chloroform extraction and ethanol precipitation. The *EGF* G61A was genotyped by the PCR-RFLP assay as described previously.[14] Briefly, we used a pair of primers of 5′-TGTCACTAAAGGA-AAGGAGGT-3′ (sense) and 5′-TTCACAGAGTTTAACAGCCC-3′ (antisense) to generate a 242 bp PCR product, which was digested by restriction enzymes of *Alu*I (New England BioLabs, Beverly, MA, USA) and separated on a 3% agarose gel. The polymorphic (61A) allele produces four fragments of 102-, 91-, 34-, and 15-bp and the wild-type (61G) results in 193-, 34- and 15-bp fragments. The SNPs of G-1380A and A-1744G of *EGF* were genotyped using a PIRA-PCR assay,[17] as we previously reported.[18] For the *EGF* G-1380A polymorphism, a mismatched C was introduced into the sense-prime to replace T at –2 bp from the polymorphic site and a restriction site of *Hpa*II was created (sense-5′-CCTTCCATTGCTGTCATCCG, antisense-5′-CATTGCTTTCTGGACTGAGTCAGA). The 148-bp PCR products were then digested by *Hpa*II (New England BioLabs) and separated on a 3% agarose gel. The wild-type (–1380G) produced two fragments of 130- and 18-bp and the polymorphic (–1380A) allele resulted in a single 148-bp fragment. Similarly, for *EGF* A-1744G, a mismatched C was introduced into the sense-primer to replace T at –2 bp from the polymorphic sites and a *Pst*I restriction site was created. The primers were sense-5′-AGAGCTACCCAACTGGGAAGGATCT and antisense-5′-GGCCTCGATGCGCTTCCGCTTCA. The 121-bp PCR product was digested by *Pst*I (New England BioLabs) and separated on a 3% agarose gel. The variant allele *EGF* –1744G produced two fragments of 97- and 24-bp and the wild-type allele –1744A produced only one fragment of 121-bp.

Genotyping was carried out without knowing the subjects' case or control status and approximately equal numbers of the cases' and the controls' samples were assayed in each 96-well PCR plate with a positive control of a sample with a known heterozygous genotype. To further confirm the genotyping results, PCR products of the three loci with different genotypes were selected for direct sequencing using an automated sequencer (ABI model 377 genetic analysis; Perkin-Elmer Applied Biosystems).

**Statistical analyses.** Differences in the distributions of demographic characteristics, selected variables, and genotypes of the *EGF* variants between the cases and controls were evaluated using the $\chi^2$ test. The associations between *EGF* genotypes and risk of gastric cancer were estimated by computing the ORs and their 95% CIs using logistic regression analyses for crude ORs and adjusted ORs when adjusting for age, sex, smoking and drinking status. We used the PHASE 2.0 program,[19] to infer haplotype frequencies based on the observed *EGF* genotypes. The Hardy–Weinberg equilibrium was tested by a goodness-of-fit $\chi^2$ test to compare the observed genotype frequencies with the expected ones among the control subjects. All statistical analyses were carried out with SAS 9.1.3 (SAS Institute, Cary, NC, USA).

**Table 1. Distribution of selected demographic variables and risk factors in gastric cancer cases and controls**

| Variable | Cases ($n$ = 617) | | Controls ($n$ = 660) | | $P$-value* |
|---|---|---|---|---|---|
| | $n$ | % | $n$ | % | |
| Age (years) | | | | | 0.275 |
| <60 | 271 | 43.9 | 310 | 47.0 | |
| ≥60 | 346 | 56.1 | 350 | 53.0 | |
| Sex | | | | | 0.433 |
| Male | 424 | 68.7 | 440 | 66.7 | |
| Female | 193 | 31.3 | 220 | 33.3 | |
| Tobacco use | | | | | 0.921 |
| Never | 376 | 60.9 | 404 | 61.2 | |
| Tobacco user | 241 | 39.1 | 256 | 38.8 | |
| Alcohol use | | | | | 0.051 |
| Never | 411 | 66.6 | 405 | 61.4 | |
| Alcohol user | 206 | 33.4 | 255 | 38.6 | |
| *H. pylori*† | | | | | 0.020 |
| Negative | 150 | 46.9 | 250 | 39.0 | |
| Positive | 170 | 53.1 | 391 | 61.0 | |

*Two-sided $\chi^2$ test. †320 gastric cancer cases and 641 controls were detected.

## Results

Among the 675 cases and 704 controls with DNA samples, the genotyping was successful for all three polymorphisms in 617 gastric cancer cases and 660 controls, resulting in an overall success rate of 92.6% (G61A, 94.8%; G-1380A, 98.0%; A-1744G, 97.0%). Therefore, a total of 617 cases and 660 controls with complete genotype information for the above three SNPs were included in the final analyses. The mean age was 60.5 years (±9.4 years, ranging 30–82 years) for the case group and 59.6 years (±10.4 years, ranging 24–84 years) for the control group ($P$ = 0.131). As shown in Table 1, there were no significant differences in terms of distributions on age (<60 and =60 years old) and gender between the cases and the controls ($P$ = 0.275 and 0.433, respectively), suggesting that our frequency matching of the demographic characteristics was satisfactory. In addition, there were no significant differences between the cases and the controls in smoking and drinking status (OR = 1.01, 95% CI = 0.81–1.27, $P$ = 0.921 for smoking and OR = 0.80, 95% CI = 0.63–1.01, $P$ = 0.051 for alcohol use, respectively). However, the positive rate of IgG antibodies against *H. pylori* was higher in control subjects (61.0%) than in gastric cancer patients (53.1%) ($P$ = 0.020) (Table 1).

The allele and genotype distributions of *EGF* G61A, G-1380A and A-1744G polymorphisms in the cases and controls are shown in Table 2. The observed genotype frequencies for these three SNPs were all in Hardy–Weinberg Equilibrium in the controls ($P$ = 0.407, 0.088 and 0.119 for *EGF* G61A, G-1380A and A-1744G, respectively). In the single locus analyses, none of the three polymorphisms achieved significant differences in the genotype distributions between the cases and the controls ($P$ = 0.062, 0.445, and 0.054 for *EGF* G61A, G-1380A and A-1744G, respectively). The logistic regression analyses revealed that the 61GA heterozygote was associated with a significantly reduced risk of gastric cancer (adjusted OR = 0.78, 95% CI = 0.62–0.99) and the 61AA homozygote was associated with a non-significantly decreased risk (adjusted OR = 0.68, 95% CI = 0.44–1.05), compared with the 61GG wild-type homozygote. When we combined the variant genotypes (61GA/AA) assuming a codominant genotype effect, the combined 61GA/AA variant genotypes were associated with a significantly decreased risk of

**Table 2. Logistic regression analyses of associations between epidermal growth factor promoter polymorphisms and risk of gastric cancer**

| Genotype | Cases ($n$ = 617) | | Controls ($n$ = 660) | | Crude OR (95% CI) | Adjusted OR† (95% CI) |
|---|---|---|---|---|---|---|
| | $n$ | % | $n$ | % | | |
| *EGF* G61A | | | | | | |
| GG | 333 | 54.0 | 314 | 47.6 | 1.00 | 1.00 |
| GA | 242 | 39.2 | 289 | 43.8 | 0.79 (0.63–0.99) | 0.78 (0.62–0.99) |
| AA | 42 | 6.8 | 57 | 8.6 | 0.70 (0.45–1.07) | 0.68 (0.44–1.05) |
| GA/AA | 284 | 46.0 | 346 | 52.4 | 0.77 (0.62–0.96) | 0.77 (0.61–0.95) |
| A allele | 326 | 26.4 | 403 | 30.5 | | |
| *EGF* G-1380A | | | | | | |
| GG | 451 | 73.1 | 462 | 70.0 | 1.00 | 1.00 |
| GA | 143 | 23.2 | 173 | 26.2 | 0.85 (0.66–1.10) | 0.83 (0.64–1.08) |
| AA | 23 | 3.7 | 25 | 3.8 | 0.94 (0.53–1.69) | 0.95 (0.53–1.70) |
| GA/AA | 166 | 26.9 | 198 | 30.0 | 0.86 (0.67–1.10) | 0.85 (0.66–1.08) |
| A allele | 189 | 15.3 | 223 | 16.9 | | |
| *EGF* A-1744G | | | | | | |
| AA | 408 | 66.1 | 403 | 61.1 | 1.00 | 1.00 |
| AG | 181 | 29.3 | 234 | 35.5 | 0.76 (0.60–0.97) | 0.77 (0.61–0.98) |
| GG | 28 | 4.5 | 23 | 3.5 | 1.20 (0.68–2.12) | 1.25 (0.71–2.22) |
| AG/GG | 209 | 33.9 | 257 | 38.9 | 0.80 (0.64–1.01) | 0.81 (0.65–1.02) |
| G allele | 237 | 19.2 | 280 | 21.2 | | |
| Combined analysis‡ | | | | | | |
| No variant locus | 117 | 19.0 | 74 | 11.2 | 1.00 | 1.00 |
| One variant locus | 343 | 55.6 | 373 | 56.5 | 0.58 (0.42–0.81) | 0.58 (0.42–0.80) |
| Two to three variant loci | 157 | 25.5 | 213 | 32.3 | 0.47 (0.33–0.67) | 0.46 (0.32–0.66) |

†Adjusted for age, sex, smoking status and drinking status. ‡Assuming dominant genetic model in each locus, that is, heterozygote and variant homozygote versus wild type homozygote. CI, confidence interval; EGF, epidermal growth factor gene; OR, odds ratio.

**Table 3. Frequencies of inferred haplotypes based on observed genotypes in gastric cancer cases and cancer-free controls**

| *EGF* haplotypes | | | Allele frequencies* | | | | OR (95% CI) |
|---|---|---|---|---|---|---|---|
| G61A alleles | G-1380A alleles | A-1744G alleles | Cases ($n$ = 1234) | | Controls ($n$ = 1320) | | |
| | | | $N$% | % | $N$ | % | |
| G | G | A | 512 | 41.5 | 435 | 32.9 | 1.00 |
| A | G | A | 305 | 24.7 | 386 | 29.3 | 0.67 (0.55–0.82) |
| G | G | G | 218 | 17.7 | 262 | 19.8 | 0.71 (0.57–0.88) |
| G | A | A | 169 | 13.7 | 217 | 16.4 | 0.66 (0.52–0.84) |
| Others† | | | 30 | 2.4 | 20 | 1.6 | 1.27 (0.71–2.28) |

*$P$ < 0.001 for haplotypes distribution among cases and controls. †Including AGG, GAG and AAA haplotyope (haplotype frequencies < 0.05). CI, confidence interval; *EGF*, epidermal growth factor gene; OR, odds ratio.

gastric cancer (adjusted OR = 0.77, 95% CI = 0.61–0.95). For *EGF* G-1380A and A-1744G SNPs, the combined variant genotypes of both –1380GA/AA and –1744AG/GG were associated with a non-significantly decreased risk (adjusted OR = 0.85, 95% CI = 0.66–1.08 for –1380GA/AA and OR = 0.81, 95% CI = 0.65–1.02 for –1744AG/GG, respectively), compared with their wild-type homozygotes, respectively. Specifically, instead of the *EGF* –1744GG variant homozygote, the –1744AG heterozygote was associated with a significantly decreased risk of gastric cancer with adjusted OR and 95% CI of 0.77 and 0.61–0.98, compared with the –1744AA wild-type genotype (Table 2).

To evaluate the combined effects of the three *EGF* promoter SNPs on gastric cancer risk, we defined the number of heterozygotes or variant homozygotes of the three loci as the number of variant locus. As shown in Table 2, compared with the subjects carrying the wild-type genotypes for all three loci, those carrying one or more variant loci had a significantly decreased risk of gastric cancer in a dose–response manner (adjusted OR = 0.58, 95% CI = 0.42–0.80 for subjects carrying one variant locus; and adjusted OR = 0.46, 95% CI = 0.32–0.66 for those with two to three variant loci; trend test: $\chi^2 = 16.14$, $P < 0.001$) (Table 2).

In the LD analyses, we found that all of the three loci were in LD with each other ($\chi^2 = 170.4$, $P < 0.001$ and $D' = 0.93$ for G61A and G-1380A; $\chi^2 = 187.5$, $P < 0.001$ and $D' = 0.85$ for G61A and A-1744G; $\chi^2 = 88.8$, $P < 0.001$ and $D' = 0.84$ for G-1380A and A-1744G), suggesting that there might be haplotype effects among these three SNPs. When we combined these three loci and carried out the haplotype inference using the PHASE 2.0 program, totally seven possible haplotypes were derived from their known genotypes (Table 3). Among them, four common (>10%) haplotypes (GGA, AGA, GGG, GAA) represented 97.6% of the chromosomes for cases and 98.4% of that for controls, and the haplotype distribution between the cases and the controls was statistically different ($\chi^2 = 31.72$, d.f. = 6, $P < 0.001$). Specifically, haplotype AGA, GGG and GAA, each containing one of the three variant alleles, were less common in the cases (AGA: 0.247, GGG: 0.177 and GAA: 0.137) than in the controls (0.293, 0.198 and 0.164, respectively) ($P < 0.001$ for AGA, $P = 0.002$ for GGG and $P < 0.001$ for GAA, respectively). Compared with the most common haplotype GGA (consisting of the common allele from each polymorphic site), the AGA, GGG and GAA haplotypes were associated with 33%, 29%


doi: 10.1111/j.1349-7006.2007.00463.x

**Table 4. Stratified analyses between the combined genotypes of epidermal growth factor promoter polymorphisms and gastric cancer risk**

| Variant loci | Cases ($n$ = 617) | | Controls ($n$ = 660) | | Adjusted OR* (95% CI) | | $P$-value* |
|---|---|---|---|---|---|---|---|
| | 0 | ≥1 | 0 | ≥1 | 0 | ≥1 | |
| | $n$ (%) | $n$ (%) | $n$ (%) | $n$ (%) | | | |
| All subjects | 117 (19.0) | 500 (81.0) | 74 (11.2) | 586 (88.8) | 1.00 | 0.53 (0.39–0.73) | <0.001 |
| Age (years) | | | | | | | |
| <60 | 51 (18.8) | 220 (81.2) | 40 (12.9) | 270 (87.1) | 1.00 | 0.63 (0.40–0.99) | 0.050 |
| ≥60 | 66 (19.1) | 280 (80.9) | 34 (9.7) | 316 (90.3) | 1.00 | 0.45 (0.29–0.70) | <0.001 |
| Sex | | | | | | | |
| Male | 76 (17.9) | 348 (82.1) | 53 (12.1) | 387 (88.0) | 1.00 | 0.61 (0.42–0.90) | 0.012 |
| Female | 41 (21.2) | 152 (78.8) | 21 (9.6) | 199 (90.5) | 1.00 | 0.40 (0.22–0.70) | 0.002 |
| Tobacco use | | | | | | | |
| Never | 71 (18.9) | 305 (81.1) | 43 (10.6) | 361 (89.4) | 1.00 | 0.50 (0.33–0.75) | 0.001 |
| Tobacco user | 46 (19.1) | 195 (80.9) | 31 (12.1) | 225 (87.9) | 1.00 | 0.58 (0.35–0.96) | 0.032 |
| Alcohol use | | | | | | | |
| Never | 78 (19.0) | 333 (81.0) | 42 (10.4) | 363 (89.6) | 1.00 | 0.49 (0.33–0.73) | 0.001 |
| Alcohol user | 39 (18.9) | 167 (81.1) | 32 (12.6) | 223 (87.5) | 1.00 | 0.60 (0.35–1.01) | 0.053 |
| *H. pylori*† | | | | | | | |
| Negative | 26 (17.3) | 124 (82.7) | 31 (12.4) | 219 (87.6) | 1.00 | 0.69 (0.39–1.24) | 0.212† |
| Positive | 32 (18.8) | 138 (81.2) | 43 (11.0) | 348 (89.0) | 1.00 | 0.51 (0.31–0.85) | 0.010† |

*Adjusted for age, sex, smoking and drinking status. †320 gastric cancer cases and 641 controls were detected for *Helicobacter pylori* infection.
CI, confidence interval; OR, odds ratio.

and 34% significantly decreased risk of gastric cancer (adjusted OR = 0.67, 95% CI = 0.55–0.82 for AGA; OR = 0.71, 95% CI = 0.57–0.88 for GGG and OR = 0.66, 95% CI = 0.52–0.84 for GAA, respectively).

The dichotomized genotypes (one or more variant loci *versus* no variant locus) were further examined for subgroups by selected variables according to the number of variant loci. As shown in Table 4, the decreased risk associated with the combined genotypes with one or more variant loci was more pronounced in subjects who were non-drinkers (OR = 0.49, 95% CI = 0.33–0.73) and carriers of *H. pylori* (OR = 0.51, 95% CI = 0.31–0.85). However, there were no significant differences in the magnitude of the associations between the combined genotypes and gastric cancer risk for subjects with different age, gender and smoking status (Table 4). In addition, we also did not find any significant gene–environment interactions in relation to risk of gastric cancer (data not shown).

## Discussion

In this population-based case-control study, we investigated the associations of three SNPs in the promoter region of *EGF* with risk of gastric cancer in a high-risk Chinese population. We found that the variant genotypes of *EGF* 61GA/AA, –1380GA/AA and –1744AG/GG were associated with a decreased risk of gastric cancer, compared with their wild-type homozygotes. In the combined analyses with these three loci of *EGF*, the more variant loci of three promoter SNPs of *EGF* there were, the lower risks of gastric cancer were observed. Compared with the most common haplotype GGA, haplotypes AGA, GGG and GAA (each containing one variant allele) were associated with a significantly decreased risk of gastric cancer. To the best of our knowledge, this is the first study that investigated the association of these three variant genotypes and haplotypes in the promoter region of *EGF* and the risk of gastric cancer.

EGF regulates cell proliferation, differentiation and survival of normal cells by binding and activating EGFR.[20] The *EGFR* and the EGF-family of peptide growth factor play an important role in the development and progression of diverse carcinoma types, including gastric cancer.[11,13,21] Dysregualtion of EGFR pathway may occur through mutations in *EGFR* and *EGF*, resulting in constitutive activation.[20,21] Therefore, genetic variants in the *EGF* gene were hypothesized to play a critical role in carcinogenesis. Shahbazi *et al*. firstly identified a G to A substitution at position 61 in the 5′ untranslated region of the *EGF* gene, and they found that the presence of the 61G allele leads to increased EGF production *in vitro* in peripheral blood mononuclear cells.[14] Furthermore, they showed that the 61G allele conferred a 2.7-fold increased risk of melanoma,[14] although this was not supported by later studies.[22–25] In addition, Bhowmick *et al*. found that *EGF* expression was significantly lower in tumors with 61AA genotype when compared with 61GA ($P < 0.011$) or 61GG ($P < 0.004$) and the frequency of G allele in GBM patients was significantly greater than that in normal controls ($P < 0.001$).[26] More recently, Vauleon *et al*. showed that the *EGF* 61G allele could result in a 40% higher luciferase activity in CHO cells than 61A allele in an *in vitro* study, because the G allele might enhance the affinity of the HNF1 transcription factor at position 61.[27,28] Taken together, this evidence suggests that the *EGF* G61A polymorphism might be a functional variant in modifying the risks for different kinds of cancers.

In a small case-control study of gastric cancer with 200 cases and 230 controls in a Japanese population, Hamai *et al*. first investigated the association of *EGF* G61A with gastric cancer risk and found that the 61GA/AA genotypes were associated with a significantly 40% decreased risk of gastric cancer (OR = 0.60, 95% CI = 0.41–0.88).[15] In another case-control study in a Japanese population (202 cases and 454 controls), Goto *et al*. also showed a protective but not significant effect on gastric cancer for the variant 61A allele.[16] In the present study with relatively a large sample size, we found that variant 61A allele of *EGF* promoter was associated with a 23% significantly decreased risk of gastric cancer in this high-risk Chinese population, which was consistent with the results reported by Hamai.[15] In addition, we found that the variant alleles of the *EGF* promoter, –1380A and –1744G, were also associated with a decreased risk of gastric cancer with borderline significance, indicating that they might play a similar role in gastric carcinogenesis through linkage disequilibriums between these loci. It is also possible, however, that these three polymorphisms might be in linkage disequilibrium with other putative etiological genetic variants.

Interestingly, we found for the first time that *EGF* haplotypes AGA, GGG and GAA were associated with a significantly decreased risk of gastric cancer when compared with the most common haplotype GGA (containing three wild-type alleles). These three haplotypes, each containing one variant allele, could represent the corresponding variant alleles of the polymorphic sites in this region, suggesting that these three SNPs might contribute to gastric cancer susceptibility in an allele-specific manner. However, these three haplotypes incorporating all three loci showed a stable and additive effect on a decreased risk of gastric cancer as compared with any single locus (with corresponding variant allele), which might result from the fact that the effect of haplotypes was greater than a single locus. Recent studies supported that haplotype analyses might be of advantage in investigating the disease associations, compared with single locus analyses.[29–31]

In the two Japanese case-control studies, the allele frequencies of *EGF* 61A in controls were 0.304 and 0.313, respectively,[15,16] which were comparable with that in this Chinese population (0.305). However, 61A allele frequencies were significantly higher in Caucasian populations (above 0.50),[14,22–26] and therefore, population stratification should be paid attention to in further studies. Because this was the first molecular epidemiological study that investigated the association between G-1380A and A-1744G polymorphisms and cancer susceptibility, no comparison between published studies could be made at this time. According to the NCBI dbSNP database (http://www.ncbi.nlm.nih.gov/), the allele frequencies in 44 unrelated Han Chinese in Beijing and 24 Han Chinese in Los Angeles were 0.068 and 0.188 for –1380A allele and 0.239 and 0.125 for –1744G allele, respectively, which was not different from our Chinese population: –1380A allele (0.169) and –1744G allele (0.212).

One limitation of the current study is that detailed information on clinical staging, metastasis and survival of gastric cancer were not available, which restricted our further analysis on the role of *EGF* in cancer progression and prognosis. Another limitation is the lack of the related phenotypic and functional evaluations on the *EGF* SNPs, which limited our inquiry into the functional consequence of these variants. Furthermore, the plasma positive rate of IgG antibodies to *H. pylori* by ELISA may not represent the real *H. pylori* infection before the disease occurred and therefore we obtained negative association between *H. pylori* infection and gastric cancer risk. A possible explanation was that the loss of *H. pylori* from the stomach occurred during gastric carcinogenesis or after the long-term antibiotic treatment. It was also possible that the immune response to *H. pylori* infection was reduced with the development of gastric cancer.[4,32] Finally, occupational exposure and certain dietary components might act as potential confounders in the analysis. Unfortunately, information on these factors in our case-control study was not available.

In summary, our results indicate that genetic variants in the promoter region of *EGF* may play a role in gastric cancer susceptibility. To explore the exact biological mechanism of *EGF* genotypes and haplotypes and their interaction with environmental factors, further functional studies and larger well-designed prospective studies are warranted.

## Acknowledgments

This study was supported in part by the National Natural Science Foundation of China, Grant numbers: 30671814 and 30571605; Jiangsu Natural Science Foundation, Grant number: BK2005143.

## References

1 Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin* 2005; **55**: 74–108.
2 Yang L. Incidence and mortality of gastric cancer in China. *World J Gastroenterol* 2006; **12**: 17–20.
3 Crew KD, Neugut AI. Epidemiology of upper gastrointestinal malignancies. *Semin Oncol* 2004; **31**: 450–64.
4 Uemura N, Okamoto S, Yamamoto S *et al*. Helicobacter pylori infection and the development of gastric cancer. *N Engl J Med* 2001; **345**: 784–9.
5 Lichtenstein P, Holm NV, Verkasalo PK *et al*. Environmental and heritable factors in the causation of cancer – analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 2000; **343**: 78–85.
6 Jorissen RN, Walker F, Pouliot N, Garrett TP, Ward CW, Burgess AW. Epidermal growth factor receptor: mechanisms of activation and signalling. *Exp Cell Res* 2003; **284**: 31–53.
7 Olayioye MA, Neve RM, Lane HA, Hynes NE. The ErbB signaling network. Receptor heterodimerization in development and cancer. *EMBO J* 2000; **19**: 3159–67.
8 Laurence DJ, Gusterson BA. The epidermal growth factor. A review of structural and functional relationships in the normal organism and in cancer cells. *Tumour Biol* 1990; **11**: 229–61.
9 Singletary SE, Baker FL, Spitzer G *et al*. Biological effect of epidermal growth factor on the in vitro growth of human tumors. *Cancer Res* 1987; **47**: 403–6.
10 Yasui W, Hata J, Yokozaki H *et al*. Interaction between epidermal growth factor and its receptor in progression of human gastric carcinoma. *Int J Cancer* 1988; **41**: 211–17.
11 Tokunaga A, Onda M, Okuda T *et al*. Clinical significance of epidermal growth factor (EGF), EGF receptor, and c-erbB-2 in human gastric cancer. *Cancer* 1995; **75**: 1418–25.
12 Coyle WJ, Sedlack RE, Nemec R *et al*. Eradication of Helicobacter pylori normalizes elevated mucosal levels of epidermal growth factor and its receptor. *Am J Gastroenterol* 1999; **94**: 2885–9.
13 Tanida S, Joh T, Itoh K *et al*. The mechanism of cleavage of EGFR ligands induced by inflammatory cytokines in gastric cancer cells. *Gastroenterology* 2004; **127**: 559–69.
14 Shahbazi M, Pravica V, Nasreen N *et al*. Association between functional polymorphism in EGF gene and malignant melanoma. *Lancet* 2002; **359**: 397–401.
15 Hamai Y, Matsumura S, Matsusaki K *et al*. A single nucleotide polymorphism in the 5′ untranslated region of the EGF gene is associated with occurrence and malignant progression of gastric cancer. *Pathobiology* 2005; **72**: 133–8.
16 Goto Y, Ando T, Goto H, Hamajima N. No association between EGF gene polymorphism and gastric cancer. *Cancer Epidemiol Biomarkers Prev* 2005; **14**: 2454–6.
17 Ke X, Collins A, Ye S. PIRA PCR designer for restriction analysis of single nucleotide polymorphisms. *Bioinformatics* 2001; **17**: 838–9.
18 Hu Z, Wang Y, Wang X *et al*. DNA repair gene XPC genotypes/haplotypes and risk of lung cancer in a Chinese population. *Int J Cancer* 2005; **115**: 478–83.
19 Stephens M, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* 2003; **73**: 1162–9.
20 Normanno N, De Luca A, Bianco C *et al*. Epidermal growth factor receptor (EGFR) signaling in cancer. *Gene* 2006; **366**: 2–16.
21 Normanno N, Bianco C, Strizzi L *et al*. The ErbB receptors and their ligands in cancer: an overview. *Curr Drug Targets* 2005; **6**: 243–57.
22 McCarron SL, Bateman AC, Theaker JM, Howell WM. EGF +61 gene polymorphism and susceptibility to and prognostic markers in cutaneous malignant melanoma. *Int J Cancer* 2003; **107**: 673–5.
23 Amend KL, Elder JT, Tomsho LP *et al*. EGF gene polymorphism and the risk of incident primary melanoma. *Cancer Res* 2004; **64**: 2668–72.
24 Randerson-Moor JA, Gaut R, Turner F *et al*. The relationship between the epidermal growth factor (EGF) 5′UTR variant A61G and melanoma/nevus susceptibility. *J Invest Dermatol* 2004; **123**: 755–9.
25 James MR, Hayward NK, Dumenil T, Montgomery GW, Martin NG, Duffy DL. Epidermal growth factor gene (EGF) polymorphism and risk of melanocytic neoplasia. *J Invest Dermatol* 2004; **123**: 760–2.
26 Bhowmick DA, Zhuang Z, Wait SD, Weil RJ. A functional polymorphism in the EGF gene is found with increased frequency in glioblastoma multiforme patients and is associated with more aggressive disease. *Cancer Res* 2004; **64**: 1220–3.
27 Vauleon E, Auger N, Benouaich-Amiel A *et al*. The 61 A/G EGF polymorphism is functional but is neither a prognostic marker nor a risk factor for glioblastoma. *Cancer Genet Cytogenet* 2007; **172**: 33–7.
28 Ktistaki E, Talianidis I. Modulation of hepatic gene expression by hepatocyte nuclear factor 1. *Science* 1997; **277**: 109–12.
29 Judson R, Stephens JC, Windemuth A. The predictive power of haplotypes in clinical response. *Pharmacogenomics* 2000; **1**: 15–26.
30 Judson R, Stephens JC. Notes from the SNP vs. haplotype front. *Pharmacogenomics* 2001; **2**: 7–10.
31 Park JY, Park JM, Jang JS *et al*. Caspase 9 promoter polymorphisms and risk of primary lung cancer. *Hum Mol Genet* 2006; **15**: 1963–71.
32 Plummer M, Vivas J, Fauchere JL *et al*. *Helicobacter pylori* and stomach cancer: a case-control study in Venezuela. *Cancer Epidemiol Biomarkers Prev* 2000; **9**: 961–5.


doi: 10.1111/j.1349-7006.2007.00463.x