RESEARCH ARTICLE

Meta-Analysis of the Association between the *rs8034191* Polymorphism in *AGPHD1* and Lung Cancer Risk

Le Zhang[1,2]*, Tian-Bo Jin[1,2,3]*, Ya Gao[4], Hui-Juan Wang[1,2], Hua Yang[1,2], Tian Feng[2], Chen Chen[2], Long-Li Kang[3], Chao Chen[1,2&]

Abstract

Background: Possible associations between the single nucleotide polymorphism (SNP) *rs8034191* in the aminoglycosidephosphotransferase domain containing 1 (*AGPHD1*) gene and lung cancer risk have been studied by many researchers but the results have been contradictory. Materials and Methods: A computerized search for publications on *rs8034191* and lung cancer risk was performed. Odds ratios (ORs) with 95% confidence intervals (CIs) were calculated to assess the association between *rs8034191* and lung cancer risk with 13 selected case-control studies. Sensitivity analysis, test of heterogeneity, cumulative meta-analysis, and assessment of bias were also performed. Results: A significant association between *rs8034191* and lung cancer susceptibility was found using the dominant genetic model (OR=1.344, 95% CI: 1.285-1.406), the additive genetic model (OR=1.613, 95% CI: 1.503-1.730), and the recessive genetic model (OR=1.408, 95% CI: 1.319-1.503). Moreover, an increased lung cancer risk was found with all genetic models after stratification of ethnicity. Conclusions: The association between *rs8034191* and lung cancer risk was significant using multiple genetic models, suggesting that *rs8034191* is a risk factor for lung cancer. Further functional studies of this polymorphism and lung cancer risk are warranted.

Keywords: Lung cancer - single nucleotide polymorphism - *AGPHD1* - genetic polymorphism - meta-analysis

Asian Pac J Cancer Prev, **16 (7)**, 2713-2717

Introduction

Lung cancer is the leading cause of cancer-related death throughout the worldwith an estimated 1.3 million new cases diagnosed annually (Shibuya et al.,2002; Herbst et al., 2008). In many countries,the morbidity and mortality of lung cancer have increased rapidly in recent years (Bhat et al., 2013; Shukla et al., 2013; Yilmaz et al., 2014). Well-known risk factors for lung cancer include cigarette smoking and exposure to ionizing radiation. Although over 80% of lung cancer cases are related to the use of tobacco (Parkin et al., 1994), only a small percentage of smokers (<20%) develop this disease. A ccumulating evidence suggests that genetic factors may contribute to variation in susceptibility to lung cancer. It is widely accepted that lung cancer is a complex multifactorial disease that is attributed to the interaction of genetic factors with environmental factors (Amos et al., 2008; Heller et al., 2010;).Despite intensive efforts devoted to investigating the genetic factors associated with lung cancer, the genes and genetic variants that drive the development of lung cancer remain unclear.

Recently,the chromosome 15q24-25.1 region has been identified as a hot spot for lung cancer susceptibility by genome-wide association (GWA) studies (Hung et al., 2008;Thorgeirsson et al., 2008; Broderick et al., 2009; Wei et al., 2011). Genes that map to this region include aminoglycoside phosphotransferase domain containing 1 (*AGPHD1*); cholinergic receptor, nicotinic, alpha 3 (CHRNA3); cholinergic receptor, nicotinic, alpha 4 (CHRNA4); cholinergic receptor, nicotinic, alpha 5 (*CHRNA5*); *PSM4*; and *LOC123688*. In particular, the relationship between the single nucleotide polymorphism (SNP) *rs8034191* in *AGPHD1* and lung cancer risk has been widely investigated but the results have been inconclusive (Mantel et al., 1959; Amos et al., 2008;Schwartz et al., 2009; Zienolddiny et al., 2009; Truong et al., 2010; Chen et al., 2011; Jaworowska et al., 2011; Sakoda et al., 2011; Wang et al., 2013).Due to insufficient sample size, these previous studies have lacked statistical power to detect common variants that have minor effects on lung carcinogenesis. Furthermore, the results of these studies are not reproducible. To address the heterogeneity and publication bias among

*[1]School of Life Sciences, Northwest University, [2]National Engineering Research Center for Miniaturized Detection Systems, [3]Key Laboratory of High Altitude Environment and Genes Related to Diseases of Tibet Autonomous Region, School of Medicine, Tibet University for Nationalities, [4]Department of Forensic Medicine, Xi'an Jiaotong University School of Medicine, Xi'an, China [&]Equal contributors *For correspondence: cchen898@nwu.edu.cn*

previous studiesand better understand the effect of the SNP *rs8034191* on the risk of lung cancer we performed a meta-analysis of 13 selected case-control studies.

Materials and Methods

Publication search andinclusion/exclusion criteria

In August 2013, we searched PubMed, Google Scholar, EMBASE, and the ChinaNational Knowledge Infrastructure using the following search terms: *AGPHD1*, *rs8034191*, lung cancer, gene, genotype, mutation, and polymorphism. Articles identifiedin the primary literaturemet our initial criteria for inclusion in the meta-analysis if they were published in English, focused on humans, and were free of obvious overlap with other studies.

From the publications identified above, two investigators independently selected articles containing information on the association between *AGPHD1* and lung cancer morbidity and checked the corresponding reference lists. If multiple studies were published on the same population or subpopulation, only the most recent or informative study was included in the meta-analysis.

Articles were included in this meta-analysis if they 1) examinedthe hypothesis that *rs8034191*is associated withlung cancer risk, 2) followed a nested case-control, case-control,or cross-sectional study design, and 3) provided estimates of ORs and corresponding 95% CIs or sufficientinformation on genotype/allele counts between cases and controlsto calculatethe ORsand 95%CIs. Articles were excluded if they included non-case-control studies, a control population containing patients with malignant tumor, orwere redundant with other published studies. A flow chart outlining the selection process for inclusion in the meta-analysis is shown in Figure 1.

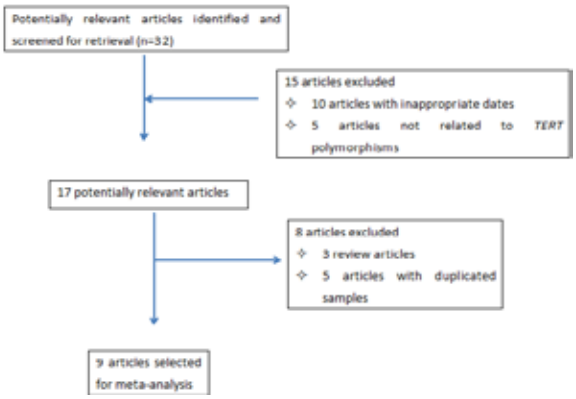

Figure 1. Study Inclusion and Exclusion Procedure

Data extraction

The following information was extracted from each study: the first author's name, the year of publication, the countryin which the study was performed, ethnicities of subjects, andthe number of cases and controlswith the TT, TC, and CC *rs8034191* genotypes. Two investigators independently extracted the data from all eligible publications, and any inconsistencies were resolved by discussion.

All statistical analyses were performed using STATA software (version 11.0; Stata Corporation, College Station, TX). Two-sided *p*-values less than 0.05 were considered statistically significant. We calculated the allelic frequenciesfor the case and control groups in each studyand assessed them for Hardy-Weinberg equilibrium (HWE) using achi-square test (Egger et al.,1997). The OR and 95% CI valueswere determinedto assess the strength of the association between each *rs8034191* polymorphism and lung cancer risk. For each study, the OR and 95% CI were assessed in adominant model, a recessive model, and an additive model. Subgroup analyses were performed based on the source of the controls and the ethnicity of the study participants. The chi-squared based Q-statistic was calculated to test for heterogeneity among the studies. If the studies were found to be heterogeneous ($p<0.05$) the pooled ORs were analyzed using a random-effects model (Higgins et al., 2002); otherwise, a fixed-effects model was used (Egger et al.,1997). The I^2statistic was then used to quantitatively estimate heterogeneity, with I^2less than 25%, between 25% and 75%, and greater than 75% representing low, moderate, and high degrees of inconsistency, respectively (Higgins et al., 2003; Hemminki et al., 2006). The significance of the combined OR was determinedusing a Z test ($p<0.05$ was considered statistically significant). Cumulative meta-analyses were performed on all eligible cancer studies according to case sample size. Additionally, sensitivity of the meta-analysis was evaluated throughthe sequential removal of each study.

Finally, we produced aBegg's funnel plot and performed an Egger's test tostatistically assess publication

Table 1. Characteristics of the studies on the association between *AGPHD1 rs8034191* polymorphisms and cancer risk included in the meta-analysis

ID	Author	Year	Ethnic group	Sample Size (Case/Control)	Source of controls	Case Alleles			Control Alleles			*p*-value (HWE)
						TT	TC	CC	TT	TC	CC	
1	Amos et al.(Texas discovery)	2008	Caucasian	1153/1137	HB	426	536	191	493	522	122	0.352
2	Amos et al.(Texas discovery)	2008	Caucasian	698/591	HB	259	328	111	269	253	69	0.421
3	Amos et al.(UK)	2008	Caucasian	1831/960	HB	670	858	303	448	415	97	0.951
4	Schwartz et al.(Caucasian)	2009	Caucasian	809/539	HB	185	264	90	326	367	116	0.44
5	Schwartz et al.(African American)	2009	African American	421/360	PB	231	119	10	300	106	15	0.148
6	Zienolddiny et al.(Norwegian)	2009	Caucasian	352/424	PB	117	178	57	176	187	61	0.324
7	Truong et al.(Caucasian)	2010	Caucasian	7259/9463	-	2586	3488	1185	4036	4256	1171	0.344
8	Truong et al.(Asian)	2010	Asian	1690/2117	-	1583	104	3	1992	122	3	0.43
9	Sakoda et al.(non-Hispanic white)	2011	Caucasian	746/1475	PB	258	369	119	625	691	159	0.117
10	Wei1 et al.(Caucasian)	2011	Caucasian	198/295	HB	64	100	34	127	137	31	0.505
11	Jaworowska et al.(Polish)	2011	Caucasian	833/831	HB	286	419	128	368	361	102	0.357
12	Chen et al.(Caucasian)	2011	Caucasian	487/974	HB	222	212	53	433	442	99	0.373
13	Wang et al.(Chinese)	2012	Asian	381/410	HB	350	29	2	385	25	0	0.524

Table 2. Stratified Analyses of the Association between *AGPHD1 rs8034191* Polymorphisms and Lung Cancer Risk

Variables	Dominant model OR(95%CI)	p^a	Recessive model OR(95%CI)	p^a	Additive model OR(95%CI)	p^a
Total	1.344(1.285-1.406)[b]	0.142	1.408(1.319-1.503)[b]	0.296	1.613(1.503-1.730)[b]	0.184
Ethnicity						
Asian	1.127(0.888-1.432)[b]	0.447	1.887(0.489-7.274)[b]	0.399	1.902(0.494-7.327)[b]	0.395
Caucasian	1.352(1.291-1.417)[b]	0.107	1.412(1.323-1.508)[b]	0.262	1.613(1.503-1.730)[b]	0.153

[a] *p*-value from chi-square test for heterogeneity; [b] A fixed-effects model was used when the *p* value from the heterogeneity test was <0.05; otherwise, a random-effects model was used

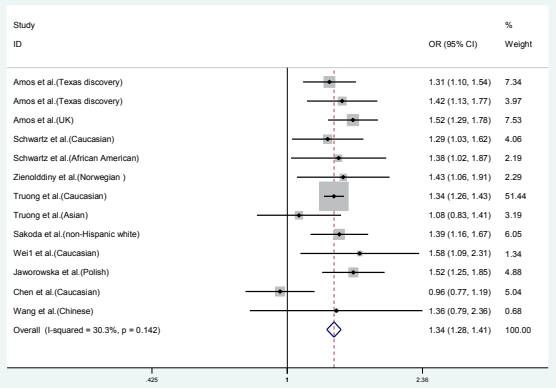

Figure 2. Overall Meta-analysis of the AGPHD1 *rs8034191* Polymorphism and Lung Cancer Risk in the Dominant Genetic Mode

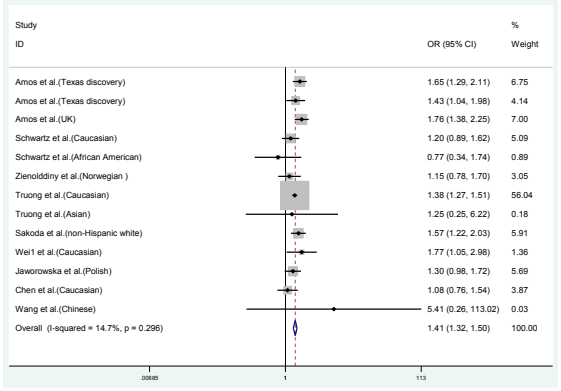

Figure 3. Overall Meta-analysis of the AGPHD1 *rs8034191* Polymorphism and Lung Cancer Risk in the Recessive Genetic Mode

bias. The Egger's test was based on the linear regression of the standard normal deviate against the precision of the standard normal deviate and was used to test the symmetry of the Begg's funnel plot. $p<0.05$ was considereda significant publication bias (Higgins et al., 2002).

Results

Eligible studiesand quality assessment

Through the procedure outlined in Figure 1 we identified ninearticles that met the inclusion criteria. These articles covered13 case-control studies including 16,858 cases and 19,576 controls.The characteristics of the studies are listed in Table 1. The sestudies focused solely on lung cancer and represented multiple ethnic populations. The alleles at *rs8034191*were in HWE in all of the studies. After evaluating these studies, all 13 were deemed to be of sufficient quality to be included in our analysis.

Meta-analysis results

After pooling the 16,858 cases and 19,5776 controls

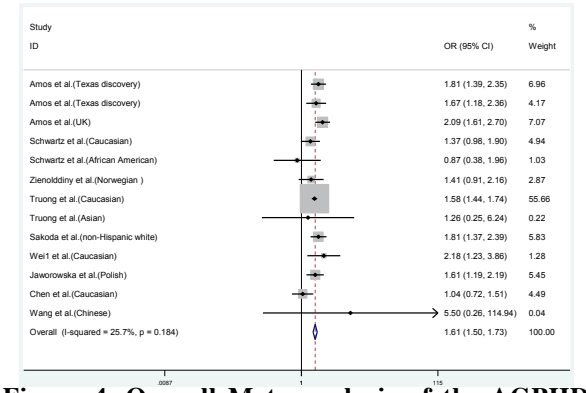

Figure 4. Overall Meta-analysis of the AGPHD1 *rs8034191* Polymorphism And Lung Cancer risk in the Additive Genetic Mode

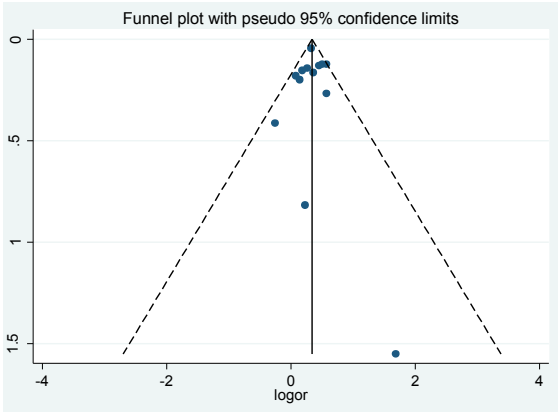

Figure 5. Begg's Funnel Plot to Test for Publication Bias in Theassociation between *rs8034191* Polymorphisms and Lung Cancer Risk

in the meta-analysis we found a significant association between *AGPHD1 rs8034191* polymorphisms and lung cancer risk using the dominant model (OR=1.344, 95% CI: 1.285-1.406), the additive model (OR=1.613, 95% CI: 1.503-1.730), and the recessive model (OR=1.408, 95% CI: 1.319-1.503).

There was no significant heterogeneity in the *rs8034191* variant genemodel. We stratified the data by dividing the participants into two subgroups based on ethnicity: Asian and Caucasian. The pooled Ors for the recessive model were 1.887 (95% CI 0.489-7.274) for the Asian subgroup and1.412 (95% CI 1.323-1.508) for the Caucasian subgroup (Table 2).

Test for heterogeneity

Based on the dominant model, there were significant heterogeneities associated with the study of Chen et al. (OR=0.955; 95% CI: 0.768-1.189) (Figure 2). Likewise, heterogeneity was detected in the study of African Americans by Schwartz et al.when using either the recessive model (OR=0.773; 95% CI: 0.343-1.743)

(Figure 3) or the additive model (OR=0.866; 95% CI: 0.382-1.963) (Figure 4).

Sensitivity of the meta-analysis

The pooled OR values were not qualitatively changed by the elimination of any individual study (data not shown), indicating that the final results of the meta-analysis were relatively stable and reliable. Likewise, the results of the meta-analysis were not influenced by departure of the allele frequency from HWE .

Publication bias

A major concern for any meta-analysis is the potential introduction of a publication bias based on the inclusion/exclusion criteria used to select studies for the analysis. To investigate whether publication bias was present in this study, a Begg's funnel plot was constructed (Figure 5). The shape of the Begg's funnel plot was relatively symmetric, indicating there was no obvious publication bias. An Egger's test based on this plot revealed no statistically significant asymmetry in the funnel plot for any of the genetic models (p=0.933).

Discussion

We performeda comprehensive meta-analysis to evaluate the associationof a common polymorphismon 15q25.1 with the risk of lung cancer. By performing subgroup analyses, we identified ethnicity as a potential source of inconsistency between studies. This is not surprising, as it is well established that genetic heterogeneity isinevitable in disease identification strategies (Higgins et al., 2003). Specifically, the overall results of this study demonstrated that the *rs8034191*-T allele of the *AGPHD1* gene might be a risk factor for the development of lung cancer in Caucasians, but not in Asians (Table 2). We also noticed remarkable differences inthe *rs8034191*-C allele between Caucasians and Asians, making it very difficult to detect weak associations in Asians unless examininga very large population. This suggests that different genetic backgrounds may differentially affect this allele or that different populations may have different linkage disequilibrium patterns; for example, the studied polymorphisms may be in linkage with another causalvariant in one ethnic population but not in another (Hemminki et al., 2006). Considering the divergent genetic backgrounds, it is necessary to construct a database of polymorphisms related to lung cancer ineach ethnic/racial group.

To the best of our knowledge, the present study is the only meta-analysis to date investigating the association of the *rs8034191* polymorphism with lung cancer susceptibility. Although potential sources of heterogeneity cannot be easily eliminated, the strengths of this study include the relatively large sample size, the lack of deviation from Hardy-Weinberg equilibrium, and the highquality of the studies involved. However, this study should be interpreted with several technical limitations in mind. First, most of the studies in this meta-analysis were case-control studies, which are susceptible to selection bias by including only non-fatalcases. Second, because only studies in the English language were considered, a publication bias may have been introduced. To address this, we performed a funnel plot and an accompanying Egger's test, which did not reveal an obvious bias. Moreover, any asymmetry in the funnel plot, either through visually interpretation or statistically testing, may result from a fundamental difference between small and large studies that arises frominherent inter-study heterogeneity There is no perfect method to test for publication bias, and the validity of the funnel plot and Egger's test have been challenged(Yu et al., 2010).Thus, we cannot completely rule out the low probability that relevant studies (for example small negative studies) are missing from the plot although the trim and fill method suggested that no missing studies were required to make the funnel plot symmetrical for either polymorphism. Third, the single locus-based nature of this meta-analysis precluded the possibility of investigating gene-geneand gene-environment interactions, as well as haplotype-based effects. In particular, further studies should investigate other markers adjacent to 15q25.1 to clarify whether the observed association is causal or due to linkage disequilibrium. It is likely that the *rs8034191*SNP alone makes aminor contribution to risk prediction in lung cancer patients,and further studies are needed to determine whether multiple polymorphisms integrated with other riskfactors will enhance the predictive capabilities. Additional studies are necessary to fully understand the relationship between the *rs8034191* SNP and lung cancer susceptibility.

In conclusion, we have expanded previous studies by providing evidence thatthe *rs8034191*-T allele of the *AGPHD1* gene mightbe a risk factor for the development of lung cancer inCaucasians, but not in Asians. Functional studies of the association between this polymorphism and cancer risk are warranted.

Acknowledgements

This work is supported by National 863 High-Technology Research and Development Program(No. 2012AA02A519).We are grateful to all the patients and individuals in the studies who made this work possible. We would also like to thank the clinicians and other hospital staff who contributed to the data collection for this study.

References

Amos CI, Wu X, Broderick P, et al (2008). Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet*. **40**, 616-22.

Bhat IA, Pandith AA, Bhat BA, et al (2013). Lack of association of a common polymorphism in the 3' -UTR of interleukin 8 with non small cell lung cancer in Kashmir. *Asian Pac J Cancer Prev*. **14**, 4403-8.

Broderick P, Wang Y, Vijayakrishnan J, et al (2009). Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. *Cancer Res*, **69**, 6633-41.

Chen J, Wu X, Pande M, et al (2011). Susceptibility locus for lung cancer at 15q25.1 is not associated with risk of pancreatic cancer. *Pancreas*, **40**, 872-5.

Egger M, Davey SG, Schneider M, Minder C (1997). Bias in

meta-analysis detected by a simple, graphical test. *BMJ*, **315**, 629-34.
Heller G, Zielinski CC, Zochbauer-Muller S (2010). Lung cancer: from single-gene methylation to methylome profiling. *Cancer Metastasis Rev*, **29**, 95-107.
Hemminki K, Lorenzo BJ, Forsti A(2006).The balance between heritable and environmental aetiology of human disease. *Nat Rev Genet*, **7**, 958-65.
Herbst RS, Heymach JV, Lippman SM (2008). Lung cancer. *N Engl J Med*, **359**, 1367-80.
Higgins JP, Thompson SG (2002). Quantifying heterogeneity in a meta-analysis. *Stat Med*, **21**, 1539-58.
Higgins JP, Thompson SG, Deeks JJ, et al (2003).Measuring inconsistency in meta-analyses. *BMJ*, **327**, 557-60.
Hung RJ, McKay JD, Gaborieau V, et al (2008). A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature*, **452** ,633-7.
Jaworowska E, Trubicka J, Lener MR, et al (2011). Smoking related cancers and loci at chromosomes 15q25, 5p15, 6p22.1 and 6p21.33 in the Polish population. *PLoS One*, **6**, 25057.
Mantel N, HaenszelW. (1959).Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*, **22**, 719-48.
Parkin DM, Pisani P, Lopez AD, et al (1994). At least one in seven cases of cancer is caused by smoking. Global estimates for 1985. *Int J Cancer*, **59**, 494-504.
Sakoda LC, Loomis MM, Doherty JA, et al (2011).Chromosome 15q24-25.1 variants, diet, and lung cancer susceptibility in cigarette smokers. *Cancer Causes Control*, **22**, 449-61.
Schwartz AG, Cote ML, Wenzlaff AS, et al (2009). Racial differences in the association between SNPs on 15q25.1, smoking behavior, and risk of non-small cell lung cancer. *J Thorac Oncol*, **4**, 1195-201.
Shibuya K, Mathers CD, Boschi-Pinto C, et al (2002).Global and regional estimates of cancer mortality and incidence by site: II. Results for the global burden of disease 2000. *BMC Cancer*, **2**, 37.
Shukla RK, Tilak AR, Kumar C, et al (2013). Associations of CYP1A1, GSTM1 and GSTT1 polymorphisms with lung cancer susceptibility in a Northern Indian population. *Asian Pac J Cancer Prev*, **14**, 3345-9.
Thorgeirsson TE, Geller F, Sulem P, et al (2008). A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature*, **452**, 638-42.
Truong T, Hung RJ, Amos CI, et al (2010). Replication of lung cancer susceptibility loci at chromosomes 15q25, 5p15, and 6p21: a pooled analysis from the International Lung Cancer Consortium. *J Natl Cancer Inst*, **102**, 959-71.
Wang H, Zhao Y, Ma J, et al (2013). The genetic variant rs401681C/T is associated with the risk of non-small cell lung cancer in a Chinese mainland population. *Genet Mol Res*, **12** ,67-73.
Wei C, Han Y, Spitz MR, et al (2011). A case-control study of a sex-specific association between a 15q25 variant and lung cancer risk. *Cancer Epidemiol Biomarkers Prev*, **20**, 2603-9.
Yilmaz M, Kacan T, Sari I, et al (2014). Lack of association between the MTHFRC677T polymorphism and lung cancer in a Turkish population. *Asian Pac J Cancer Prev*, **15**, 6333-7.
Yu K, Zhang J, Zhang J, et al (2010). Methionine synthase A2756G polymorphism and cancer risk: a meta-analysis. *Eur J Hum Genet*, **18**, 370-8.
Zienolddiny S, Skaug V, Landvik NE, et al (2009). The TERT-CLPTM1L lung cancer susceptibility variant associates with higher DNA adduct formation in the lung. *Carcinogenesis*, **30**, 1368-71.