Int J Mol Epidemiol Genet 2010;1(1):11-18
www.ijmeg.org/IJMEG906001

Original Article

Candidate variants at 6p21.33 and 6p22.1 and risk of non-small cell lung cancer in a Chinese population

Mingfeng Zhang[1], Lingmin Hu[1], Hao Shen[1], Jing Dong[1], Yongqian Shu[1,2], Lin Xu[3], Guangfu Jin[1], Tian Tian[1], Zhibin Hu[1], Hongbing Shen[1,2]

[1]Department of Epidemiology and Biostatistics, Cancer Center, Nanjing Medical University, Nanjing, China; [2]Department of Oncology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China; [3]Department of Thoracic Surgery, Jiangsu Cancer Hospital, Nanjing, China;

Received June 17, 2009; accepted July 20, 2009; available online July 26, 2009

Abstract: Chromosome 6p21.33, containing *BAT3* and *MSH5* genes, together with chromosome 6p22.1 were recently identified as susceptible regions for lung cancer in Caucasian populations. These findings interest us in assessing whether genetic variants in these regions also contribute to lung cancer risk in Chinese populations. We genotyped the most significant single nucleotide polymorphism (SNP) (rs9295740) reported in Caucasian populations at Chromosome 6p22.1 and one common potentially functional variant (rs2075789) located at exon 2 of *MSH5* in a case-control study including 1009 histologically confirmed non-small cell lung cancer (NSCLC) cases and 1127 cancer-free controls in a Chinese population. We found that the distributions of genotypes of both SNPs between cases and controls were not significantly different (P = 0.624 for rs9295740 and P = 0.937 for rs2075789). Logistic regression analyses revealed neither of the two SNPs was significantly associated with altered risk of NSCLC in dominant or recessive genetic models. When we compared the combined variant genotypes (GA+AA) with the common homozygote GG, assuming a dominant genetic model, the adjusted ORs were 1.03 (95% CI = 0.86-1.25) for rs9295740 and 1.03 (95% CI = 0.85-1.25) for rs2075789. In addition, no significant associations were observed in subgroups stratified by age, gender, smoking status or histologic types. Our results indicate that the most significant SNP rs9295740 identified in Caucasians in 6p22.1 and the potentially functional SNP rs2075789 in 6p21.33, seem not applicable to Chinese populations as susceptible markers for lung cancer. Re-sequencing and fine-mapping this region, along with extensive functional evaluations, is required.

Key words: Polymorphism; lung cancer; 6p21.33; 6p22.1; MSH5

Introduction

Lung cancer (#211980) is the most common cancer as well as the leading cause of cancer related deaths worldwide [1], and approximately 80% were non-small cell lung carcinomas (NSCLCs). Although tobacco smoking is the most important cause of lung cancer, less than 20% of the smokers develop lung cancer [2], suggesting that genetic susceptibility plays an important role in the etiology of lung cancer. More and more studies have focused on the heritable component of lung cancer and there is increasing evidence supporting that host genetic factors may modify the risk of lung cancer development [3]. However, the genetic basis of susceptibility to lung cancer, especially for sporadic cases, is still undefined.

The genome-wide association (GWA) approach has been the most powerful and efficient way thus far in identifying genetic variants, especially single nucleotide polymorphisms (SNPs), associated with complex human diseases. This approach enables investigators to identify novel loci or genes for various diseases and quantitative traits. The first wave of large-scale, high-density genome-wide association studies has improved our understanding of the genetic basis of several kinds of diseases, including cancer. Three genome-wide association studies in populations of European ancestry (EA) have mapped

a lung cancer susceptibility locus to chromosome 15q25 containing nicotinic acetylcholine receptor genes (*CHRNA3*, *CHRNA5* and *CHRNB4*) [4-6]. More recently, Wang *et al.* [7] further found strong evidence for a lung cancer susceptible region at 6p21.33. Two highly correlated markers (r^2 = 0.99) for lung cancer risk were identified, rs3117582 localized to intron 1 of *BAT3* and rs3131379 localized to intron 10 of *MSH5*. Besides, a risk locus (rs9295740) at the chromosome 6p22.1, which has extensive linkage disequilibrium (LD) with 6p21.33, was also identified. rs4324798 located at chromosome 6p was also identified as a possible risk locus of lung cancer ($P<5\times10^{-5}$) by Hung *et al.* [4], although not exceeding the genome-wide significance level. These findings interest us in assessing whether these genetic variants identified in Caucasians also contribute to lung cancer risk in Chinese populations. Because SNPs rs3131379 and rs3117582 located at 6p21.33 as well as rs4324798 were not polymorphic in Asians, we therefore genotyped the most significant SNP rs9295740 in 6p22.1 and a potentially functional SNP rs2075789 in 6p21.33, in an ongoing case-control study with 1009 NSCLC cases and 1127 frequency-matched cancer-free controls to evaluate the associations of genetic variants in these regions with lung cancer risk in our Chinese population.

Materials and methods

Study population

This ongoing molecular epidemiological study of lung cancer was approved by the Institutional Review Board of Nanjing Medical University. Details of the study design were described previously [8]. Briefly, all subjects were genetically-unrelated ethnic Han Chinese and were from Nanjing City and surrounding regions in Jiangsu Province of eastern China. The incident lung cancer patients were histopathologically diagnosed and recruited from July 2003 to July 2008 at the Cancer Hospital of Jiangsu province (Nanjing), the First Affiliated Hospital of Nanjing Medical University and the Nanjing Thoracic Hospital, Nanjing, China, without the restrictions of age, sex and histology. As a results, a total of 1009 NSCLC cases were obtained. The 1127 cancer-free controls were randomly selected from a pool of 30,000 individuals participated in a community-based screening program for non-infectious diseases conducted in Jiangsu province during the same time period as the cases were recruited. The control subjects had no history of cancer and were frequency-matched to the cases on age and sex. Each participant was scheduled for an interview after written informed consent was obtained, and a structured questionnaire was administered by trained interviewers to collect information on demographic data and environmental exposure history. After interview, approximately 5 ml venous blood sample was collected from each participant.

Individuals who smoked at least one cigarette per day for over one year were defined as ever smokers, otherwise they were considered as nonsmokers. Cumulative smoking dose was measured by pack-years smoked ([cigarettes per day/20] × smoking years). Light and heavy smokers were categorized according to the 50th percentile of pack-years from the control subjects (i.e. ≤25 pack-years and >25 pack-years).

With such a sample size, the power of our test was 38.04% and 33.87% to detect the allele effect of rs9295740 and rs2075789, respectively, in our population, with an estimated OR of 1.20.

Selections of SNPs

Three SNPs, rs3131379 and rs3117582 located at 6p21.33 and rs9295740 located at 6p22.1, were identified as the most significant loci associated with lung cancer risk at 6p in populations of European ancestry according to the results of the GWAS conducted by Wang *et al*. [7]. rs4324798 located at chromosome 6p was also identified as a possible risk locus of lung cancer by Hung *et al.* [4]. However, by searching the HapMap database (HapMap Data Rel 23a/phase II Mar 08), we found that rs3131379, rs3117582 and rs4324798 were not polymorphic in Asians with minor allele frequencies (MAFs) = 0 in both Chinese Beijing (CHB) and Japanese Tokyo (JPT) populations. Since rs3131379 and 3117582 localized in intron 1 of *BAT3* and intron 10 of *MSH5*, respectively, which were two strong candidate genes for lung cancer susceptibility, we further screened the potentially functional variants (including variants in the promoter, 5'- and 3'-untranslated regions, as well as non-synonymous variants in the coding region) of the *BAT3* and *MSH5* genes in the NCBI dbSNP

(http://www.ncbi. nlm.nih.gov/SNP, build 129) to search for candidate variants with MAF > 0.05 in Chinese populations. Only one SNP (rs2075789 G>A) located at exon 2 of *MSH5* with amino acid change from Pro to Ser was identified. Finally, the SNPs rs9295740 in 6p22.1 and rs2075789 in 6p21.33 were chosen for genotyping in the present study.

Genotyping

Genomic DNA was extracted from a leukocyte pellet by phenol-chloroform method. Both SNPs (rs9295740 and rs2075789) were genotyped using PCR-restriction fragment length polymorphism (RFLP) method. We used a pair of primers of 5'- GGCAGTTCAAATTCTTT GATATTA -3' (sense) and 5'- GAGCACTTAGAG TTCTGGGG -3' (antisense) to amplify a 129-bp PCR product for rs9295740, which was digested by restriction enzyme of *Ase*I (New England BioLabs, Beverly, MA) and separated on a 3% agarose gel. The A allele produces two fragments of 107- and 22-bp while the G allele results in a single 129-bp fragment. For rs2075789, the primers were 5'- CGGCTCC CAACCCTCTTT -3' (sense) and 5'- TTCCGAACCG CCCTCACT -3' (antisense) and the 262-bp PCR fragment was digested by *Bsr*I (New England BioLabs, Beverly, MA). The product with G allele are divided into 134- and 128-bp fragments while the A allele keeps one fragment of 262-bp.

All genotyping analyses were blinded without knowing the case/control status for quality control. To validate the genotyping results, 10% random selected samples were repeated and the concordance rate achieved 100%. The call rates of genotyping reached 98.7% and 99.3% in the cases and the controls, respectively, for both SNPs. The failure of genotyping was due to inadequate quality or quantity of DNA sample. Thus, genotyping of 996 cases and 1119 controls were finally included in the final analyses.

Statistical analyses

Differences in the distributions of demographic variables, selected variables, and frequencies of the genotypes between the cases and controls were evaluated by two-sided χ^2 test. Unconditional logistic regression was used to estimate crude odds ratios (ORs), adjusted ORs and their 95% confidential intervals (CIs), with adjustment for age, sex and smoking, where appropriate. Hardy-Weinberg equilibrium was tested by a goodness-of-fit χ^2 test to compare the observed genotype frequencies to the expected ones among the control subjects. All the statistical analyses were performed with Statistical Analysis System software (v.9.1.3; SAS Institute, Cary, NC). All statistical tests were two-sided and $P < 0.05$ was used as the criterion of statistical significance.

Results

The characteristics of the 1009 NSCLC patients and 1127 controls included in the analysis are summarized in **Table 1**. There were no significant differences in terms of distributions on age (≤ 60 and > 60 years) and sex between the cases and the controls (P = 0.643 and 0.772, respectively), suggesting that our frequency matching of the demographic characteristics was satisfactory. As expected, compared with the control subjects, the cases were more likely to be smokers ($P < 0.001$). Among the smokers, about 68.7% of the NSCLC cases smoked more than 25 pack-years, which was significantly higher than that of the controls (52.2%) ($P < 0.001$). Of the total 1009 NSCLC cases, 588 (58.3%) were classified as adenocarcinoma, 310 (30.7%) as squamous cell carcinoma and 111 (11.0%) as large cell, mixed cell carcinomas or undifferentiated carcinoma (**Table 1**).

The genotype distributions of the selected SNPs, rs9295740 and rs2075789, in the cases and controls are shown in **Table 2**. The observed genotype frequencies for these two polymorphisms were both in agreement with that expected under the Hardy-Weinberg equilibrium among the controls (P = 0.775 for rs9295740, and 0.950 for rs2075789). Genotype frequencies of rs9295740 (GG, GA and AA) in the cases (66.3%, 30.6% and 3.1%) seemed comparable to those in the controls (67.5%, 29.0% and 3.6%) (P = 0.624). For rs2075789, similar distributions of the genotypes (GG, GA and AA) were also observed among the cases (71.6%, 26.4% and 2.0%) and the controls (72.0%, 25.8% and 2.1%) (P = 0.937). Logistic regression analyses revealed that neither SNPs were significantly associated with the altered risk of NSCLC in dominant or recessive genetic models. As shown in **Table 2**, the adjusted ORs (95 CI%) of variant allele (A allele) for rs9295740 and rs2075789 were

Table 1. Distributions of select characteristics among lung cancer cases and cancer-free controls

Characteristics	Cases (n = 1009)		Controls (n = 1127)		*P**
	No.	%	No.	%	
Age (years)					0.643
≤60	519	51.4	591	52.4	
>60	490	48.6	536	47.6	
Sex					0.772
Male	754	74.7	836	74.2	
Female	255	25.2	291	25.8	
Smoking status					< 0.001
Nonsmoker	357	35.4	625	55.5	
Smoker	652	64.6	502	44.5	
Smoking level†					< 0.001
≤25 pack-years	204	31.3	240	47.8	
>25 pack-years	448	68.7	262	52.2	
Histological type§					
SCC	310	30.7			
AC	588	58.3			
Other	111	11.0			

* Two-sided χ^2 test. † Categorized by using the 50th percentile of pack-years (cigarettes per day/20 × years smoked) of the control subjects as the cutoff points. § SCC, squamous cell carcinoma; AC, adenocarcinoma; Other, includes large cell, mixed cell and undifferentiated carcinoma.

1.04 (95% CI = 0.85-1.26) and 1.04 (95% CI = 0.85-1.27), respectively, for one copy and 1.01 (95% CI = 0.62-1.66) and 0.92 (95% CI = 0.50-1.71), respectively, for two copies, compared with the common homozygote GG. When we combined the variant genotypes (GA+AA) assuming a dominant genetic model, the adjusted ORs were 1.03 (95% CI = 0.86-1.25) for rs9295740 and 1.03 (95% CI = 0.85-1.25) for rs2075789, respectively. Similar results were also observed in the recessive genetic model by comparing the variant homozygote AA with the combined genotypes (GA+GG), being ORs of 1.00 (95% CI = 0.61-1.64) for rs9295740 and 0.91 (95% CI = 0.49-1.69) for rs2075789, respectively.

To control the potential confounding factors and detect the phenotype-specific associations, the subgroup analyses were performed according to age, gender, smoking status and histologic types of lung cancer, when assuming a dominant genetic model to enhance the statistic power (**Figure 1**). However, no significant associations were observed in any subgroups for either SNP (**Figure 1**).

Discussion

In this case-control study of lung cancer in a Chinese population, we did not find any significant associations between genetic variants of 6p21.33 and 6p22.1 regions and lung cancer risk, which were identified in Caucasian populations, indicating the different susceptibility markers in different ethnic populations. To the best of our knowledge, this is the first report that extended the significant associations of SNPs of 6p regions with lung cancer from published GWAS of Caucasians to Asian populations, showing completely different findings between Chinese populations and populations of European ancestry.

The GWAS conducted by Wang *et al.* [7] demonstrated a convincing association between polymorphisms at the chromosome 6p21.33 and lung cancer risk in Caucasians. rs3117582 at intron 1 of *BAT3* and rs3131379 at intron 10 of *MSH5* were identified as the most significant loci in this region. Both *BAT3* and *MSH5* represent strong candidate genes for lung cancer susceptibility. BAT3 is an essential regulator of p53-mediated responses to genotoxic stress, and controls DNA damage-induced acetylation of p53 [9]. MSH5 is known to play functional roles in an array of cellular processes such as DNA damage response as well as meiotic homologous recombination, suggesting critical roles in genome stability

Table 2. Logistic regression analyses on associations of the SNPs rs9295740 and rs2075789 with risk of lung cancer

Locus	Genotype	Cases† (n = 996)		Controls§ (n = 1119)		Crude OR	Adjusted OR*
		No.	%	No.	%		
rs9295740							
Additive model	GG	660	66.3	755	67.5	1.00	1.00
	GA	305	30.6	324	29.0	1.07(0.89-1.30)	1.04(0.85-1.26)
	AA	31	3.1	40	3.6	0.89(0.55-1.43)	1.01(0.62-1.66)
Dominant model	GG	660	66.3	755	67.5	1.00	1.00
	GA+AA	336	33.7	364	32.6	1.06(0.88-1.27)	1.03(0.86-1.25)
Recessive model	GG+GA	965	96.9	1079	96.4	1.00	1.00
	AA	31	3.1	40	3.6	0.87(0.54-1.40)	1.00(0.61-1.64)
rs2075789							
Additive model	GG	713	71.6	806	72.0	1.00	1.00
	GA	263	26.4	289	25.8	1.03(0.85-1.25)	1.04(0.85-1.27)
	AA	20	2.0	24	2.1	0.94(0.52-1.72)	0.92(0.50-1.71)
Dominant model	GG	713	71.6	806	72.0	1.00	1.00
	GA+AA	283	28.4	313	28.0	1.02(0.85-1.24)	1.03(0.85-1.25)
Recessive model	GG+GA	976	98	1095	97.9	1.00	1.00
	AA	20	2.0	24	2.1	0.94(0.51-1.70)	0.91(0.49-1.69)

* Adjusted by age, sex and smoking status. † Thirteen of the 1009 cases failed to be genotyped for both SNPs. § Eight of the 1127 controls failed to be genotyped for both SNPs.

and integrity during DNA replication [10]. However, neither rs3131379 nor rs3117582 was polymorphic in Asian populations according to the HapMap database. Then we screened the potentially functional variants of these two genes and found the SNP rs2075789 with Pro-to-Ser change (P29S) at exon 2 of *MSH5* with MAF>0.05 in Chinese populations. This missense variant located within the interacting domain of MSH5 and resulted in a weakened protein interaction with MSH4 [11]. However, we did not find any significant association of rs2075789 with the risk of lung cancer in our Chinese population, suggesting that the functional variant *MSH5* rs2075789 might contribute little to the lung carcinogenesis. Since the SNP rs2075789 was selected based on the functional consideration and could not tag the variants in 6p21.33 region, failure to detect the effect of rs2075789 did not necessarily invalidate the associations between variants at chromosome 6p21.33 and lung cancer risk. Besides, ethnic differences in both genetic makeups and environment exposures may interplay in the etiology of lung cancer. For example, it appears that Asian smokers are less susceptible to lung cancer compared with African American smokers [12] and different genetic polymorphisms have been identified in Chinese compared with those in Caucasians [13]. Therefore, different genetic markers and diverse environmental exposures might account for the different effect across the Caucasians and Asian populations.

The SNP rs9295740 was identified as a risk locus for lung cancer at another region of chromosome 6p (6p22.1) by Wang *et al*, which was in extensive LD with the region of 6p21.33 [7]. The genotype frequencies of rs9295740 among controls in the present study (GG: 67.5%; GA: 29%; and AA: 3.6%) were very close to those in Caucasian populations (GG: 68.8%; GA: 28.1%; and AA: 3.1%). However, similar distributions of the genotypes (GG, GA and AA) were also observed among the NSCLC cases and the controls in this study, suggesting that there was no main effect of this variant on the risk of NSCLC in this Chinese population. This discrepancy may be explained by different ethnic background and rs9295740 could not be the true causal locus for lung cancer. The association with lung cancer risk identified in Caucasians might be mediated by linkage disequilibrium with one or more causal loci, while it is possible that rs9295740 is not correlated with the causal loci in Chinese populations. For example, both rs3131379 and rs3117582 were in moderate LD with rs9295740 (r^2 = 0.38 and 0.39, respectively) in the study of Wang *et al* [7], but rs3131379 and 3117582 were not polymorphic in Chinese populations and such correlation could not be observed.

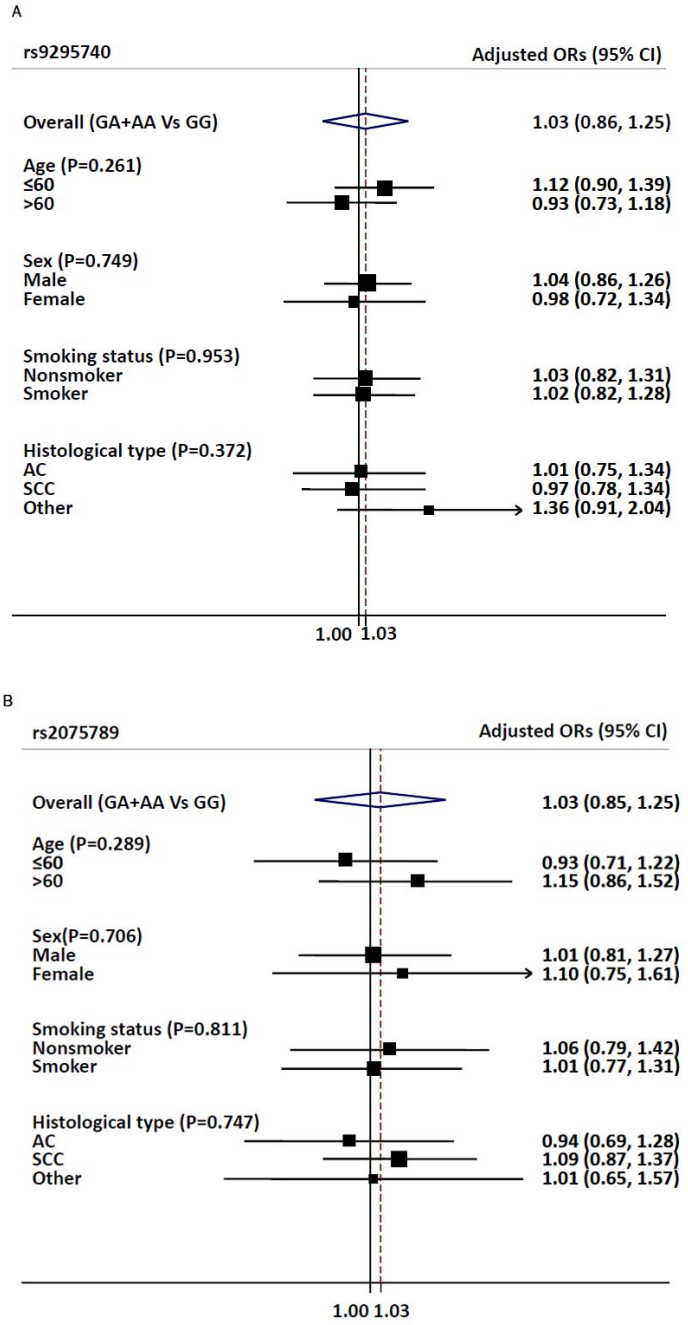

Figure 1. Forest plots represent the effect of the SNPs rs9295740 (Figure 1A) in 6p22.1 and rs2075789 (Figure 1B) in 6p21.33 on lung cancer risk among different subgroups. *P* values are from heterogeneity tests. AC, adenocarcinoma; SCC, squamous cell carcinoma; other includes large cell, mixed cell and undifferentiated carcinoma.

The results from this study indicate that the SNP (rs9295740) identified in Caucasian populations seemed not applicable to Chinese populations as a biomarker for lung cancer risk, and the non-synonymous SNP rs2075789 was not likely to be the causal locus of lung cancer in the region across 6p21.33–6p22.1. However, our results were not enough to deny the associations between variants at 6p21.33–6p22.1 regions and the risk of lung

cancer in Chinese populations. To further clarify such associations, re-sequencing and fine-mapping this region in Chinese populations is strongly proposed. In addition, more extensive examination of the transcripts in this region is required to find out other candidate genes and to determine the causal variants for lung cancer risk. The ethnic difference in copy number variations in those regions may also need to be examined in the future.

Besides, there was still some limitation in our study. Although we conducted the replication study with a large sample size, the power of our test was only 38.04% to detect the allele effect of rs9295740 with the OR of 1.20, which was detected in Caucasians, given the genotype frequency in this Chinese population and a significant level of 0.05. However, we noticed that the ORs of both SNPs detected in our population were quite around the null value (OR=1.02 for rs9295740 and OR=1.01 for rs2075789), which may support the hypothesis of null association.

In conclusion, this is the first study to replicate the GWAS findings observed in populations of European ancestry and the results indicate that the SNP rs9295740 seems not applicable to Chinese populations as the susceptible markers for lung cancer, and that the non-synonymous SNP rs2075789 at exon 2 of *MSH5* may not be causative. Re-sequencing and fine-mapping in this region, along with extensive examination of potential functional variants and further replications in different populations, are required to identify the exact casual variants of lung cancer.

Acknowledgements

This work was supported in part by Key Grant of National Natural Science Foundation of China (30730080); National Outstanding Youth Science Foundation of China (30425001), the China National Key Basic Research Program Grants (2002CB512902) and Natural Science Foundation of Jiangsu Province (BS2006005).

Please address correspondences to: Hongbing Shen, PhD, Department of Epidemiology and Biostatistics, Cancer Center, Nanjing Medical University, 140 Hanzhong Rd., Nanjing 210029, China. Tel. (Fax.): 86-25-8686-2756. E-mail: hbshen@njmu.edu.cn

References

[1] Schwartz AG, Prysak GM, Bock CH, Cote ML. The molecular epidemiology of lung cancer. Carcinogenesis 2007;28:507–18.

[2] Hecht SS. Cigarette smoking and lung cancer: chemical mechanisms and approaches to prevention. Lancet Oncol 2002;3:461–9.

[3] Kiyohara C, Otsu A, Shirakawa T, Fukuda S, Hopkin JM. Genetic polymorphisms and lung cancer susceptibility: a review. Lung Cancer. 2002;37(3):241-56.

[4] Hung RJ, McKay JD, Gaborieau V, Boffetta P, Hashibe M, Zaridze D, Mukeria A, Szeszenia-Dabrowska N, Lissowska J, Rudnai P, Fabianova E, Mates D, Bencko V, Foretova L, Janout V, Chen C, Goodman G, Field JK, Liloglou T, Xinarianos G, Cassidy A, McLaughlin J, Liu G, Narod S, Krokan HE, Skorpen F, Elvestad MB, Hveem K, Vatten L, Linseisen J, Clavel-Chapelon F, Vineis P, Bueno-de-Mesquita HB, Lund E, Martinez C, Bingham S, Rasmuson T, Hainaut P, Riboli E, Ahrens W, Benhamou S, Lagiou P, Trichopoulos D, Holcátová I, Merletti F, Kjaerheim K, Agudo A, Macfarlane G, Talamini R, Simonato L, Lowry R, Conway DI, Znaor A, Healy C, Zelenika D, Boland A, Delepine M, Foglio M, Lechner D, Matsuda F, Blanche H, Gut I, Heath S, Lathrop M, Brennan P. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature 2008;452, 633–7.

[5] Thorgeirsson TE, Geller F, Sulem P, Rafnar T, Wiste A, Magnusson KP, Manolescu A, Thorleifsson G, Stefansson H, Ingason A, Stacey SN, Bergthorsson JT, Thorlacius S, Gudmundsson J, Jonsson T, Jakobsdottir M, Saemundsdottir J, Olafsdottir O, Gudmundsson LJ, Bjornsdottir G, Kristjansson K, Skuladottir H, Isaksson HJ, Gudbjartsson T, Jones GT, Mueller T, Gottsäter A, Flex A, Aben KK, de Vegt F, Mulders PF, Isla D, Vidal MJ, Asin L, Saez B, Murillo L, Blondal T, Kolbeinsson H, Stefansson JG, Hansdottir I, Runarsdottir V, Pola R, Lindblad B, van Rij AM, Dieplinger B, Haltmayer M, Mayordomo JI, Kiemeney LA, Matthiasson SE, Oskarsson H, Tyrfingsson T, Gudbjartsson DF, Gulcher JR, Jonsson S, Thorsteinsdottir U, Kong A, Stefansson K. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 2008;452, 638–42.

[6] Amos CI, Wu X, Broderick P, Gorlov IP, Gu J, Eisen T, Dong Q, Zhang Q, Gu X, Vijayakrishnan J, Sullivan K, Matakidou A, Wang Y, Mills G, Doheny K, Tsai YY, Chen WV, Shete S, Spitz MR, Houlston RS. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet 2008;40, 616–22.

[7] Wang Y, Broderick P, Webb E, Wu X, Vijayakrishnan J, Matakidou A, Qureshi M, Dong Q, Gu X, Chen WV, Spitz MR, Eisen T, Amos CI, Houlston RS. Common 5p15.33 and

6p21.33 variants influence lung cancer risk. Nat Genet 2008;40(12):1407-9.
[8] Hu Z, Wang H, Shao M, Jin G, Sun W, Wang Y, Liu H, Wang Y, Ma H, Qian J, Jin L, Wei Q, Lu D, Huang W, Shen H. Genetic variants in MGMT and risk of lung cancer in Southeastern Chinese: a haplotype-based analysis. Hum Mutat 2007;28(5):431-40.
[9] Sasaki T, Gan EC, Wakeham A, Kornbluth S, Mak TW, Okada H. HLA-B-associated transcript 3 (Bat3)/Scythe is essential for p300-mediated acetylation of p53. Genes Dev 2007;21: 848–61.
[10] Stojic L, Brun R, Jiricny J. Mismatch repair and DNA damage signaling. DNA Repair (Amst) 2004;3(8-9):1091–101.
[11] Yi W, Wu X, Lee TH, Doggett NA, Her C. Two variants of MutS homolog hMSH5:Prevalence in humans and effects on protein interaction. Biochemical and Biophysical Research Communications 2005;332:524–32.
[12] Haiman CA, Stram DO, Wilkens LR, Pike MC, Kolonel LN, Henderson BE, Le Marchand L. Ethnic and Racial Differences in the Smoking-Related Risk of Lung Cancer. N Engl J Med. 2006 Jan 26; 354(4):333-42.
[13] Wu C, Hu Z, Yu D, Huang L, Jin G, Liang J, Guo H, Tan W, Zhang M, Qian J, Lu D, Wu T, Lin D, Shen H. Genetic variants on chromosome 15q25 associated with lung cancer risk in Chinese populations. Cancer Res. 2009 Jun 15;69(12):5065-72.