

## RESEARCH ARTICLE

# Genetic Epidemiological Analysis of Esophageal Cancer in High-incidence Areas of China

Kai-Juan Wang<sup>1,2\*</sup>, Jun-Xia Yang<sup>1,2</sup>, Jia-Chen Shi<sup>1</sup>, Song-Yuan Deng<sup>2</sup>, Xiao-Qin Cao<sup>1,2</sup>, Chun-Hua Song<sup>1,2</sup>, Peng Wang<sup>1,2</sup>

### Abstract

Genetic epidemiological studies have shown that genetic susceptibility to esophageal cancer (EC) is an important cause of its high incidence within families in some areas of China. The purpose of this study was to obtain evidence of a genetic basis of EC in Xin-an and Xin-xiang counties in China. Familial aggregation and complex segregation analyses were performed of 79 EC families in these counties. The heritability of EC was examined using Falconer's method and complex segregation analysis was conducted with the SEGREG program in Statistical Analysis for Genetic Epidemiology (SAGE version 5.3.1). The results showed that the distribution of EC in families did not fit well into a binomial distribution. The heritability of EC among first-degree and second-degree relatives was  $67.0 \pm 7.31\%$  and  $43.1 \pm 9.80\%$ , respectively, and the summing up powered heritability was  $53.2 \pm 6.74\%$ . The segregation ratio was 0.045. Complex segregation analysis showed that the genetic model of EC was additive. The current results provide evidence for an inherited propensity to EC in certain high-risk groups in China, and support efforts to identify the genes that confer susceptibility to this disease.

**Keywords:** Esophageal cancer - familial aggregation - genetic model - complex segregation analysis - heritability

*Asian Pac J Cancer Prev*, 15 (22), 9859-9863

### Introduction

Esophageal carcinoma is the sixth leading cause of cancer-related mortality and the eighth most common cancer worldwide (Pennathur et al., 2013). In China, esophageal cancer ranks the 5th most common diagnosed cancer and 4th leading cause of cancer related mortality (Chen et al., 2013). In most cases, EC is characterized by rapid progression and fatal prognosis. The highest incidence is within the age group between 50 and 70 years old. Three to five times more men are diagnosed than women. The histological type is most frequently squamous cell carcinoma (Kollarova et al., 2007).

In China, high-incidence areas are mainly along the northern borders of the three provinces Hebei, Henan, and Shanxi that about the southern flank of the Taihang Mountains in the north-central region (Olopade and Pichert, 2001; Tran et al., 2005; Guohong et al., 2010; Tang et al., 2014). Over the last two decades, some studies have sought epidemiological evidence of genetic susceptibility to EC, as well as effective ways for screening individuals who are at high risk of the disease (Carter et al., 1992; Zhang et al., 2000; Qiao et al., 2001; Engel et al., 2003). Familial aggregation of EC was found in these high-incidence areas (Hu et al., 1992). Epidemiological and other scientific investigations have

clearly indicated that certain environmental factors are associated with EC, including nitrosamines, mycotoxins, Human papillomavirus, and microelement, etc. (Dong et al., 2008). In recent years, the etiology of EC has been more extensively studied from environmental, mental, and hereditary perspectives and less from genetic epidemiology. Furthermore, the data was incomplete in pedigree investigation.

Xin-an and Xin-xiang counties in Henan Province are rural areas, which have not undergone any epidemiologic surveillance for malignant tumors. In 2003, the official registration of death data showed that EC is the foremost cause of cancer death in local residents. There has been increasing awareness of the importance of genetic factors in the pathogenesis of EC as research has provided a greater understanding of the mechanisms involved (Masao Kanamori et al., 2001). The aim of the present study was to obtain evidence of inherited susceptibility to EC in high-risk populations. Our results support the need for further studies to identify genes that confer a high familial risk of EC and localize them in the population.

### Materials and Methods

#### *Study subjects*

This study comprised 472 subjects from 79 EC

<sup>1</sup>Department of Epidemiology, College of Public Health, Zhengzhou University, <sup>2</sup>Henan Key Laboratory of Tumor Epidemiology, Zhengzhou, Henan Province, China \*For correspondence: [kjwang@163.com](mailto:kjwang@163.com)

families. All subjects were recruited from a population-based study conducted on EC probands obtained from a census carried out from August 2005 to August 2007 in Xin-an and Xin-xiang counties of Henan Province in China.

#### *Subject interview and data management*

The pedigrees of probands from four generations were analyzed. Patients were diagnosed by histopathology conducted at the hospitals at the county level or above. All of the investigators received specific instructions on how to conduct the study and the questionnaires were completed with face-to-face interviews or by assessing patients' medical records. A structured questionnaire was administered to the patients, who were asked to provide verbal answers. The collected information included lifestyle habits (e.g., tobacco smoking and alcohol drinking) and diet, as well as the family history of four successive generations. Whenever a proband was unsure of an answer, additional family members or older neighbors were interviewed to ensure the accuracy of the information. Families were ascertained through the single proband. Information regarding familial EC occurrence was obtained through face-to-face interviews. The date of diagnosis was recorded for the probands and all the affected relatives. All of the questionnaires were reviewed and checked for quality control. Data from 472 subjects in 79 families were then sorted and placed into a database using EpiData 3.1 software (EpiData, Odense, Denmark). All data were reviewed twice for accuracy.

Analyses were performed using Statistical Package for Social Sciences (SPSS, Version 13.0) and Statistical Analysis for Genetic Epidemiology (S.A.G.E. Version.5.3.1) (Hanna et al., 2005) software. The chi-squared ( $\chi^2$ ) test and t-test were used to compare differences in the data. The heritability ( $h^2$ ), segregation ratio (PLM), and odds ratios (ORs) with 95% confidence intervals (95%CI) were used to describe the genetic risk of EC. This study was conducted with the permission of the Institutional Review Board of Zhengzhou University. All subjects provided signed informed consent prior to participating in the study, and all remained anonymous, with information kept as coded data.

#### *Familial aggregation analysis of EC*

Familial aggregation analysis of EC was performed by testing the distribution of EC cases among the subjects of 79 families. The null hypothesis of non-familial aggregation of EC was tested by Tarone's one-sided score test for binomial distributions (Tarone R E, 1979). A low and non-significant score indicated an expected random distribution of a variable.

#### *Estimation heritability of EC*

The estimation of heritability was determined with Falconer's threshold models (Douglas S. Falconer, 1981), which is a typical approach for determining binary and ordered categorical traits. In accordance with Falconer's method, values for  $h^2$  and the 95%CI were obtained by using the prevalence rate of EC among relatives of the selected proband.

#### *Segregation analysis*

Segregation analysis was performed using the Li-Mantel method (Li and Mantel, 1968). This method is based on an assumption of complete ascertainment whereby all affected cases are ascertained independently from their affected siblings.

#### *Complex segregation analysis*

Complex segregation analysis was performed using the S.A.G.E. program SEGREG in order to evaluate possible inheritance models of EC (Stricker et al., 1995). SEGREG is a program based on the regression model by Bonney (Stricker et al., 1995) and is configured for segregation analysis under a class A regressive model (with the possibility of including a common sibship component that is dependent on the proportion of siblings affected) either for a truncated trait, such as the age at onset of disease that follows a logistic distribution (possibly after transformation; Model 1), or for susceptibility to the disease (Model 2).

The disease is a discrete trait with a variable age of onset. Under Model 1, the genotype is presumed to influence the age of disease onset through location susceptibility, which is defined as the probability of being affected by age taken to infinity. Susceptibility may be different for the two classes. Under Model 2, the genotype is presumed to influence the susceptibility to the affected state, but not to affect the age of disease onset (Lee et al., 2008). Model 1 is appropriate for the analysis of EC.

Under Model 1, six hypotheses were tested against the likelihood of a general (unrestricted) model, in which all parameters were unrestricted and allowed to fit the empirical data. Therefore, the general model would give the best fit for the data. The six hypotheses of transmission are: major gene type, Mendelian dominant, Mendelian recessive, Mendelian additive, purely environmental effect, and no transmission. Twice the difference between the natural log likelihood (lnL) for the data under the hypothesis of interest and the lnL under the unrestricted model is in accordance with  $\chi^2$  distribution. Therefore, it is used to statistically assess the degree of departure from expectation. The degrees of freedom (df) for the  $\chi^2$  statistic is given by the differences in the number of estimated parameters between the hypothesis and the unrestricted model. If one or more parameters are fixed at a bound at the end of the estimation process, a range of df and P values are given when appropriate. A non-significant  $\chi^2$  test indicates that the hypothetical model cannot be rejected. When covariates and major gene effects are considered simultaneously, the number of potential hypothesis tests is large and the natural hierarchy of the models may not be clear. In this case, models can also be compared by using Akaike's information criterion (AIC) (Sun et al., 2010b), which is defined as  $\pm 2 \times$  (number of parameters estimated). The one with the lowest AIC value is the model with a better fit.

## **Results**

In Xin-an and Xin-xiang counties, of 472 subjects in 79 pedigrees, 122 had EC. The incidence of EC in the

pedigrees was 25.85%. More males than females were affected by EC in both populations. The ratio of males-to-females among the probands was 1.84. The EC patients ranged from 30 to 89 yr old and the peak age category was 50-59 yr old. The average age of EC onset (at diagnosis) was 58.08±11.82 yr old (Table 1). Third-degree relatives were not analyzed due to insufficient availability of data as most of the individuals were younger and did not reach the age of disease onset or were older and had a high proportion of missing information. The mean age of disease onset in probands with EC with a family history of the disease was 53.17±2.11 yr. The mean age of those without a family history was 58.65±2.29 yr.

#### Family aggregation of EC

The  $\chi^2$  test of goodness-of-fit for binomial distribution was used to evaluate the familial aggregation of EC. A comparison of observed and expected cases of EC is shown in Table 2. Among the 79 families, 34 were found to contain non-affected individuals; 16 had one member with EC; 19 with two EC patients; and 10 with three or above affected members were also identified. The binomial distribution analysis showed a significant

**Table 1. Descriptive of the Esophageal Cancer Patients**

Age	No. of Male (%)	No. of Female (%)	No. of Total (%)
30~	6(7.6)	2(4.7)	8(6.6)
40~	15(19.0)	10(23.3)	25(20.5)
50~	28(35.4)	19(44.2)	47(38.5)
60~	17(21.5)	9(20.9)	26(21.3)
70~	7(8.9)	3(7.0)	10(8.2)
80~	6(7.6)	0(0.0)	6(4.9)
Total	79(100)	43(100)	122(100)

**Table 2. Observed and Expected Distribution of EC in Families of Probands**

No. of relatives affected*	No. of families observed (A)	Observed Prevalence (p)	Expected counts(T)	(A-T) 2/T	$\chi^2$	P
0	34	0.0147	25.42	2.9	10.46	0.002
1	16		9.51	4.43		
2	19		4.89	1.13		
≥3	10		8.02	0.49		
Total	79			8.95		

\*Excluded the probands

**Table 3. Heritability of Esophageal Cancer in Relatives**

Groups	Prevalence (%)	X	a	b	Vb×1000	$h^2 \pm 1.96S(h^2)$ (%)
First-degree relatives	9.09	1.4	1.85	0.34	3.94	67.02±7.31
Second -degree relatives	2.83	1.89	2.27	0.22	6.6	43.08±9.80
Weighted average hereditary						53.16±6.74
General population*	1.16	2.49	2.78			

\*X: susceptibility of population; a: susceptibility of the pedigrees; b: the regression coefficient between the proband and relatives; V<sub>b</sub>: sample variability. \*: China Han population

**Table 5. Complex Segregation Analysis Results of Esophageal Cancer Using the S.A.G.E. SEGREG Program**

Genetic model	qa	$\tau_{aa}$	$\tau_{ab}$	$\tau_{bb}$	$\beta_{aa}$	$\beta_{ab}$	$\beta_{bb}$	-2 lnL	AIC	$\chi^2$	df	P
Dominant	0.32	1	0.5	0	4.83	8.31	-8.19	314.6	328.6	10.27	4	0.02
Recessive	0.02	1	0.5	0	8.35	8.35	-4.25	317.9	329.9	15.33	5	0.04
Additive	0.37	1	0.5	0	11.23	-8.24	-8.24	314.62	320.33	9.19	5	0.07
Major gene	0.02	1	0.5	0	25.23	10.24	-4.25	317.9	330.89	14.79	5	0.01
Non- transmitted	Fixed	N	N	N	-3.12	-3.12	-3.12	320.09	328.01	16.61	7	0.02
Environment	0.1	0.1	0.1	0.1	-3.12	-3.12	-3.12	320.09	328.01	17.65	7	0.03
General	0.09	0.13	0	0	-2.72	21.81	-7.14	304.42	325.21			

\*qa: frequency of a; : transmission probability (for a general model, this is arbitrary);  $\beta_{aa}$ ,  $\beta_{ab}$ ,  $\beta_{bb}$ : the baseline parameter for a subpopulation with the aa genotype; lnL: natural logarithm of likelihood; AIC: Akaike information criterion; df: degrees of freedom; N, there is no transmission probability for an environmental model

deviation from the expected distribution ( $\chi^2=10.46$ , df=1,  $p=0.002$ ). When comparing theoretical cases with actual cases, familial aggregation in the occurrence of EC in relatives of probands was found.

#### Analysis of heritability

According to Falconer's method, the heritability ( $h^2$ ) of EC in first-degree and second-degree relatives of probands was 67.02±7.31% and 43.08±9.80%, respectively (Table 3). The prevalence rate in first-degree relatives was higher than that of second-degree relatives. The weighted average heritability of first and second degree relatives was 53.16±6.74%. A clear heritability tendency was observed in EC pathogenesis.

#### Parameter estimate in the segregation analysis

The Li-Mantel method was used to estimate the segregation ratio of EC. The distribution of probands and their affected siblings among the 79 families is shown in Table 4. The segregation proportion PLM was 0.045.

#### Complex segregation analysis using the S.A.G.E. SEGREG program

**Table 4. Distribution of Families by Size and Number of Affected Cases**

No. of Siblings by No. of Families	No. of Families	No. of Siblings	No. of Affected Siblings per Family	No. of Families by exactly one present
1	15	15	2	2
2	7	14	3	2
3	7	21	9	3
4	16	64	5	0
5	10	50	1	1
6	13	78	1	4
7	7	49	3	1
8	2	16	2	0
9	1	9	1	0
10	1	10	0	0
Total	79	326(T)	27(R)	13(J)

\*T: the total number of siblings; R: the total number of affected siblings per family; J: the total number of families by exactly one present

The complex segregation analysis results are summarized in Table 5. In this analysis, the SEGREG program of S.A.G.E. was used to fit several genetic models, including Mendelian inheritance (major gene, dominant, recessive, or co-dominant) and non-Mendelian inheritance (non-transmitted, environmental, or general effects). The parameters in each model were estimated using maximum likelihood methods. The likelihood of each nested model was compared with that of a general model via a likelihood ratio test. Models with likelihoods that were not statistically worse than the likelihood of the general model were considered to fit the data better than models that were statistically different from the general model. AIC identified the model with the most parsimonious fit to the data. In the 79 pedigrees of EC, the Mendelian inheritance types major gene, dominant, and recessive were rejected, while Mendelian additive inheritance (co-dominant) hypotheses were accepted ( $\chi^2=5$ ,  $df=9.19$ ,  $p=0.07$ ). Based on the AIC value (320.33), the best fit was the additive model of transmission. This implies that EC might be caused by the additive effect of multiple minigenes, as well as the common family environment. The additive model also suggested that approximately 37% of the population were expected to carry the candidate gene (or gene pattern).

## Discussion

In recent years it has been recognized that there is a familial component associated with the pathogenesis of cancer (Wu et al., 2011), and this phenomenon was observed in our current study of EC. We conducted a genetic epidemiology study for populations in Xin-an and Xin-xiang counties of Henan Province in China, and also found that the age of onset for EC was consistent with that previously reported (Layke and Lopez, 2006); the peak age of disease onset was 50-59 years and the mean age of onset was  $58.08 \pm 11.82$  yr. An unequal gender distribution (Feng et al., 2014; Tang et al., 2014) (1.84 males to every female) was also observed in this study.

Earlier studies have described an association between family history and the age of EC onset (Demeester, 2009). In the present study, the age at disease onset in patients with a family history of EC was significantly earlier than that of patients without a family history of the disease ( $53.17 \pm 2.11$  yr compared with  $58.65 \pm 2.29$  yr). Our findings showed that a family history of EC has a potential influence on the age and frequency of disease onset, which suggests that genetic factors are important in the development of this cancer.

In the present study we used Falconer's model of susceptible threshold value and found a significant familial aggregation of EC. A significant increased risk of EC was detected in both first-degree ( $h^2=67.02 \pm 7.31\%$ ) and second-degree relatives ( $h^2=43.08 \pm 9.80\%$ ). The weighted average heritability of EC was  $53.16 \pm 6.74\%$ . The heritability of EC in first-degree relatives was similar to that found by Zhang et al. (2000). Based on the heritability of the disease in first-degree and second-degree relatives, it seems that this susceptibility was due to a multifactorial inheritance. However, since heritability was less than 70%,

environmental factors such as diet may also be involved in the etiology of EC (Sun et al., 2010a; Dawsey et al., 2014; Wang et al., 2014).

The heritability of EC is different in high-incidence areas of China. One possible reason for these reported differences might be that the researchers used a different estimation method. For example, some researchers have used the incidence rate of a region as a basis of comparison, while others used the incidence rate of the general national population. Also, the gender structure of probands or differences in family environment may affect the degree of genetic influence.

Data of the current study were consistent with a multigene mode of inheritance in the simple segregation analysis procedure, since the segregation ratio interval estimates included the value  $PLM=0.045$ , which was significantly lower than 0.25, considered by Li and Mantel as the theoretical maximum likelihood estimate (Li and Mantel, 1968). Genetic factors may be polygenic, oligogenic, or Mendelian with any mode of inheritance, whereas environmental factors may include both familial and non-familial variables. These simple approaches for performing segregation analyses often encounter difficulties when several hypotheses are more or less likely to occur. Development of advanced methods of analysis may lead to a better understanding of the etiology of EC. However, a similar formal analysis had not been performed previously in this population. Therefore, a complex segregation analysis may provide useful information related to the mode of inheritance in this area of China.

In our present study, we performed a complex segregation analysis using the SEGREG program. Complex segregation analysis is a powerful tool to detect gene variation. It has been developed to reduce the restrictions placed on assumptions that are made for the model tested. Our results suggest that the genetic model of EC does not fit a single-gene recessive or dominant mode of inheritance. The Mendelian hypothesis, environment, and non-major-gene (sporadic) models of inheritance were all rejected. The inheritance of EC follows the polygenetic additive model and suggests that genetics are an important factor in the pathogenesis of EC. In addition, the results also confirm that the prevalence rate in relatives of probands is higher than those in the general population. However, studies from regions with a high incidence rate of EC in China found an autosomal recessive mode of inheritance (Zhang et al., 2000). The subjects in these different studies were different and the race, age, and gender ratios also varied. However the analysis of heritability and the complex segregation analysis are only suitable for a specific population, environment and time interval. When any of these factors change, the results are altered.

Although our hypothesis provides the best fit genetic model to the data, segregation analysis has limitations. In particular, segregation analysis cannot distinguish the effect of a single locus behind a trait from the effect of two or more independently acting loci. Several studies have searched for EC-susceptibility genes (Dong et al., 2008), but by using segregation analysis it is difficult to determine



which factor is more important for the pathogenesis of EC and how the genes affect the etiology of the disease.

In summary, our study revealed that a family history of EC is associated with an increased risk of EC occurrence among the offspring. We found that familial aggregation applies to EC, and genetic factors have a certain role in the tumorigenesis of the disease. Simple and complex segregation analyses showed that the genetic model of EC etiology in China may be Mendelian multi-gene additive. This study provides modeling parameters of EC for further linkage studies. Future molecular epidemiologic studies may be performed in this population to determine the specific molecular abnormalities associated with EC that tend to lead to the familial disorder, and how these lesions are inherited.

## Acknowledgements

This research is supported by the National Natural Science Foundation of China (No. 81373097), Major Program of Chinese Ministry of Health and Henan Provincial Medical Science Foundation [2013].

## References

- Carter CL, Hu N, Wu M, et al (1992). Segregation analysis of esophageal cancer in 221 high-risk Chinese families. *J Natl Cancer Inst*, **84**, 771-6.
- Chen W, He Y, Zheng R, et al (2013). Esophageal cancer incidence and mortality in China, 2009. *J Thorac Dis*, **5**, 19-26.
- Dawsey SM, Fagundes RB, Jacobson BC, et al (2014). Diet and esophageal disease. *Ann N Y Acad Sci*, **1325**, 127-37.
- Demeester SR (2009). Epidemiology and biology of esophageal cancer. *Gastrointest Cancer Res*, **3**, 2-5.
- Dong LM, Potter JD, White E, et al (2008). Genetic susceptibility to cancer: the role of polymorphisms in candidate genes. *JAMA*, **299**, 2423-36.
- Douglas S. Falconer 1981. Introduction to Quantitative Genetics . Longman Press.
- Engel LS, Chow WH, Vaughan TL, et al (2003). Population attributable risks of esophageal and gastric cancers. *J Natl Cancer Inst*, **95**, 1404-13.
- Feng XS, Yang YT, Gao SG, et al (2014). Prevalence and age, gender and geographical area distribution of esophageal squamous cell carcinomas in North China from 1985 to 2006. *Asian Pac J Cancer Prev*, **15**, 1981-7.
- Guohong Z, Min S, Duenmei W, et al (2010). Genetic heterogeneity of oesophageal cancer in high-incidence areas of southern and northern China. *PLoS One*, **5**, 9668.
- Hanna GL, Fingerlin TE, Himle JA, et al (2005). Complex segregation analysis of obsessive-compulsive disorder in families with pediatric probands. *Hum Hered*, **60**, 1-9.
- Hu N, Dawsey SM, Wu M, et al (1992). Familial aggregation of oesophageal cancer in Yangcheng County, Shanxi Province, China. *Int J Epidemiol*, **21**, 877-82.
- Kollarova H, Machova L, Horakova D, et al (2007). Epidemiology of esophageal cancer--an overview article. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*, **151**, 17-20.
- Layke JC, Lopez PP (2006). Esophageal cancer: a review and update. *Am Fam Physician*, **73**, 2187-94.
- Lee JH, Barral S, Cheng R, et al (2008). Age-at-onset linkage analysis in Caribbean Hispanics with familial late-onset Alzheimer's disease. *Neurogenetics*, **9**, 51-60.
- Li CC, Mantel N (1968). A simple method of estimating the segregation ratio under complete ascertainment. *Am J Hum Genet*, **20**, 61-81.
- Kanamori, M Schnell AH, Inoue M, et al (2001). Segregation analysis of gastric cancer in a Japanese population. *Int J Hum Genet*, **1**, 263-70.
- Olopade OI, Pichert G (2001). Cancer genetics in oncology practice. *Ann Oncol*, **12**, 895-908.
- Pennathur A, Gibson MK, Jobe BA, et al (2013). Oesophageal carcinoma. *Lancet*, **381**, 400-12.
- Qiao YL, Hou J, Yang L, et al (2001). [The trends and preventive strategies of esophageal cancer in high-risk areas of Taihang Mountains, China]. *Zhongguo Yi Xue Ke Xue Yuan Xue Bao*, **23**, 10-4.
- Stricker C, Fernando RL, Elston RC (1995). Linkage analysis with an alternative formulation for the mixed model of inheritance: the finite polygenic mixed model. *Genetics*, **141**, 1651-6.
- Sun X, Chen W, Chen Z, et al (2010a). Population-based case-control study on risk factors for esophageal cancer in five high-risk areas in China. *Asian Pac J Cancer Prev*, **11**, 1631-6.
- Sun X, Elston R, Barnholtz-Sloan J, et al (2010b). A segregation analysis of Barrett's esophagus and associated adenocarcinomas. *Cancer Epidemiol Biomarkers Prev*, **19**, 666-74.
- Tang WR, Fang JY, Wu KS, et al (2014). Epidemiological characteristics and prediction of esophageal cancer mortality in China from 1991 to 2012. *Asian Pac J Cancer Prev*, **15**, 6929-34.
- Tarone R E (1979). Testing the goodness of fit of the binomial distribution. *Biometrika*, **66**, 585-90.
- Tran GD, Sun XD, Abnet CC, et al (2005). Prospective study of risk factors for esophageal and gastric cancers in the Linxian general population trial cohort in China. *Int J Cancer*, **113**, 456-63.
- Wang AH, Liu Y, Wang B, et al (2014). Epidemiological studies of esophageal cancer in the era of genome-wide association studies. *World J Gastrointest Pathophysiol*, **5**, 335-43.
- Wu M, Zhang ZF, Kampman E, et al (2011). Does family history of cancer modify the effects of lifestyle risk factors on esophageal cancer? A population-based case-control study in China. *Int J Cancer*, **128**, 2147-57.
- Zhang W, Bailey-Wilson JE, Li W, et al (2000). Segregation analysis of esophageal cancer in a moderately high-incidence area of northern China. *Am J Hum Genet*, **67**, 110-9.