Association of the variable number of tandem repeats polymorphism in the promoter region of the *SMYD3* gene with risk of esophageal squamous cell carcinoma in relation to tobacco smoking

Haijian Wang,[1,6] Yang Liu,[1,6] Wen Tan,[2,6] Yang Zhang,[1] Naiqing Zhao,[3] Yan Jiang,[1] Chengzhao Lin,[1] Bingtao Hao,[4] Dan Zhao,[2] Ji Qian,[1] Daru Lu,[1] Li Jin,[1] Qingyi Wei,[5] Dongxin Lin[2,7] and Fuchu He[1,4,7]

[1]MOE Key Laboratory of Contemporary Anthropology, State Key Laboratory of Genetic Engineering, School of Life Sciences and Institutes of Biomedical Sciences, Fudan University, Shanghai 200433; [2]Department of Etiology and Carcinogenesis, Cancer Institute, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021; [3]Department of Biostatistics, School of Public Health, Fudan University, Shanghai 200032; [4]State Key Laboratory of Proteomics, Beijing Proteome Research Center, Beijing Institute of Radiation Medicine, Beijing 102206, China; [5]Department of Epidemiology, The University of Texas M.D. Anderson Cancer Center, Houston, TX 77030, USA

The etiology of esophageal squamous cell carcinoma (ESCC) has been shown to be multifactorial, including genetic, epigenetic, and environmental factors, such as tobacco smoking. A variable number of tandem repeats (VNTR) polymorphism in the promoter region of *SMYD3*, a recently characterized histone lysine methyltransferase gene that is implicated in cell proliferation and carcinogenesis, has been shown to be functional, but its association with cancer risk has not been well established because of apparently discrepant results in different populations. In this case-control study, we genotyped 567 patients with newly diagnosed ESCC and 567 healthy controls and found an increased ESCC risk (odds ratio [OR] = 1.42, 95% confidence interval [CI] = 1.05–1.91) associated with the common *SMYD3* VNTR genotype. Stratification analysis revealed that the increased risk was limited to smokers (OR = 1.99; 95% CI = 1.27–3.12). Furthermore, compared with the reference group of non-smokers carrying the homozygous or heterozygous genotype, ORs (95% CI) of the wild genotype for non-smokers, smokers who smoked <25, and ≥25 pack-years were 1.03 (0.70–1.53), 2.80 (1.66–4.70), and 4.76 (2.67–8.46), respectively (*P* < 0.001 for trend test), suggesting an interaction between this genetic polymorphism and smoking status. These findings provide additional evidence that the common VNTR polymorphism in the promoter region of *SMYD3* gene may be a susceptibility factor for human cancers such as ESCC by interacting with tobacco carcinogens. (*Cancer Sci* 2008; 99: 787–791)

Esophageal cancer is one of the most frequent cancers in the world, with a marked regional variation in incidence and mortality worldwide.[1] In northern China, esophageal squamous cell carcinoma (ESCC) is prevalent, with a five-year survival rate of less than 10%, and it is estimated that there are about 250 000 newly diagnosed ESCC cases and 157 000 deaths each year.[2] Epidemiological investigations have revealed that the etiological factors of ESCC include several environmental risk factors, such as tobacco smoking, heavy alcohol drinking, micronutrient deficiency, and dietary carcinogen exposure.[3,4] However, studies in high-risk areas have also suggested a strong familial aggregation or clustering of cases within families, indicating that genetic susceptibility factors may also play an important role in the development of ESCC.

It is believed that cancer development depends not only on genetic alterations but also on epigenetic changes such as dysfunction of histone modifications, which crucially regulate chromatin functions and gene expression.[5] Mutation screening and expression profiling of histone modifier genes in human solid cancers have suggested that they are important in neoplastic transformation and progression.[6] Although individual predisposition to ESCC has been shown to vary with the genetic status of a variety of critical genes involved in carcinogen metabolism,[7] DNA repair,[8] cell cycle control,[9] and apoptosis,[10,11] few studies have addressed the association between ESCC risk and polymorphisms of genes regulating the modifications of histone and the structural and functional dynamics of chromatin.

Histone lysine methylation is part of the 'histone code', and deregulation of this modification process might cause carcinogenesis.[12] To date, all known lysine histone methyltransferases contain a conserved domain termed the SET (Su(var)3–9, Enhancer of zeste, and Trithorax] domain. The SET and MYND (myeloid translocation protein 8, Nervy, and DEAF1) domain-containing protein 3 (SMYD3), a member of the SMYD family of SET domain-containing proteins, is a recently characterized histone lysine methyltransferase that specifically methylates Lys4 of histone H3 (H3-K4).[13] Evidence has accumulated suggesting that SMYD3 exerts an oncogenic effect by activating transcription of its downstream target genes, and three mechanisms have been proposed for its transactivating function.[13,14] First, SMYD3 recruits RNA polymerase II through its interaction with the RNA helicase HELZ, forming a transcriptional complex to elongate transcription. Second, SMYD3 methylates the H3-K4 nearby the target genes, which alters the chromatin folding, leading to increased accessibility of DNA to proteins that mediate transcription. Third, SMYD3 transactivates a set of genes through the interaction between its MYND-type zinc finger domain, a cysteine-rich structure present in proteins generally implicated in gene regulation and carcinogenesis[15] and the *cis*-regulatory binding motif CCCTCC in the promoter region of target genes including oncogenes, homeobox genes, and genes known to be associated with cell-cycle regulation.[13] For example, it was currently revealed that the telomerase reverse transcriptase gene (*TERT*), which codes the catalytic component of the telomerase complex that is a prerequisite for cell immortalization and transformation, is a direct target of SMYD3.[16]

Recently, a tandem repeat variation [(CCGCC)ₙ] in the promoter region of *SMYD3* has been reported to be associated with increased risks of colorectal cancer, hepatocellular carcinoma (HCC), and breast cancer.[17] This tandem-repeat sequence was shown to be an E2F-1 binding site and the wild-type three-repeat

---

allele has enhanced binding affinity compared to the two-repeat allele.[17] However, inconsistent results also exist which suggest no association between the *SMYD3* VNTR genotypes and risk of familial breast cancer[18] or HCC.[19] These apparent discrepancies strongly suggest that additional studies in different populations with different cancer types are needed to test the hypothesis proclaiming the *SMYD3* VNTR polymorphism as a risk for human cancers.

In this study, we examined whether the *SMYD3* VNTR polymorphism is associated with the risk of developing ESCC in a Chinese population. We observed that the three-repeat allele of the gene may be a susceptibility factor for ESCC by interacting with tobacco smoking.

Materials and Methods

Study subjects. The study population consisted of 567 patients with ESCC and 567 healthy controls. All subjects were unrelated ethnic Han Chinese and residents in Beijing and the surrounding regions. Case patients were recruited between July 1999 and December 2001 at the Cancer Hospital, Chinese Academy of Medical Sciences (Beijing, China). All incident patients with histologically confirmed ESCC were enrolled with a response rate of 94% and there was no sex and age restriction. The exclusion criteria included previous cancer and previous chemotherapy or radiotherapy. Most case patients were participants in a molecular epidemiological study of esophageal cancer described previously.[8] Control subjects were cancer-free individuals and randomly selected from a nutritional survey database conducted in the same regions during the same time period of case collection, with a response rate of 89%. The characteristics of control subjects were described previously.[8] The selection criteria for control subjects included no individual history of cancer and frequency matching to case patients on sex and age. At recruitment, informed consent was obtained from each subject and personal data from each participant regarding demographic characteristics such as sex and age and related risk factors including tobacco smoking were collected by questionnaire. This study was approved by the institutional review board of the Chinese Academy of Medical Sciences Cancer Hospital and Institute.

SMYD3 VNTR analysis. Genotyping of the VNTR polymorphism in the promoter region of *SMYD3* was performed by direct sequencing of polymerase chain reaction (PCR) products containing the VNTR site, with the same protocol and primers of the same sequences as those used in the two previous studies:[17,18] forward: 5′-CGC CTG TCT TCT GCG CAG TCG-3′; reverse: 5′-CAC CTT CAG CGG CTC CAT CCT C-3. PCR products were sequenced with ABI PRISM Dye Terminator Sequencing Kits (Applied Biosystems, Foster City, CA, USA) and loaded onto an ABI 3730XL sequencer.

Statistical analysis. The deviations of genotype frequencies in the control subjects and in the case patients from those expected under Hardy–Weinberg equilibrium were assessed by using the standard χ^2-test. The χ^2-test was also used to examine differences in demographic variables, smoking, and distribution of genotypes between cases and controls. The associations between genotype and risk of ESCC were estimated by calculating odds ratio (OR) and their 95% confidence interval (95% CI) with unconditional logistic regression models. The ORs were adjusted for age, sex, and pack-years smoked. Smokers were considered current smokers if they smoked up to one year before the date of cancer diagnosis or if they smoked up to one year before the date of the interview for control subjects. Information was collected on the number of cigarettes smoked per day, the age at which the subjects started smoking, and the age at which the smokers stopped smoking. Subjects who had never smoked or smoked for a period of less than one year before the date of cancer

Table 1. Distributions of select characteristics by case-control status

Variable	Controls (n = 567)	Cases (n = 567)	P
Sex (%)			0.525[†]
Male	379 (66.8)	389 (68.6)	
Female	188 (33.2)	178 (31.4)	
Mean age in years (SD)	56.7 (9.5)	57.0 (7.9)	0.560[‡]
Smoking status (%)			<0.001[†]
Never	369 (65.1)	256 (45.1)[†]	
Ever	198 (34.9)	311 (54.9)[†]	
<25 pack-year	112 (56.6)	142 (45.7)[‡]	0.016[†]
≥25 pack-year	86 (43.4)	169 (54.3)[‡]	
Mean pack-years (SD)	25.0 (16.4)	28.9 (19.7)	0.020[‡]

[†]Pearson χ^2 test for the difference between cases and controls.
[‡]Student *t*-test with unequal variances and Satterthwaite's degrees of freedom for the difference between cases and controls.

diagnosis for case patients or the date of interview for control subjects were defined as non-smokers. The light and heavy smokers were categorized by using the 50% ile pack-year [(cigarettes per day/20) × (years smoked)] values of the controls as the cut-off points (i.e. <25 and ≥25 pack-years). The probability level of <0.05 was used as the criterion of statistical significance, and all statistical tests were two-sided. We tested the null hypothesis of additive and multiplicative gene–environment interaction between the *SMYD3* VNTR polymorphism and tobacco smoking, and evaluated departures from additive and multiplicative interaction models by including main effect variables and their product terms in the logistic regression model.[20] The homogeneity test was done to compare the difference between smoking-related ORs among different genotypes or between the product of related ORs and the joint effect OR. All statistical analyses were performed by using Statistical Analysis System software (version 8.0; SAS Institute, Cary, NC, USA).

Results

The distributions of age, sex, and smoking status among the subjects are summarized in Table 1. There were no significant differences between case patients and control subjects in terms of the distributions of sex and age, suggesting that the frequency matching was adequate. However, more smokers were presented in the case group compared with controls (54.9% *versus* 34.9%; χ^2 = 45.5, P < 0.001), which is consistent with findings showing that smoking is one major risk factor for ESCC.[3,4]

The allelic frequency and genotype distributions of the *SMYD3* promoter VNTR polymorphism in cases and controls are shown in Table 2. The allelic frequencies for the three tandem repeats (denoted as *3*) and the two tandem repeats (denoted as *2*) were 88.0% and 12.0% among controls, and 90.4% and 9.6% among cases, respectively, with the three tandem repeat allele being slightly more prevalent in cases than in controls. The distributions of *SMYD3* VNTR genotypes in cases and controls were consistent with that predicted by the Hardy–Weinberg equilibrium. The overall distribution of genotypes in cases differed significantly from that in controls (χ^2 = 5.55, degrees of freedom [d.f.] = 2, P < 0.05), and the frequency of the hetero-zygotes (*3/2*) was significantly lower in cases than in controls (16.0% *versus* 22.6%; χ^2 = 7.36, d.f. = 1, P < 0.01). Because the *2/2* genotype was rare, it was combined with the *3/2* genotype as the reference group for further analysis, as used in previous association studies of the *SMYD3* VNTR polymorphism.[17] Multivariate logistic regression analysis showed that the *SMYD3* *3/3* genotype was associated with significantly increased risk for the development of ESCC (OR 1.42; 95% CI 1.05–1.91) compared with the *3/2* or *2/2* genotype (Table 2).

Table 2. Genotype distribution of the *SMYD3* VNTR polymorphism between cases and controls and their association with risk of ESCC

Genotype	Controls (n = 567) n (%)	Cases (n = 567) n (%)
SMYD3 VNTR		
3/3	435 (76.7)	467 (82.4)
3/2	128 (22.6)	91 (16.0)
2/2	4 (0.7)	9 (1.6)
2 allele frequency	0.120	0.096
Adjusted OR (95% CI)[†]		1.42 (1.05–1.91)
	χ^2 with 2 degree of freedom = 5.55 and *P*-value = 0.022	

[†]Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated by logistic regression with the *3/2* and *2/2* genotypes as a reference group and adjusted for age, sex and pack-years within the strata. ESCC, esophageal squamous cell carcinoma; VNTR, variable number of tandem repeats.

Table 3. Risk of ESCC related to *SMYD3* VNTR genotypes by sex, age, and smoking status

Characters	Genotype of *SMYD3* VNTR (number of cases/controls)				
	3/2 + 2/2	OR (95% CI)[†]	3/3	OR (95% CI)[†]	P
Gender					
Female	36/41	1.00 (ref)	142/147	1.10 (0.66–1.82)	0.722
Male	64/91	1.00 (ref)	325/288	1.62 (1.12–2.34)	0.011
Age					
<55 years-old	38/55	1.00 (ref)	162/196	1.04 (0.64–1.70)	0.878
≥55 years-old	62/77	1.00 (ref)	305/239	1.71 (1.18–2.50)*	0.005
Smoking status					
Non-smoker	55/82	1.00 (ref)	201/287	1.03 (0.70–1.53)	0.874
Smoker	45/50	1.00 (ref)	266/148	1.99 (1.27–3.12)**	0.003

[†]Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated by unconditional logistic regression with the (*3/2* + *2/2*) genotypes as a reference group and adjusted for age, sex and pack-years within the strata.
*$P < 0.05$, test for homogeneity between age-related ORs among *3/2* + *2/2* and *3/3* genotype.
**$P < 0.05$, test for homogeneity between smoking-related ORs among *3/2* + *2/2* and *3/3* genotype.
ESCC, esophageal squamous cell carcinoma; VNTR, variable number of tandem repeats.

The associations of *SMYD3* VNTR genotypes with risk of ESCC were further examined with stratification by age, sex, and smoking status (Table 3). It was shown that the association between the *3/3* genotype and increased risk of ESCC was more pronounced in males (OR 1.62; 95% CI 1.12–2.34, *P* = 0.011)

and in subjects aged ≥55 years (OR 1.71; 95% CI 1.18–2.05, *P* = 0.005). Significant association of the *3/3* wild genotype with risk of the cancer was observed for male older smokers as compared with the reference group of female younger non-smokers (data not shown). As to the stratum of smoking status, among non-smokers, the *SMYD3 3/3* genotype was not associated with elevated risk of ESCC (OR 1.03; 95% CI 0.70–1.53); however, among smokers, the association was significant and substantial (OR 1.99; 95% CI 1.27–3.12; *P* = 0.046, test for homogeneity), suggesting that the *SMYD3* VNTR polymorphism is a susceptibility factor for smoking-related ESCC.

Because tobacco smoking is a well-established risk factor for ESCC and because the risk associated with the *SMYD3 3/3* genotype was more pronounced in smokers, we therefore examined whether an interaction existed between the *SMYD3* VNTR polymorphism and smoking status (Table 4). Compared with non-smokers carrying the *3/2* or *2/2* genotypes, the risk for the presence of both smoking and the *3/3* genotype (OR = 3.39; 95% CI = 2.09–5.49) was greater than the product of the OR for smoking (OR = 1.78; 95% CI = 0.95–3.31) and the OR for the *3/3* genotype (OR = 1.03; 95% CI = 0.70–1.53), with marginal statistical significance in the test for interaction (*P* = 0.087). These data suggested a potential multiplicative joint effect between smoking and the *SMYD3* VNTR polymorphism. Moreover, when the risk associated with the *3/3* genotype was further stratified by smoking levels (<25 or ≥25 pack-years smoked), the risk increased consistently with cumulative smoking dose among smokers, in which a multiplicative interaction between the susceptible genotype and categories of pack-year smoked was also observed (*P* = 0.07, test for interaction). Compared with the *3/2* and *2/2* genotype carriers who were non-smokers, the OR for *3/3* genotype carriers who did not smoke, smoked <25, and ≥25 pack-year were 1.03 (95% CI = 0.70–1.53), 2.80 (95% CI = 1.66–4.70), and 4.76 (95% CI = 2.67–8.46), respectively (*P* < 0.001 for trend test). Thus, these data suggest an interaction between the *SMYD3* VNTR polymorphism and smoking in the etiology of ESCC in this study population.

Discussion

In the present study, we examined whether the VNTR polymorphism in the promoter region of the *SMYD3* gene has an impact on risk for developing ESCC. On the basis of analysis of 567 patients and 567 controls in a Chinese population, we observed a significant association between this genetic polymorphism and risk of the cancer. Subjects having the *SMYD3 3/3* VNTR genotype were at a substantially increased risk of ESCC in relation to tobacco smoking.

Our results in the present study are consistent with previous findings of molecular assays and population association studies and therefore are biologically plausible. The tandem-repeat

Table 4. Risk of ESCC related to *SMYD3* VNTR genotypes by smoking status and pack-years smoked

Characters	Genotype of *SMYD3* VNTR (number of cases/controls)					
	3/2 + 2/2	OR (95% CI)[†]	P	3/3	OR (95% CI)[†]	P
Smoking status						
Non-smoker	55/82	1.00 (ref)		201/287	1.03 (0.70–1.53)	0.874
Smoker	45/50	1.78 (0.95–3.31)	0.070	266/148	3.39 (2.09–5.49)	<0.001
<25 pack-year	19/29	1.26 (0.60–2.66)	0.540	123/84	2.80 (1.66–4.70)	<0.001
≥25 pack-year	26/21	2.66 (1.22–5.80)	0.014	143/64	4.76 (2.67–8.46)[‡]	<0.001

[†]The odds ratios (ORs) and 95% confidence intervals (CIs) for groups with different smoking level were built by unconditional logistic regression in comparison with the reference group of non-smokers carrying the (*3/2* + *2/2*) genotypes and adjusted for age and sex.
[‡]$P < 0.05$, test for homogeneity between smoking-related ORs among *3/2* + *2/2* and *3/3* genotype.
ESCC, esophageal squamous cell carcinoma; VNTR, variable number of tandem repeats.

sequence of CCGCC motif in the *SMYD3* promoter region is a binding site for E2F1, a crucial transcriptional factor in cell-cycle control, DNA synthesis, DNA repair, replication, and apoptosis through transcriptional activation of a number of downstream genes.[17] The major allele of three repeats motif (corresponding to the high-risk allele) have enhanced binding affinity to E2F-1 and increased transactivation capacity compared to the two repeats motif (the low-risk allele).[17] It has been reported that augmented expression of *SMYD3* promotes growth of cancer cells through its histone methyltransferase activity, and genetic knockdown with small interfering RNAs results in significant growth suppression in cancer cells.[13] Enhanced cell growth in human colorectal, HCC, and mammary carcinomas by up-regulation of SMYD3 has been observed[13,21] which is consistent with the results of population studies showing higher risk of the *3/3* genotype for these types of cancer.[17] In addition, it has been shown that other histone methyltransferases, such as SUV39H1, SUV39H2, and EZH2, display differential expression profiles between various cancers and normal tissues[6] and genetic polymorphisms in *SUV39H2* and *EZH2* are, respectively, associated with risk of the development of lung cancer[22] and prostate cancers.[23]

In the present study, we observed a more than three-fold increased ESCC risk associated with the *SMYD3 3/3* genotype among smokers but not non-smokers, suggesting a possible gene–environment interaction between the *SMYD3* polymorphism and smoking in the etiology of ESCC in this Chinese population. While numerous studies have demonstrated genetoxic effects of cigarette smoke leading to carcinogenesis, little is known about molecular events and the nature that drives possible epigenetic effect in tumorigenesis by tobacco carcinogens. However, the effect of tobacco smoke on epigenetic silencing of tumor-suppressor genes *p16(INK4a)* in tumorigenesis via hypermethylation of CpG islands in the gene promoters has been documented.[23] Our study showing *SMYD3 3/3* genotype as a risk factor for ESCC among smokers suggested that cigarette smoke may affect not only the methylation status of the regulatory DNA sequence of some tumor-suppressor genes, but also the methylation status of histone through interaction with the function of histone modifier enzymes such as SMYD3. It has also been documented that tobacco smoke causes abnormal inflammatory response in the lungs through disrupting the balance of histone acetylation:deacetylation by increasing histone acetyltransferase and decreasing histone deacetylase activity, which is known to play an important role in transcriptional regulation of proinflammatory genes.[24–26] But to our knowledge, no study on smoking and histone methylation status has been reported. Therefore, further functional studies are warranted to elucidate the relationship between tobacco smoking and histone methylation in carcinogenesis. Furthermore, because tobacco smoking is the leading preventable cause of cancer and because the cancer-prone genotype of the *SMYD3* VNTR polymorphism is very prevalent in humans, our findings also have implications for the prevention of tobacco smoking–related cancers.

The present study may have some limitations because it is a hospital-based, case-control study with a relatively small sample size. However, we used incident cases and the procedures for the ascertainment of cases and controls were well defined.[8] The fact that genotype frequencies among controls and cases fit the Hardy–Weinberg law further supports the randomness of subject selection. Furthermore, both patients and controls were subjects in our previous comprehensive association studies of ESCC and genetic polymorphisms in the DNA repair pathway, with most of the findings replicated in follow-up studies in other ethnic populations or in other cancer types.[8,27–29] Therefore, it is at least possible that ascertainment bias and confounding factors could account for the association of *SMYD3* VNTR polymorphism and risk of ESCC. In addition to tobacco smoking, other environmental factors such as alcohol consumption have been suggested to be associated with risk of ESCC, and this might also modulate the effect of *SMYD3* VNTR genotype. Unfortunately, information on exposure other than smoking is not available in the present study, which prevents more comprehensive evaluation of the role of gene–environment interaction in ESCC development.

In conclusion, the present case-control study in a Chinese population demonstrates a significant association between the *SMYD3* promoter VNTR polymorphism and risk of ESCC, and a suggestive gene–environment interaction between this polymorphism and tobacco smoking in the etiology of cancer. These molecular epidemiological findings provide additional evidence that the common VNTR polymorphism in *SMYD3* may be a susceptibility factor for human cancer.

Acknowledgments

References

1 Pisani P, Parkin DM, Ferlay J. Estimates of the worldwide mortality from eighteen major cancers in 1985. Implications for prevention and projections of future burden. *Int J Cancer* 1993; **55**: 891–903.

2 Yang L, Parkin DM, Li L, Chen Y. Time trends in cancer mortality in China: 1987–99. *Int J Cancer* 2003; **106**: 771–83.

3 Gao YT, McLaughlin JK, Blot WJ *et al*. Risk factors for esophageal cancer in Shanghai, China. I. Role of cigarette smoking and alcohol drinking. *Int J Cancer* 1994; **58**: 192–6.

4 Engel LS, Chow WH, Vaughan TL *et al*. Population attributable risks of esophageal and gastric cancers. *J Natl Cancer Inst* 2003; **95**: 1404–13.

5 Gibbons RJ. Histone modifying and chromatin remodelling enzymes in cancer and dysplastic syndromes. *Hum Mol Genet*.: 14 *Spec*, 2005; **1**: R85–92.

6 Ozdag H, Teschendorff AE, Ahmed AA *et al*. Differential expression of selected histone modifier genes in human solid cancers. *BMC Genomics* 2006; **7**: 90.

7 Tan W, Song N, Wang GQ *et al*. Impact of genetic polymorphisms in cytochrome P450 2E1 and glutathione S-transferases M1, T1, and P1 on susceptibility to esophageal cancer among high-risk individuals in China. *Cancer Epidemiol Biomarkers Prev* 2000; **9**: 551–6.

8 Hao B, Wang H, Zhou K *et al*. Identification of genetic variants in base excision repair pathway and their associations with risk of esophageal squamous cell carcinoma. *Cancer Res* 2004; **64**: 4378–84.

9 Smeds J, Berggren P, Ma X, Xu Z, Hemminki K, Kumar R. Genetic status of cell cycle regulators in squamous cell carcinoma of the oesophagus. the CDKN2A (p16 (INK4a) and p14 (ARF) and p53 genes are major targets for inactivation. *Carcinogenesis* 2002; **23**: 645–55.

10 Sun T, Miao X, Zhang X, Tan W, Xiong P, Lin D. Polymorphisms of death pathway genes FAS and FASL in esophageal squamous-cell carcinoma. *J Natl Cancer Inst* 2004; **96**: 1030–6.

11 Sun T, Gao Y, Tan W *et al*. A six-nucleotide insertion-deletion polymorphism in the CASP8 promoter is associated with susceptibility to multiple cancers. *Nat Genet* 2007; **39**: 605–13.

12 Dillon SC, Zhang X, Trievel RC, Cheng X. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol* 2005; **6**: 227.

13 Hamamoto R, Furukawa Y, Morita M *et al*. SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat Cell Biol* 2004; **6**: 731–40.

14 Sims RJ 3rd, Reinberg D. From chromatin to cancer: a new histone lysine methyltransferase enters the mix. *Nat Cell Biol* 2004; **6**: 685–7.

15 Spadaccini R, Perrin H, Bottomley MJ, Ansieau S, Sattler M. Structure and functional analysis of the MYND domain. *J Mol Biol* 2006; **358**: 498–508.

16 Liu C, Fang X, Ge Z *et al*. The telomerase reverse transcriptase (hTERT) gene is a direct target of the histone methyltransferase SMYD3. *Cancer Res* 2007; **67**: 2626–31.

17 Tsuge M, Hamamoto R, Silva FP *et al*. A variable number of tandem repeats

polymorphism in an E2F-1 binding element in the 5′ flanking region of SMYD3 is a risk factor for human cancers. *Nat Genet* 2005; **37**: 1104−7.

18 Frank B, Hemminki K, Wappenschmidt B *et al*. Variable number of tandem repeats polymorphism in the SMYD3 promoter region and the risk of familial breast cancer. *Int J Cancer* 2006; **118**: 2917−8.

19 Wang XQ, Miao X, Cai Q, Garcia-Barcelo MM, Fan ST. SMYD3 tandem repeats polymorphism is not associated with the occurrence and metastasis of hepatocellular carcinoma in a Chinese population. *Exp Oncol* 2007; **29**: 71−3.

20 Brennan P. Gene–environment interaction and aetiology of cancer: what does it mean and how can we measure it? *Carcinogenesis* 2002; **23**: 381−7.

21 Hamamoto R, Silva FP, Tsuge M *et al*. Enhanced SMYD3 expression is essential for the growth of breast cancer cells. *Cancer Sci* 2006; **97**: 113−8.

22 Yoon KA, Hwangbo B, Kim IJ *et al*. Novel polymorphisms in the SUV39H2 histone methyltransferase and the risk of lung cancer. *Carcinogenesis* 2006; **27**: 2217−22.

23 Bachmann N, Hoegel J, Haeusler J *et al*. Mutation screen and association study of EZH2 as a susceptibility gene for aggressive prostate cancer. *Prostate* 2005; **65**: 252−9.

24 Marsit CJ, Karagas MR, Danaee H *et al*. Carcinogen exposure and gene promoter hypermethylation in bladder cancer. *Carcinogenesis* 2006; **27**: 112−6.

25 Kim DH, Nelson HH, Wiencke JK *et al*. p16 (INK4a) and histology-specific methylation of CpG islands by exposure to tobacco smoke in non-small cell lung cancer. *Cancer Res* 2001; **61**: 3419−24.

26 Kraunz KS, Nelson HH, Lemos M, Godleski JJ, Wiencke JK, Kelsey KT. Homozygous deletion of p16 (INK4a) and tobacco carcinogen exposure in nonsmall cell lung cancer. *Int J Cancer* 2006; **118**: 1364−9.

27 Lockett KL, Hall MC, Xu J *et al*. The ADPRT V762A genetic variant contributes to prostate cancer susceptibility and deficient enzyme function. *Cancer Res* 2004; **64**: 6344−8.

28 Zhang X, Miao X, Liang G *et al*. Polymorphisms in DNA base excision repair genes ADPRT and XRCC1 and risk of lung cancer. *Cancer Res*, 2005; **65**: 722−6.

29 Hung RJ, Brennan P, Canzian F *et al*. Large-scale investigation of base excision repair genetic polymorphisms and lung cancer risk in a multicenter study. *J Natl Cancer Inst* 2005; **97**: 567−76.