Published in final edited form as:
Cancer Epidemiol Biomarkers Prev. 2010 February ; 19(2): 517–524. doi: 10.1158/1055-9965.EPI-09-0791.

Cumulative Effect of Multiple Loci on Genetic Susceptibility to Familial Lung Cancer

Pengyuan Liu[1], **Haris G. Vikis**[1], **Yan Lu**[1], **Yian Wang**[1], **Ann G. Schwartz**[2], **Susan M. Pinney**[3], **Ping Yang**[4], **Mariza de Andrade**[4], **Adi Gazdar**[5], **Colette Gaba**[6], **Diptasri Mandal**[7], **Juwon Lee**[3], **Elena Kupert**[3], **Daniela Seminara**[8], **John Minna**[5], **Joan E. Bailey-Wilson**[9], **Christopher I. Amos**[10], **Marshall W. Anderson**[3], and **Ming You**[1]

[1]Washington University, St. Louis, Missouri [2]Karmanos Cancer Institute, Detroit, Michigan [3]University of Cincinnati, Cincinnati, Ohio [4]Mayo Clinic, Rochester, Minnesota [5]University of Texas Southwestern Medical Center, Dallas, Texas [6]University of Toledo College of Medicine, Toledo, Ohio [7]Louisiana State University Health Science Center, New Orleans, Louisiana [8]National Cancer Institute, Bethesda, Maryland [9]National Human Genome Research Institute, Baltimore, Maryland [10]M. D. Anderson Cancer Center, Houston, Texas

Abstract

Background—Genetic factors play important roles in lung cancer susceptibility. In this study, we replicated the association of 5p15.33 and 6p21.33 with familial lung cancer. Taking into account the previously identified genetic susceptibility variants on 6q23-25/*RGS17* and 15q24-25.1, we further determined the cumulative association of these four genetic regions and the population attributable risk percent of familial lung cancer they account for.

Methods—One hundred ninety-four case patients and 219 cancer-free control subjects from the Genetic Epidemiology of Lung Cancer Consortium were used for the association analysis. Each familial case was chosen from one high-risk lung cancer family that has three or more affected members. Single nucleotide polymorphisms (SNP) on chromosomal regions 5p15.33, 6p21.33, 6q23-25/*RGS17*, and 15q24-25.1 were assessed for their associations with familial lung cancer. The cumulative association of the four chromosomal regions with familial lung cancer was evaluated with the use of a linear logistic model. Population attributable risk percent was calculated for each SNP using risk ratio.

Results—SNP rs31489 showed the strongest evidence of familial lung cancer association on 5p15.33 ($P = 2 \times 10^{-4}$; odds ratio, 0.57; 95% confidence interval, 0.42-0.77), whereas rs3117582 showed a weak association on 6p21.33 ($P = 0.09$; odds ratio, 1.47; 95% confidence interval, 0.94-2.31). Analysis of a combination of SNPs from the four regions provided a stronger cumulative association with familial lung cancer ($P = 6.70 \times 10^{-6}$) than any individual SNPs. The risk of lung cancer was increased to 3- to 11-fold among those subjects who had at least one copy of risk allele at each region compared with subjects without any of the risk factors. These four genetic regions contribute to a total of 34.6% of familial lung cancer in smokers.

Corresponding Author: Ming You, Department of Surgery and The Alvin J. Siteman Cancer Center, Washington University, 660 Euclid Avenue, Box 8109, St. Louis, MO 63110. Phone: 314-362-9294; Fax: 314-362-9366. youm@wudosis.wustl.edu.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Conclusions—The SNPs in four chromosomal regions have a cumulative and significant association with familial lung cancer and account for about one-third of the population attributable risk for familial lung cancer.

Introduction

Lung cancer is the leading cause of cancer death for both men and women in the United States. In 2008, there was an estimated 215,000 cases of lung cancer diagnosed and only 15% of those patients are expected to survive 5 years (1). Although cigarette smoking is the major risk factor for lung cancer, genetic factors also affect lung cancer susceptibility (2). For example, increased lung cancer rates are observed in several genetic syndromes including Li-Fraumeni syndrome, hereditary retinoblastoma, familial breast cancer, and Bloom syndrome (3-7). Direct evidence for a genetic predisposition to lung cancer is highlighted by several genome-wide association studies (GWAS) that have been recently done on lung cancer populations with European ancestry. These GWAS identified the associations of common variants on chromosome 15q24-25.1, 5p15.33, and 6p21.33 with lung cancer susceptibility (8-13).

Lung cancer can occur in multiple members of the same family and constitute familial lung cancer. We have previously identified the association of the chromosomal region 15q24-25.1 with familial lung cancer through a GWAS of 194 case patients with family history and 219 cancer-free control subjects collected by the Genetic Epidemiology of Lung Cancer Consortium (GELCC). The odds ratio (OR) associated with the 15q24-25.1 locus was higher than that observed in sporadic lung cancer samples (10). In addition, recent fine mapping studies identified a major candidate gene, *RGS17*, for the familial lung cancer susceptibility locus on chromosome 6q23-25. Two single nucleotide polymorphisms (SNP; rs9479510 and rs4083914) in the *RGS17* gene are associated with familial lung cancer, but do not associate with sporadic lung cancer (14). *RGS17* encodes a recently identified member of the regulator of G protein signaling (RGS) family. RGS proteins negatively regulate G protein–related signaling at least in part by accelerating the GTPase activity of Gα subunits. *RGS17* is highly expressed in tumor tissues. The loss of *RGS17* transcript inhibits the growth of xenografted tumors and the proliferation of tumor cells, whereas overexpression of *RGS17* increases the rate of proliferation of tumor cells (14,15).

Here, we describe the association of 5p15.33 and 6p21.33 with familial lung cancer. Taking all of the four chromosomal regions together, we assess their cumulative association with familial lung cancer and estimate the population attributable risk (PAR%) of these genetic variants underlying lung cancer susceptibility. We showed that a combination of SNPs from the chromosomal regions have a stronger association with familial lung cancer than any individual SNPs.

Materials and Methods

Patient Samples and SNP Genotyping

To confirm the association of 5p15.33 and 6p21.33 with familial lung cancer, we analyzed SNPs within these two chromosomal regions from our previous study (10). We have previously conducted a GWAS among individuals with a familial history of lung cancer using the Affymetrix SNP chips. These individuals are members of families with three or more members with lung cancer that were collected as part of the GELCC collections. To ensure genetic independence among subjects, each case patient with familial lung cancer was chosen from one high-risk lung cancer family. All the case patients in this study are histologically confirmed non–small cell lung cancer. Non-cancer control subjects were obtained from a combination of unaffected spouses from GELCC families and of unaffected

individuals from the Coriell Institute for Medical Research and the Fernald Medical Monitoring Program. These control subjects had no blood relationship with any selected case patients. Totally, 194 case patients with familial lung cancer and 219 cancer-free control subjects were used. Detailed information on the GELCC samples was described in Supplementary Table S1.

We analyzed 144 SNPs on 5p15.33 (1.1-1.7 Mb) and 2,337 SNPs on 6p21.33 (27.0-34.0 Mb) from our Affymetrix SNP chips. To comprehensively cover these two regions, we typed additional 11 SNPs on 5p15.33 and 21 SNPs on 6p21.33 in the GELCC collection. These additional SNPs are those showing strong association in previous studies (11,13) but were not presented in the Affymetrix SNP chips. SNP genotyping was done using the Sequenom mass array spectrometry system at the Human Genetics Division Genotyping Core of Washington University (St. Louis, MO). Genotyping results are initially evaluated according to the fraction of assay successes per plate. The genotyping call rate is >90% for each SNP in our study. In addition to assay success rates, genotyping plates are reviewed for results from positive and negative DNA control wells that are organized in specific patterns to assist in the quality check (QC) process and to ensure correct plate orientations during processing and data review.

Statistical Analyses

Hardy-Weinberg equilibrium for each SNP was examined among control subjects with the use of Fisher exact test. The statistical significance of the association between SNP allele and disease status was assessed primarily with the Cochran-Armitage trend test with a 1 degree of freedom and with the Fisher exact test, implemented in the PLINK software.[11] Allelic ORs associated with each SNP and 95% confidence intervals (CI) were estimated. To rule out the confounding effects of smoking behavior on lung cancer risk, the association analysis was adjusted by sex, age, and pack-years of cigarette exposure with the use of logistic regression analysis.

We tested the cumulative effects of the four chromosomal regions (5p15.33, 6p21.33, 15q24-25.1, and 6q23-25/*RGS17*) on familial lung cancer by counting the number of risk allele at each locus in logistic regression model. One SNP was chosen from each of the four chromosomal regions. The chosen SNPs are those that have been validated with association with sporadic lung cancer and/or familial lung cancer. To improve the stability of estimates of OR, we combined the sporadic lung cancer samples from three recent GWAS (8,9,13) with our familial samples in the data analysis. The cumulative ORs for subjects carrying different copies of risk alleles were estimated by comparing them with those carrying none of these risk alleles.

PAR% was estimated for each locus, which defines what percentage of the total risk for lung cancer is due to the genetic effect of that locus: $PAR = \Sigma p_i \times (OR_i - 1)/(\Sigma p_i \times (OR_i - 1) + 1)$, in which p_i is the prevalence of risk allele at *i*th locus associated with lung cancer among control subjects, and OR_i is OR of risk allele at *i*th locus. The joint PAR was calculated on the basis of the individual PAR of each associated SNP: $1 - \Pi(1 - PAR_i)$ in which PAR_i is the individual PAR for each associated SNP.

[11] http://pngu.mgh.harvard.edu/~purcell/plink/

Results

Association of 5p15.33

In the GELCC collection of 194 familial cases and 214 cancer-free controls, we observed a statistically significant association between 5p15.33 and familial lung cancer (Supplementary Table S2; Table 1). Twelve SNPs on the 5p15.33 have *P* value of <0.01. Four of five SNPs identified in previous studies (11, 13) were confirmed in the GELCC collections. Among these SNPs, rs31489 is the most significant one with *P* value of 2.0×10^{-4} (OR, 0.57; 95% CI, 0.42-0.77). The frequency of the minor allele A of rs31489 was 32.49% in cancer patients versus 45.79% in disease-free subjects. Subjects carrying one copy of minor allele A are at 1.75-fold reduced risk of developing lung cancer. The association of the 5p15.33 locus with familial lung cancer remains statistically significant after adjusting sex, age, and smoking quantity ($P = 8.0 \times 10^{-4}$). We did not observe the association of 5p15.33 with smoking quantity in the familial samples (Supplementary Table S3).

We then compared the association of 5p15.33 in cases from families with three affected members and in cases from families with four or more affected members. Generally, the association of 5p15.33 tends to be stronger in cases from families with four or more affected members than in cases from families with three affected members. However, there is no statistically significant difference in allelic ORs between these two groups of familial cases (Table 2). When comparing the association of 5p15.33 in familial and sporadic cases, we observed that the effect size of 5p15.33 is significantly larger in familial case than in sporadic cases ($P = 4.8 \times 10^{-3}$ from the Woof's test; OR, 0.57; 95% CI, 0.42-0.77 in familial cases; OR, 0.88; 95% CI, 0.84-0.84 in sporadic cases; Table 3; ref. 13).

In the GELCC collections, the susceptibility region on 5p15.33 contains three genes: *TERT*, *CLPTM1L*, and *SLC6A3*. SNPs with the strongest associations are located within *CLPTM1L* (Fig. 1A). Interestingly, we identified an additional significant SNP rs466630 ($P = 4.5 \times 10^{-3}$; OR, 1.85; 95% CI, 1.24-2.77) that is independent of previously reported SNPs ($r^2 < 0.001$; Supplementary Table S2). This SNP is located within *SLC6A3*.

Association of 6p21.33

We observed a weak association between 6p21.33 and familial lung cancer in the GELCC collections (Supplementary Table S4; Table 1). Among the three SNPs on 6p21.33 reported in a recent study (13), rs3117582 achieved a *P* value of 0.09 (OR, 1.47; 95% CI, 0.94-2.31). The minor allele *C* of rs3117582 is enriched in cases compared with disease-free controls, which is consistent with that observed in the previous study (13). The weak association observed on 6p21.33 may be due to the small sample size of the GELCC collections in this study and to the low population frequency of the risk allele of rs3117582. However, we found that several SNPs (e.g., $P = 6 \times 10^{-4}$; OR, 1.71; 95% CI, 1.25-2.33 for rs1634718) nearby rs3117582 show stronger associations with familial lung cancer (Supplementary Table S4; Fig. 1B). There is no significant difference in the strength of association of 6p21.33 by the number of familial affecteds in the kindred of origin (Table 2).

Cumulative Association of Several Genetic Variants

We then evaluated what percentage of the total risk for familial lung cancer is due to the genetic effect of the 5p15.33, 6p21.33, and 15q24-25.1. The 15q24-25.1 region has been recently reported to be associated with both familial and sporadic lung cancer (8-10,12). At least two common variants on 15q24-25.1, SNPs rs8034191and rs1051730, were confirmed in the GELCC collections (Supplementary Table S5 and S6). We observed that 5p15.33, 6p21.33, and 15q24-25.1 account for 7.5%, 2.7%, and 8.5% risk of lung cancer in the

familial population, respectively. The three genetic regions contribute to a total of 17.6% PAR for familial lung cancer (Supplementary Table S7; Table 4). The risk of lung cancer ranged from 2- to 4-fold (that is cumulative OR, 2-4) among those subjects who had at least one copy of risk allele at each of the three regions compared with subjects without any of risk alleles.

RGS17 has been recently identified as a major candidate gene for the familial lung cancer susceptibility locus on chromosome 6q23-25 (Supplementary Table S5 and S6; ref. 14). We therefore took *RGS17* into consideration in the analysis of the cumulative association of familial lung cancer. The 6q23-35/*RGS17* accounts for an additional 20.66% PAR for lung cancer in the familial samples. The total risk of familial lung cancer contributed by these four genetic regions was further increased to 34.6% (Table 4). Analysis of a combination of SNPs from these regions provided a stronger cumulative association with familial lung cancer ($P = 6.70 \times 10^{-6}$). The cumulative OR associated with these four regions ranged from 3- to 11-fold among those subjects who had at least one copy of risk allele at each region, compared with subjects without any of the risk factors. However, <2% carriers of risk alleles have cumulative ORs larger than five in the populations (Table 5).

Discussion

In this study, we replicated the overall association of 5p15.33 and 6p21.33 with familial lung cancer. We found that the SNP rs31489 shows the strongest evidence of familial lung cancer association on 5p15.33 ($P = 2 \times 10^{-4}$; OR, 0.57; 95% CI, 0.42-0.77), whereas rs3117582 shows a relatively weak association on 6p21.33 ($P = 0.09$; OR, 1.47; 95% CI, 0.94-2.31). Although each of these chromosomal regions was only moderately associated with lung cancer, we observed that they have a strong cumulative association with the disease. We show 3- to 11-fold increase in lung cancer risk based on the cumulative association of 5p15.33, 6p21.33, 15q24-25.1, and 6q23-25/*RGS17* among those subjects who had at least one copy of risk allele at each of the four regions compared with subjects without any of the risk factors. These four genetic regions account for 34.6% of familial lung cancer. These results further confirm the importance of these genomic regions underlying lung cancer susceptibility.

Our results should encourage further studies characterizing candidate genes for each of the loci and the mechanism about how these genetic variants contribute to the risk of lung cancer. The candidate genes in the 5p15.33 locus include *TERT* and *CLPTM1L* in previous studies (11,13). In the present study, fine mapping was done on the 5p15.33 locus using a more dense set of SNPs from the Affymetrix SNP 6.0 array and additional SNPs genotyped by the Sequenom mass array spectrometry system. The familial data show very strong associations on the 5p15.33, and the strongest association was located within *CLPTM1L*. Several SNPs in *SLC6A3* also showed strong associations with familial lung cancer (Fig. 1A). Gain in 5p15.33 is one of the most frequent events observed in early-stage non–small cell lung cancer (16). The 5p15.33 locus has been recently identified to associate with multiple cancer types including lung, urinary bladder, prostate, and cervical cancer (17). *TERT* is the reverse transcriptase component of telomerase and is an attractive candidate for the 5p15.33 locus. Telomere attrition limits the replicative potential of most somatic cells. In contrast, tumor cells acquire immortality by continuous telomere maintenance, which is predominantly due to the transcriptional upregulation of *TERT* (18). Interestingly, *TERT* has also been found to directly promote proliferation (19,20). *CLPTM1L*, originally named for its relationship to a gene in the cleft palate susceptibility, was founded to sensitize ovarian cancer cells to cisplatin-induced apoptosis (21). *SLC6A3* (also named the dopamine transporter, DAT1) mediates the active reuptake of dopamine from the synapse and is a principal regulator of dopaminergic neurotransmission. The *SLC6A3* gene has been

implicated in human disorders such as Parkinsonism, Tourette syndrome, and substance abuse (22).

We also observed a weak association on 6p21.33 in the familial data. Among many genes at this locus, BAT3 and MSH5 are attractive candidates. BAT3/SCYTHE is known to regulate apoptosis and is essential for p53 acetylation and p53-mediated DNA damage response (23). Thus, variants with decreased function may confer the risk of lung cancer through decreased apoptosis in response to DNA damage. MSH5 is a member of the MUTS family of proteins involved in DNA mismatch repair and can stabilize the double-strand break repair intermediately (24). Missense mutations found in MSH5 have been shown to result in increased tolerance to killing by DNA alkylating agents (25). It is conceivable that variants of MSH5 could confer tolerance to DNA alkylation–induced apoptosis.

We have previously identified the association between 15q24-25.1 and familial lung cancer in the GELCC collections. The 15q25.1 contains *IREB2, LOC123688, PSMA4, CHRNA5 CHRNA3*, and *CHRNB4*. The *CHRN* genes encode for subunits of the nicotinic acetylcholine receptor subunits, seem biologically relevant, and are initially attractive candidates. Among the remaining genes in the locus, IREB2 encodes an iron regulator protein 2 and plays a central role in maintaining cellular iron homeostasis (26); PSMA4 encodes a structural protein of the 20S proteasome core (27,28); and *LOC123688* is a hypothetical gene.

Our study is unique because of its use of familial lung cancer samples from the GELCC collections. Each case was chosen from one family who has three or more members with lung cancer. However, several caveats for our findings should be acknowledged. First, familial lung cancer occurs in <1 of 1,000 lung cancer patients diagnosed. This results in small sample sizes in family studies and limits statistical power in the association analysis. Second, lung cancer is a late-onset disease. Tobacco smoking is well established as the major risk factor for lung cancer, contributing to a 10-fold increase in risk in long-term smokers compared with nonsmokers (29). To increase the statistical power to detect genetic variants and reduce confounding effects, in addition to spousal controls, we chose old, healthy, heavy smokers as controls. This sample strategy may lead to biased estimation of covariate effects and risk effects at genetic loci in the logistical regression model. However, these caveats should not affect our conclusion that a combination of SNPs from the chromosomal regions has a stronger association with familial lung cancer than any individual SNPs at chromosome 5p15.33, 6p21.33, 6q23-25, and 15q24-25.1. As a whole, the risk of lung cancer is increased significantly among those subjects who had at least one copy of risk allele at these chromosome regions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the Fernald Medical Monitoring Program for sharing their biospecimens and data with us, the lung cancer families who participated in this research, the Vanderbilt University Microarray Shared Resource, Washington University Genotyping Core, Mayo Clinic Genotyping Shared Resource (supported in part by P30CA 15083) for the high caliber service, and Shaw Levy and Jennifer Baker for their assistance in various aspects of this work.

Grant Support

NIH grants U01CA76293 (GELCC), R01CA058554, R01CA093643, R01CA099147, R01CA099187, R01ES012063, R01ES013340, R03CA77118, R01CA80127, P30ES06096, P50CA70907 (Specialized Program of Research Excellence), N01HG65404, N01-PC35145, P30CA22453, R01CA63700, DE-FGB-95ER62060, Mayo

Clinic intramural research funds, and Department of Defense VITAL grant. This study was supported in part by NIH, the Intramural Research Programs of the National Cancer Institute, and the National Human Genome Research Institute.

References

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008. CA Cancer J Clin 2008;58:71–96. [PubMed: 18287387]
2. Matakidou A, Eisen T, Houlston RS. Systematic review of the relationship between family history and lung cancer risk. Br J Cancer 2005;93:825–33. [PubMed: 16160696]
3. German J. Bloom syndrome: a mendelian prototype of somatic mutational disease. Medicine (Baltimore) 1993;72:393–406. [PubMed: 8231788]
4. Johannsson O, Ostermeyer EA, Hakansson S, et al. Founding BRCA1 mutations in hereditary breast and ovarian cancer in southern Sweden. Am J Hum Genet 1996;58:441–50. [PubMed: 8644702]
5. Li FP, Fraumeni JF Jr. Mulvihill JJ, et al. A cancer family syndrome in twenty-four kindreds. Cancer Res 1988;48:5358–62. [PubMed: 3409256]
6. Sanders BM, Jay M, Draper GJ, Roberts EM. Non-ocular cancer in relatives of retinoblastoma patients. Br J Cancer 1989;60:358–65. [PubMed: 2789942]
7. Strong LC, Herson J, Haas C, et al. Cancer mortality in relatives of retinoblastoma patients. J Natl Cancer Inst 1984;73:303–11. [PubMed: 6589425]
8. Amos CI, Wu X, Broderick P, et al. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nat Genet 2008;40:616–22. [PubMed: 18385676]
9. Hung RJ, McKay JD, Gaborieau V, et al. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature 2008;452:633–7. [PubMed: 18385738]
10. Liu P, Vikis HG, Wang D, et al. Familial aggregation of common sequence variants on 15q24-25.1 in lung cancer. J Natl Cancer Inst 2008;100:1326–30. [PubMed: 18780872]
11. McKay JD, Hung RJ, Gaborieau V, et al. Lung cancer susceptibility locus at 5p15.33. Nat Genet 2008;40:1404–6. [PubMed: 18978790]
12. Thorgeirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 2008;452:638–42. [PubMed: 18385739]
13. Wang Y, Broderick P, Webb E, et al. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. Nat Genet 2008;40:1407–9. [PubMed: 18978787]
14. You M, Wang D, Liu P, et al. Fine mapping of chromosome 6q23-25 region in familial lung cancer families reveals RGS17 as a likely candidate gene. Clin Cancer Res 2009;15:2666–74. [PubMed: 19351763]
15. James MA, Lu Y, Liu Y, Vikis HG, You M. RGS17, an overexpressed gene in human lung and prostate cancer, induces tumor cell proliferation through the cyclic AMP-PKA-CREB pathway. Cancer Res 2009;69:2108–16. [PubMed: 19244110]
16. Kang JU, Koo SH, Kwon KC, Park JW, Kim JM. Gain at chromosomal region 5p15.33, containing TERT, is the most frequent genetic event in early stages of non-small cell lung cancer. Cancer Genet Cytogenet 2008;182:1–11. [PubMed: 18328944]
17. Rafnar T, Sulem P, Stacey SN, et al. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. Nat Genet 2009;41:221–7. [PubMed: 19151717]
18. Janknecht R. On the road to immortality: hTERT upregulation in cancer cells. FEBS Lett 2004;564:9–13. [PubMed: 15094035]
19. Choi J, Southworth LK, Sarin KY, et al. TERT promotes epithelial proliferation through transcriptional control of a Myc- and Wnt-related developmental program. PLoS Genet 2008;4:e10. [PubMed: 18208333]
20. Sarin KY, Cheung P, Gilison D, et al. Conditional telomerase induction causes proliferation of hair follicle stem cells. Nature 2005;436:1048–52. [PubMed: 16107853]
21. Woods DC, Alvarez C, Johnson AL. Cisplatin-mediated sensitivity to TRAIL-induced cell death in human granulosa tumor cells. Gynecol Oncol 2008;108:632–40. [PubMed: 18191995]

22. Vandenbergh DJ, Persico AM, Uhl GR. A human dopamine transporter cDNA predicts reduced glycosylation, displays a novel repetitive element and provides racially-dimorphic TaqI RFLPs. Brain Res Mol Brain Res 1992;15:161–6. [PubMed: 1359373]

23. Bawa S, Xiao W. A mutation in the MSH5 gene results in alkylation tolerance. Cancer Res 1997;52:2715–20. [PubMed: 9205082]

24. Harfe BD, Jinks-Robertson S. DNA mismatch repair and genetic instability. Annu Rev Genet 2000;34:359–99. [PubMed: 11092832]

25. 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. [PubMed: 17403783]

26. LaVaute T, Smith S, Cooperman S, et al. Targeted deletion of the gene encoding iron regulatory protein-2 causes misregulation of iron metabolism and neurodegenerative disease in mice. Nat Genet 2001;27:209–14. [PubMed: 11175792]

27. Davoli R, Fontanesi L, Russo V, et al. The porcine proteasome subunit A4 (PSMA4) gene: isolation of a partial cDNA, linkage and physical mapping. Anim Genet 1998;29:385–8. [PubMed: 9800328]

28. Nandi D, Woodward E, Ginsburg DB, Monaco JJ. Intermediates in the formation of mouse 20S proteasomes: implications for the assembly of precursor β subunits. EMBO J 1997;16:5363–75. [PubMed: 9311996]

29. Doll R, Peto R. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. J Natl Cancer Inst 1981;66:1191–308. [PubMed: 7017215]

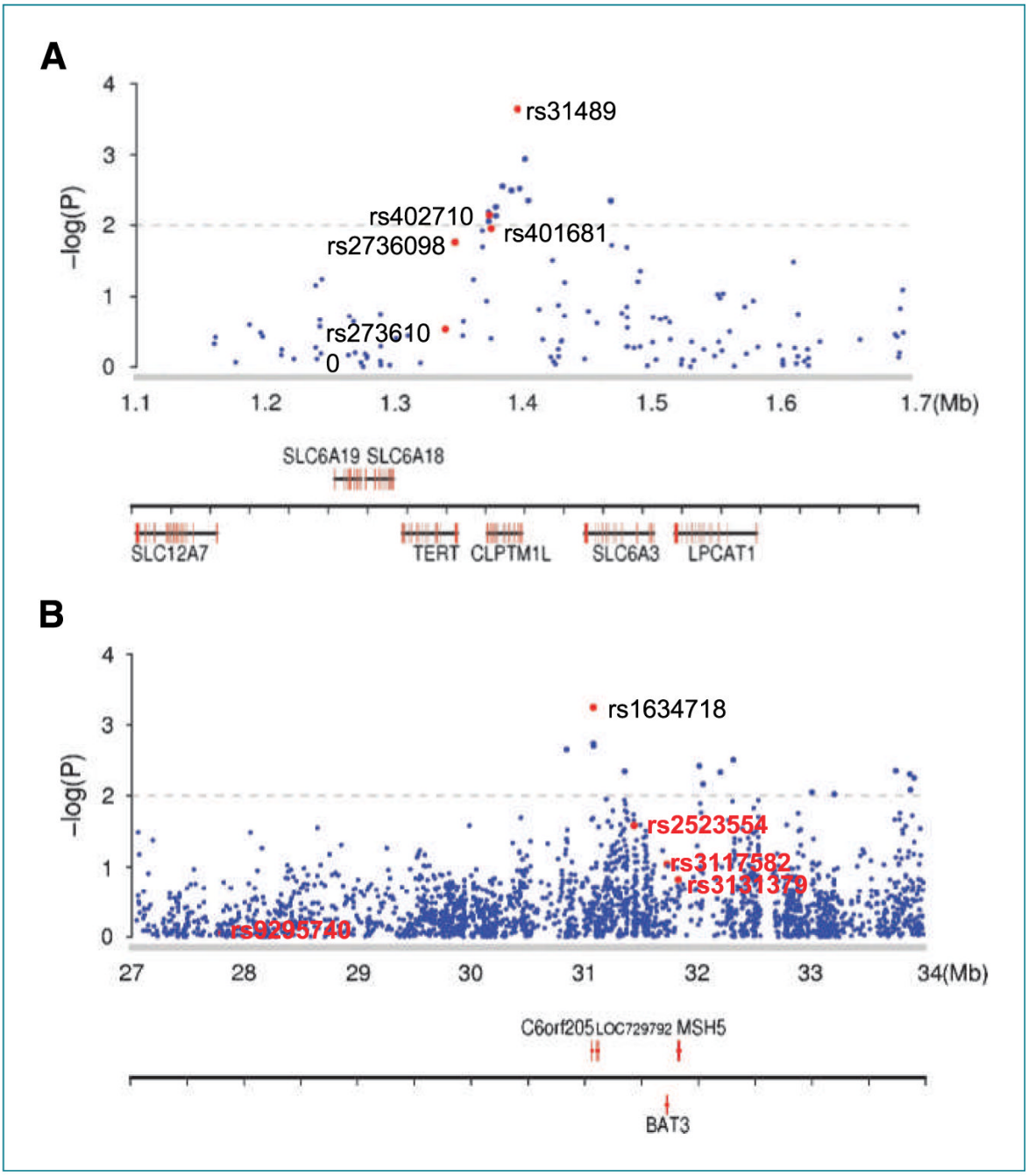

Figure 1.
Association results and genomic context of chromosomes 5p15.33 (A) and 6p21.33 (B). Association analyses were done on 194 familial cases and 219 controls from the GELCC collections.

Table 1

Association of SNPs on chromosomes 5p15.33 and 6p21.33 with familial lung cancer

dbSNP*	Chr	Pos	MA	Frequency		*P*	OR (95% CI)	Adjusted by sex, age, and pack-years	
				Cases	Controls			*P*	OR (95% CI)
rs2736100	5	1339516	A	0.4234	0.4630	0.2952	0.86 (0.65-1.15)	0.6035	1.12 (0.73-1.72)
rs2736098	5	1347086	T	0.3288	0.2512	0.0172	1.46 (1.07-2.00)	0.2120	1.31 (0.86-2.02)
rs402710	5	1373722	T	0.2778	0.3709	0.0071	0.65 (0.48-0.88)	0.0075	0.57 (0.38-0.86)
rs401681	5	1375087	A	0.3711	0.4589	0.0111	0.70 (0.53-0.92)	0.0189	0.63 (0.43-0.93)
rs31489	5	1395714	A	0.3249	0.4579	0.0002	0.57 (0.42-0.77)	0.0008	0.48 (0.31-0.74)
rs9295740	6	27797481	A	0.1865	0.1820	0.8735	1.03 (0.72-1.48)	0.7925	1.07 (0.65-1.75)
rs3117582	6	31728499	C	0.1290	0.0914	0.0916	1.47 (0.94-2.31)	0.1455	1.58 (0.85-2.92)
rs3131379	6	31829012	A	0.1198	0.0890	0.1518	1.39 (0.89-2.19)	0.1854	1.51 (0.82-2.78)

Abbreviatons: Chr, chromosome; Pos, position; MA, minor allele.

*These SNPs were validated in the studies of McKay et al. (2008) and/or Wang et al. (2008).

Table 2

Comparison of the association of 5p15.33 and 6p21.33 in cases from families with three affecteds and that in cases from families four or more affecteds

dbSNP	Chr	Pos	MA	3−Aff*		4+Aff†		P‡
				P	OR (95% CI)	*P*	OR (95% CI)	
rs2736100	5	1339516	A	0.3699	0.82 (0.56-1.22)	0.5018	0.89 (0.64-1.24)	0.7671
rs2736098	5	1347086	T	0.0266	1.63 (1.07-2.47)	0.0968	1.37 (0.96-1.95)	0.5382
rs402710	5	1373722	T	0.0417	0.64 (0.42-0.98)	0.0205	0.66 (0.47-0.93)	0.9279
rs401681	5	1375087	A	0.0209	0.63 (0.43-0.93)	0.0633	0.74 (0.53-1.01)	0.5574
rs31489	5	1395714	A	0.0079	0.57 (0.37-0.86)	0.0013	0.57 (0.41-0.80)	0.9627
rs9295740	6	27797481	A	0.4497	1.22 (0.75-1.97)	0.7493	0.93 (0.61-1.41)	0.4023
rs3117582	6	31728499	C	0.5056	1.23 (0.66-2.32)	0.0667	1.62 (0.98-2.66)	0.5085
rs3131379	6	31829012	A	0.6224	1.17 (0.63-2.19)	0.1123	1.53 (0.93-2.53)	0.5122

*Cases came from those families with three affected members with lung cancer.

†Cases came from those families with four or more affected members with lung cancer.

‡Allelic ORs estimated from the above groups were compared by Woolf's test.

Table 3

Comparison of the association of 5p15.33 and 6p21.33 in familial and sporadic cases

SNPs/samples*	Frequency		*P*	Allelic OR (95% CI)	*p*†
	Cases	Controls			
rs3117582 (C)‡					
GELCC	0.1290	0.0913	0.0916	1.47 (0.94-2.31)	—
UK-GWA	0.1568	0.1180	6.24×10^{-6}	1.39 (1.20-1.61)	0.810
Texas-GWA	0.1157	0.1118	0.7097	1.04 (0.86-1.25)	0.159
IARC-GWA	0.1081	0.0797	4.41×10^{-6}	1.40 (1.21-1.62)	0.831
rs401681 (A)					
GELCC	0.3711	0.4589	0.0111	0.70 (0.53-0.92)	—
UK-GWA	0.4090	0.4426	0.0056	0.87 (0.79-0.96)	0.136
Texas-GWA	0.4094	0.4424	0.0250	0.87 (0.78-0.98)	0.140
IARC-GWA	0.3831	0.4134	0.0039	0.88 (0.81-0.96)	0.113
rs31489 (A)					
GELCC	0.3249	0.4579	0.0002	0.57 (0.42-0.77)	—
UK-GWA	0.3801	0.4073	0.0235	0.89 (0.81-0.99)	0.005
Texas-GWA	0.3803	0.4098	0.0428	0.88 (0.78-1.00)	0.007
IARC-GWA	0.3570	0.3874	0.0034	0.88 (0.80-0.96)	0.006

*GELCC indicates familial lung cancer samples from the GELCC; UK-GWA and Texas-GWA indicates two lung cancer GWAS on sporadic lung cancer from the studies of Wang et al. (2008) and Amos et al. (2008); IARC GWA indicates the lung cancer GWAS on sporadic lung cancer from the study of Hung et al. (2008).

†Allelic ORs estimated from the GELCC and from each of sporadic lung cancer samples were compared by Woolf's test.

‡Letter in parenthesis is SNP minor allele.

Table 4

PARs for representative SNPs on chromosomes 5p15.33, 6p21.33, and 15q24-25.1, and *RGS17* gene

dbSNPs	Region/gene	Alleles*	Frequency of risk allele†		OR‡(95% CI)	*P*	PAR (%)
			Cases	Controls			
rs31489	5p15.33	**C**/A	0.6751	0.5421	1.15 (1.10-1.20)	1.19×10^{-6}	7.48
rs3117582	6p21.33	**C**/A	0.1290	0.0914	1.31 (1.20-1.43)	1.41×10^{-9}	2.74
rs1051730	15q24-25.1	**A**/G	0.4075	0.2948	1.31 (1.24-1.39)	3.76×10^{-21}	8.45
rs9479510	RGS17	**C**/G	0.5054	0.3774	1.69 (1.27-2.24)	0.0003	20.66
All 4 SNPs							34.64

*Letter in bold is risk allele.

†Frequency of risk allele was measured in the GELCC familial samples.

‡ORs were estimated from combined samples from sporadic lung cancer samples and our familial samples.

Table 5

Cumulative association of SNPs on chromosomes 5p15.33, 6p21.33, 15q24-25.1, and *RGS17* gene with familial lung cancer

Region/gene	**5p15.33**	**6p21.33**	**15q24-25.1**	**RGS17**	**Cumulative OR (95% C.I.)**
SNPs	**rs31489**	**rs3117582**	**rs1051730**	**rs9479510**	
Risk allele	**C**	**C**	**A**	**C**	
Geno. Freq. (%)*	**Copies of risk alleles**				
2.98	0	0	0	0	Reference
1.76	1	1	1	1	3.34 (2.83-3.84)
1.10	2	1	1	1	3.83 (3.32-4.35)
0.10	1	2	1	1	4.36 (3.82-4.91)
0.06	2	2	1	1	5.01 (4.46-5.56)
0.36	1	1	2	1	4.38 (3.86-4.90)
0.22	2	1	2	1	5.03 (4.50-5.56)
0.02	1	2	2	1	5.73 (5.17-6.29)
0.01	2	2	2	1	6.58 (6.02-7.15)
0.56	1	1	1	2	5.64 (4.66-6.62)
0.35	2	1	1	2	6.48 (5.49-7.46)
0.03	1	2	1	2	7.37 (6.37-8.37)
0.02	2	2	1	2	8.47 (7.47-9.48)
0.12	1	1	2	2	7.40 (6.41-8.39)
0.07	2	1	2	2	8.50 (7.51-9.50)
0.01	1	2	2	2	9.68 (8.67-10.69)
0.004	2	2	2	2	11.12 (10.11-12.14)

*Geno. Freq.: the frequency for a joint genotype at four loci in the population was calculated by assuming Hardy-Weinberg equilibrium at each allele and linkage equilibrium among the four genetic loci.