Table 2
Odds ratios and 95% confidence intervals from the stratified meta-analyses of the association between *MTHFR*[a] *C677T* and head and neck cancer and lung cancer

Tumour site		No. of cases	No. of controls	No. *MTHFR*[a] 677 TT cases	No. *MTHFR* 677 TT controls	OR[a,b]	95% CI[a]	*p* value for heterogeneity within strata	*p* value for heterogeneity across strata
Head and Neck[c]	Never drinkers	47	285	15	61	1.84	0.62–5.50	0.21	0.24
	Ever drinkers	517	1684	96	315	0.94	0.73–1.22	0.39	
	High folate intake	359	1127	63	247	0.85	0.63–1.16	0.72	0.06
	Low folate intake[d]	204	841	45	129	1.37	0.92–2.06	0.88	
Lung[e]	Never drinkers	279	549	49	123	0.90	0.33–2.47	0.006	0.71
	Ever drinkers	1480	2143	268	406	1.09	0.91–1.30	0.70	
	High folate intake	1447	2124	270	437	0.94	0.79–1.12	0.54	0.06
	Low folate intake[d]	718	842	132	131	1.28	0.97–1.68	0.36	

[a] OR, Odds Ratio; CI, Confidence Interval; *MTHFR*, methylenetetrahydrofolate reductase.
[b] The comparison is *MTHFR* 677 TT *versus* CC.
[c] Hung et al. [7], Suzuki et al. [27] and Weinstein at al. [32] studies were included.
[d] Low folate intake defined on the lower quartile estimated in the control population as provided by the authors and defined as: for head and neck <7 times/week [7], <245.7 μg/day [27] and <332.7 μg/day [32]; for lung <290.0 μg/day [6], <5 times/week [7], <253.6 μg/day [18]. See Section 3.3 for details.
[e] Shi et al. [6], Hung et al. [7] and Suzuki et al. [18] studies were included.

levels [40]. Our meta-analyses, however, failed to demonstrate overall a statistically significant risk of lung and head and neck cancer associated with *MTHFR* C677T homozygous variant genotype. Since *MTHFR* C677T appears to act as beneficial or deleterious depending on subjects' folate status, one would expect that the homozygous variant genotype would have no effect on cancer risk in population with high prevalence of folate supplement intake. More than a quarter of the weight in the results of both our meta-analyses' on lung and head and neck cancer was accounted for by studies conducted in the USA [6,24,25,30], where some common food items are regularly fortified with folate since 1998 [39]. Those studies actually showed almost the weakest association between 677 TT and lung and head and neck cancer, in fact by excluding these studies we showed a significant increased risk of lung cancer for *MTHFR* 677 TT genotype. Even if caution needs to be used when interpreting these results, both in view of the lengthy induction time for lung and head and neck cancer and the lag-time for an effect of folic acid, our results suggest a possible chemopreventive effect of folate more prominent in *MTHFR* 677 TT individuals, and a possible stronger role for the gene in those with low folate intake, which need to be addressed more in depth.

In the stratified meta-analysis according to alcohol consumption, we were unable to observe any effect modification, which is in line with the pooled analysis on gastric cancer [40]. However the information on alcohol did not take into account the amount or duration of alcohol which might be relevant especially for head and neck cancer. The present meta-analysis on *MTHFR* 677 TT and lung cancer included two additional studies [17,23] respect to the one previously published [34]. In addition, we further investigate the effect modification by folate status, which was lacking in the previous meta-analysis.

MTHFR A1298C has been reported to be in negative linkage disequilibrium with *C677T* [6,7]. The results of our meta-analyses revealed fluctuating estimates and overall null findings, which would suggest that C677T is the main *MTHFR* variant that is associated with cancer risk.

Some limitations should be considered when interpreting the results, in addition to those inherited from the meta-analysis. First, the data of the estimated dietary folate intake across the studies is collected using different food frequency questionnaires, and different cut-off values defines the lower quartile of folate intake in the studies, depending on the distribution of this variable in each specific population. Both situations, however, could lead to non-differential misclassification of the exposure and biased effect measures toward the null. If such bias is present in our data, it would indicate that the underlying true effect modification should be stronger than what we observed. Second, the subgroup meta-analyses on folate intake and alcohol consumption are based on a small number of studies with such information available. Nevertheless the total number of subjects included in this part of the analysis comprise the largest sample size so far.

Despite all these remarks, the observed increased risk for lung cancer among *MTHFR* 677 homozygous variant carriers with low dietary folate intake suggests that dietary folate might be protective in carcinogenesis especially in situation of impaired folate status as recently shown for gastric cancer [40]. Since more than half of the included studies were based on a limited number of cases (<200) it is critical that larger prospective studies, collecting detailed lifestyle habits data and repeated serological dosage of folate levels, are performed, in order to clarify the preventive role of folate in tobacco- and alcohol-related cancers. To overcome the limitation of meta-analysis, a coordinated genotyping of *MTHFR C677T* is now underway in the International Lung Cancer Consortium (http://ilcco.iarc.fr/), which will allow us to investigate the role of *MTHFR* in lung carcinogenesis and its potential effect modification by folate consumption.

Conflict of interest statement

Authors disclose any financial and personal relationships with other people or organisations that could inappropriately influence their work.

Funding

This study was supported by a training fellowship of Environmental Cancer Risk, Nutrition and Individual Susceptibility (ECNIS), a FP6 EC Network of Excellence (Contract No. ETP/14/2007).

Acknowledgments

We thank Takeshi Suzuki, Stephanie J.Weinstein and Paolo Vineis for collaboration and availability to share their data for the present meta-analysis.

References

[1] S.W. Choi, J.B. Mason, Folate and carcinogenesis: an integrated scheme, J. Nutr. 130 (2000) 129–132.
[2] B.C. Blount, M.M. Mack, C.M. Wehr, J.T. MacGregor, R.A. Hiatt, G. Wang, S.N. Wickramasinghe, R.B. Everson, B.N. Ames, Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage, Proc. Natl. Acad. Sci. USA 94 (1997) 3290–3295.
[3] S.J. Duthie, Folic acid deficiency and cancer: mechanisms of DNA instability, Br. Med. Bull. 55 (1999) 578–592.
[4] A. Parle-McDermott, J.L. Mills, A.M. Molloy, N. Carroll, P.N. Kirke, C. Cox, M.R. Conley, F. Pangilinan, L.C. Brody, J.N. Scott, The MTHFR 1298CC and 677TT genotypes have opposite associations with red cell folate levels, Mol. Genet. Metab. 88 (2006) 290–294.
[5] G. Friedman, N. Goldschmidt, Y. Friedlander, A. Ben-Yehuda, J. Selhub, S. Babaey, M. Mendel, M. Kidron, H. Bar-On, A common mutation A1298C in human methylenetetrahydrofolate reductase gene: association with plasma total homocysteine and folate concentrations, J. Nutr. 129 (1999) 1656–1661.
[6] Q. Shi, Z. Zhang, G. Li, P.C. Pillow, L.M. Hernandez, M.R. Spitz, Q. Wei, Sex differences in risk of lung cancer associated with methylenetetrahydrofolate reductase polymorphisms, Cancer Epidemiol. Biomarkers Prev. 14 (2005) 1477–1484.
[7] R.J. Hung, M. Hashibe, J. McKay, V. Gaborieau, N. Szeszenia-Dabrowska, D. Zaridze, J. Lissowska, P. Rudnai, E. Fabianova, I. Mates, L. Foretova, V. Janout, V. Bencko, A. Chabrier, N. Moullan, F. Canzian, J. Hall, P. Boffetta, P. Brennan, Folate related genes and the risk of tobacco-related cancers in Central Europe, Carcinogenesis 28 (2007) 1334–1340.
[8] P. Frosst, H.J. Blom, R. Milos, P. Goyette, C.A. Sheppard, R.G. Matthews, G. Boers, M. den Heijer, L.A. Kluijtmans, L.P. van den Heuve, R. Rozen, A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase, Nat. Genet. 10 (1995) 111–113.
[9] I. Weisberg, P. Tran, B. Christensen, S. Sibani, R. Rozen, A second genetic polymorphism in methylenetetrahydrofolate reductase (MTHFR) associated with decreased enzyme activity, Mol. Genet. Metab. 64 (1998) 169–172.
[10] K.J. Lievers, G.H. Boers, P. Verhoef, M. den Heijer, L.A. Kluijtmans, N.M. van der Put, F.J. Trijbels, H.J. Blom, A second common variant in the methylenetetrahydrofolate reductase (MTHFR) gene and its relationship to MTHFR enzyme activity, homocysteine, and cardiovascular disease risk, J. Mol. Med. 79 (2001) 522–528.
[11] L.B. Bailey, Folate, methyl-related nutrients, alcohol, and the MTHFR 677C ->T polymorphism affect cancer risk: intake recommendations, J. Nutr. 133 (2003) 3748S–3753S.
[12] Vainio, F. Bianchini (Eds), IARC Handbooks of Cancer Prevention, vol. 8, Fruit and Vegetables, IARC Press, Lyon, 2003, pp. 1–376.
[13] M.A. Kane, The role of folates in squamous cell carcinoma of the head and neck, Cancer Detect. Prev. 29 (2005) 46–53.
[14] E. Cho, D.J. Hunter, D. Spiegelman, D. Albanes, W.L. Beeson, P.A. van den Brandt, G.A. Colditz, D. Feskanich, A.R. Folsom, G.E. Fraser, J.L. Freudenheim, E. Giovannucci, R.A. Goldbohm, S. Graham, A.B. Miller, T.E. Rohan, T.A. Sellers, J. Virtamo, W.C. Willett, S.A. Smith-Warner, Intakes of vitamins A, C and E and folate and multivitamins and lung cancer: a pooled analysis of 8 prospective studies, Int. J. Cancer 118 (2006) 970–978.
[15] IARC Working Group on the evaluation of carcinogenic risks to humans schistosomes, liver flukes and *Helicobacter pylori*, IARC Monogr. Eval. Carcinog. Risks Hum. 61 (1994) 1–241.
[16] A. Jatoi, B.D. Daly, G. Kramer, J.B. Mason, Folate status among patients with non-small cell lung cancer: a case-control study, J. Surg. Oncol. 77 (2001) 247–252.
[17] P. Vineis, F. Veglia, S. Garte, C. Malaveille, G. Matullo, A. Dunning, M. Peluso, L. Airoldi, K. Overvad, O. Raaschou-Nielsen, F. Clavel-Chapelon, J.P. Linseisen, R. Kaaks, H. Boeing, A. Trichopoulou, D. Palli, P. Crosignani, R. Tumino, S. Panico, H.B. Bueno-De-Mesquita, P.H. Peeters, E. Lund, C.A. Gonzalez, C. Martinez, M. Dorronsoro, A. Barricarte, C. Navarro, J.R. Quiros, G. Berglund, B. Jarvholm, T.J. Key, R. Saracci, E. Riboli, H. Autrup, Genetic susceptibility according to three metabolic pathways in cancers of the lung and bladder and in myeloid leukemias in nonsmokers, Ann. Oncol. 18 (2007) 1230–1242.
[18] T. Suzuki, K. Matsuo, A. Hiraki, T. Saito, S. Sato, Y. Yatabe, T. Mitsudomi, T. Hida, R. Ueda, K. Tajima, Impact of one-carbon metabolism-related gene polymorphisms on risk of lung cancer in Japan: a case control study, Carcinogenesis 28 (2007) 1718–1725.
[19] X.M. Zhang, X.P. Miao, W. Tan, S.N. Qu, T. Sun, Y.F. Zhou, D.X. Lin, Association between genetic polymorphisms in methylentetrahydrofolate reductase and risk of lung cancer, Zhongguo Yi Xue Ke Xue Yuan Xue Bao 27 (2005) 700–703.
[20] M. Shen, N. Rothman, S.I. Berndt, X. He, M. Yeager, R. Welch, S. Chanock, N. Caporaso, Q. Lan, Polymorphisms in folate metabolic genes and lung cancer risk in Xuan Wei, China, Lung Cancer 49 (2005) 299–309.
[21] Y.L. Jeng, M.H. Wu, H.B. Huang, W.Y. Lin, S.L. You, T.Y. Chu, C.J. Chen, C.A. Sun, The methylenetetrahydrofolate reductase 677C ->T polymorphism and lung cancer risk in a Chinese population, Anticancer Res. 23 (2003) 5149–5152.
[22] K. Siemianowicz, J. Gminski, W. Garczorz, N. Slabiak, M. Goss, M. Machalski, H. Magiera-Molendowska, Methylenetetrahydrofolate reductase gene C677T and A1298C polymorphisms in patients with small cell and non-small cell lung cancer, Oncol. Rep. 10 (2003) 1341–1344.
[23] B.T. Heijmans, J.M. Boer, H.E. Suchiman, C.J. Cornelisse, R.G. Westendorp, D. Kromhout, E.J. Feskens, P.E. Slagboom, A common variant of the methylenetetrahydrofolate reductase gene (1p36) is associated with an increased risk of cancer, Cancer Res. 63 (2003) 1249–1253.
[24] H. Shen, M.R. Spitz, L.E. Wang, W.K. Hong, Q. Wie, Polymorphisms of methylene-tetrahydrofolate reductase and risk of lung cancer: a case-control study, Cancer Epidemiol. Biomarkers Prev. 10 (2001) 397–401.
[25] D.T. Hsiung, C.J. Marsit, E.A. Houseman, K. Eddy, C.S. Furniss, M.D. McClean, K.T. Kelsey, Global DNA methylation levels in whole blood as a biomarker in head and neck squamous cell carcinoma, Cancer Epidemiol. Biomarkers Prev. 16 (2007) 108–114.
[26] A. Reljic, A.M. Simundic, E. Topic, N. Nikolac, D. Justinic, M. Stefanovic, The methylenetetrahydrofolate reductase (MTHFR) C677T polymorphism and cancer risk: the Croatian case-control study, Clin. Biochem. 40 (2007) 981–985.
[27] T. Suzuki, K. Matsuo, Y. Hasegawa, A. Hiraki, K. Wakai, K. Hirose, T. Saito, S. Sato, R. Ueda, K. Tajima, One-carbon metabolism-related gene polymorphisms and risk of head and neck squamous cell carcinoma: case-control study, Cancer Sci. 98 (2007) 1439–1446.
[28] E. Vairaktaris, C. Yapijakis, P. Kessler, A. Vylliotis, J. Ries, J. Wiltfang, S. Vassiliou, S. Derka, F.W. Neukam, Methylenetetrahydrofolate reductase polymorphism and minor increase of risk for oral cancer, J. Cancer Res. Clin. Oncol. 132 (2006) 219–222.
[29] P. Capaccio, F. Ottaviani, V. Cuccarini, S. Cenzuales, B.M. Cesana, L. Pignataro, Association between methylenetetrahydrofolate reductase polymorphisms alcohol intake and oropharyngolaryngeal carcinoma in northern Italy, J. Laryngol. Otol. 119 (2005) 371–376.
[30] A.S. Neumann, H.J. Lyons, H. Shen, Z. Liu, Q. Shi, E.M. Sturgis, S. Shete, M.R. Spitz, A. El-Naggar, W.K. Hong, Q. Wei, Methylenetetrahydrofolate reductase polymorphisms and risk of squamous cell carcinoma of the head and neck: a case-control analysis, Int. J. Cancer 115 (2005) 131–136.
[31] N. Kureshi, S. Ghaffar, S. Siddiqui, I. Salahuddin, P.M. Frossard, Head and neck cancer susceptibility: a genetic marker in the methylenetetrahydrofolate reductase gene, ORL J. Otorhinolaryngol. Relat. Spec. 66 (2004) 241–245.
[32] S.J. Weinstein, G. Gridley, L.C. Harty, S.R. Diehl, L.M. Brown, D.M. Winn, E. Bravo-Otero, R.B. Hayes, Folate intake serum homocysteine and methylenetetrahydrofolate reductase (MTHFR) C677T genotype are not associated with oral cancer risk in Puerto Rico, J. Nutr. 132 (2002) 762–767.

[33] F. Gemignani, S. Landi, N. Szeszenia-Dabrowska, D. Zaridze, J. Lissowska, P. Rudnai, E. Fabianova, D. Mates, L. Foretova, V. Janout, V. Bencko, V. Gaborieau, L. Gioia-Patricola, I. Bellini, R. Barale, F. Canzian, J. Hall, P. Boffetta, R.J. Hung, P. Brennan, Development of lung cancer before the age of 50: the role of xenobiotic metabolizing genes, Carcinogenesis 28 (2007) 1287–1293.
[34] R. Mao, Y. Fan, Y. Jin, J. Bai, S. Fu, Methylenetetrahydrofolate reductase gene polymorphisms and lung cancer: a meta-analysis, J. Hum. Genet. 53 (2008) 340–348.
[35] R. DerSimonian, N. Laird, Meta-analysis in clinical trials, Control. Clin. Trials 7 (1986) 177–188.
[36] J.J. Deeks, D.G. Altman, M.J. Bradburn, Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis, in: M. Egger, G. Davey Smith, D.G. Altman (Eds.), Systematic Reviews in Health Care - Meta-analysis in Context, BMJ Books, London, 2005, pp. 285–312.
[37] M. Egger, G. Davey Smith, M. Schneider, C. Minder, Bias in meta-analysis detected by a simple, graphical test, BMJ 315 (1997) 629–634.
[38] J.P. Ioannidis, T.A. Trikalinos, The appropriateness of asymmetry tests for publication bias in meta-analyses: a large survey, CMAJ 176 (2007) 1091–1096.
[39] M.R. Malinow, P.B. Duell, D.L. Hess, P.H. Anderson, W.D. Kruger, B.E. Phillipson, R.A. Gluckman, P.C. Block, B.M. Upson, Reduction of plasma homocyst(e)ine levels by breakfast cereal fortified with folic acid in patients with coronary heart disease, N. Engl. J. Med. 338 (1998) 1009–1015.
[40] S. Boccia, R. Hung, G. Ricciardi, F. Gianfagna, M.P. Ebert, J.Y. Fang, C.M. Gao, T. Götze, F. Graziano, M. Lacasaña-Navarro, D. Lin, L. López-Carrillo, et al, Meta- and pooled analyses of the methylenetetrahydrofolate reductase C677T and A1298C polymorphisms and gastric cancer risk: a huge-GSEC review, Am. J. Epidemiol. 167 (2008) 505–516.

DOI: 10.1093/jnci/djq178
Advance Access publication on June 14, 2010.

ARTICLE

Replication of Lung Cancer Susceptibility Loci at Chromosomes 15q25, 5p15, and 6p21: A Pooled Analysis From the International Lung Cancer Consortium

Therese Truong, Rayjean J. Hung, Christopher I. Amos, Xifeng Wu, Heike Bickeböller, Albert Rosenberger, Wiebke Sauter, Thomas Illig, H.-Erich Wichmann, Angela Risch, Hendrik Dienemann, Rudolph Kaaks, Ping Yang, Ruoxiang Jiang, John K. Wiencke, Margaret Wrensch, Helen Hansen, Karl T. Kelsey, Keitaro Matsuo, Kazuo Tajima, Ann G. Schwartz, Angie Wenzlaff, Adeline Seow, Chen Ying, Andrea Staratschek-Jox, Peter Nürnberg, Erich Stoelben, Jürgen Wolf, Philip Lazarus, Joshua E. Muscat, Carla J. Gallagher, Shanbeh Zienolddiny, Aage Haugen, Henricus F. M. van der Heijden, Lambertus A. Kiemeney, Dolores Isla, Jose Ignacio Mayordomo, Thorunn Rafnar, Kari Stefansson, Zuo-Feng Zhang, Shen-Chih Chang, Jin Hee Kim, Yun-Chul Hong, Eric J. Duell, Angeline S. Andrew, Flavio Lejbkowicz, Gad Rennert, Heiko Müller, Hermann Brenner, Loïc Le Marchand, Simone Benhamou, Christine Bouchardy, M. Dawn Teare, Xiaoyan Xue, John McLaughlin, Geoffrey Liu, James D. McKay, Paul Brennan, Margaret R. Spitz

Manuscript received July 17, 2009; revised March 26, 2010; accepted April 26, 2010.

Correspondence to: Paul Brennan, PhD, Genetic Epidemiology Group, International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon Cedex 08, France (e-mail: Brennan@iarc.fr).

Background Genome-wide association studies have identified three chromosomal regions at 15q25, 5p15, and 6p21 as being associated with the risk of lung cancer. To confirm these associations in independent studies and investigate heterogeneity of these associations within specific subgroups, we conducted a coordinated genotyping study within the International Lung Cancer Consortium based on independent studies that were not included in previous genome-wide association studies.

Methods Genotype data for single-nucleotide polymorphisms at chromosomes 15q25 (rs16969968, rs8034191), 5p15 (rs2736100, rs402710), and 6p21 (rs2256543, rs4324798) from 21 case–control studies for 11 645 lung cancer case patients and 14 954 control subjects, of whom 85% were white and 15% were Asian, were pooled. Associations between the variants and the risk of lung cancer were estimated by logistic regression models. All statistical tests were two-sided.

Results Associations between 15q25 and the risk of lung cancer were replicated in white ever-smokers (rs16969968: odds ratio [OR] = 1.26, 95% confidence interval [CI] = 1.21 to 1.32, $P_{trend} = 2 \times 10^{-26}$), and this association was stronger for those diagnosed at younger ages. There was no association in never-smokers or in Asians between either of the 15q25 variants and the risk of lung cancer. For the chromosome 5p15 region, we confirmed statistically significant associations in whites for both rs2736100 (OR = 1.15, 95% CI = 1.10 to 1.20, $P_{trend} = 1 \times 10^{-10}$) and rs402710 (OR = 1.14, 95% CI = 1.09 to 1.19, $P_{trend} = 5 \times 10^{-8}$) and identified similar associations in Asians (rs2736100: OR = 1.23, 95% CI = 1.12 to 1.35, $P_{trend} = 2 \times 10^{-5}$; rs402710: OR = 1.15, 95% CI = 1.04 to 1.27, P_{trend} = .007). The associations between the 5p15 variants and lung cancer differed by histology; odds ratios for rs2736100 were highest in adenocarcinoma and for rs402710 were highest in adenocarcinoma and squamous cell carcinomas. This pattern was observed in both ethnic groups. Neither of the two variants on chromosome 6p21 was associated with the risk of lung cancer.

Conclusions In this international genetic association study of lung cancer, previous associations found in white populations were replicated and new associations were identified in Asian populations. Future genetic studies of lung cancer should include detailed stratification by histology.

J Natl Cancer Inst 2010;102:959–971

Replication of initial genome-wide association findings is considered a gold standard for reporting genotype–phenotype associations. Three human genomic regions at chromosomes 15q25, 5p15, and 6p21 that were found to be associated with susceptibility to lung cancer in genome-wide association studies merit such replication.

The region at 15q24–25.1, which contains three nicotinic acetylcholine receptor subunit genes (*CHRNA5*, *CHRNA3*, and *CHRNB4*), was associated with the risk of lung cancer in three independently conducted genome-wide association studies that gave remarkably consistent results for associations between three

CONTEXT AND CAVEATS

Prior knowledge

Genome-wide association studies that were conducted in white populations have identified three chromosomal regions at 15q25, 5p15, and 6p21 as being associated with the risk of lung cancer. Whether genetic variants at these regions are associated with risk of lung cancer in other populations is unclear.

Study design

A coordinated genotyping study of six single-nucleotide polymorphisms at these three chromosomal regions using data from 21 independent case–control studies that included Asians studies and were not included in previous genome-wide association studies.

Contribution

The 15p25 locus–risk of lung cancer association in whites was replicated, but there was no association between this locus and the risk of lung cancer in white lifetime never-smokers. There was no association between 15q25 and the risk of lung cancer in Asians. The chromosome 5p15 locus–risk of lung cancer association was replicated in whites, and a similar association was found in Asians. In both whites and Asians, the two variants in 5p15 were more strongly associated with adenocarcinoma than with other histology groups. Chromosome 6p21 was not associated with the risk of lung cancer.

Implications

Future genetic studies of lung cancer should include detailed stratification by histology.

Limitations

Some of the variants at chromosome 15q25 had low minor allele frequencies in Asians. Replication of variants in Asians that were originally identified in studies of whites may not be relevant. The fact that different studies with different genotyping protocols were included could have led to heterogeneity.

From the Editors

single-nucleotide polymorphisms (SNPs) at this locus and the risk of lung cancer: The MD Anderson Cancer Center study reported an odds ratio (OR) of 1.32 ($P = 10^{-17}$) for rs8034191 (1), the International Agency for Research on Cancer (IARC) study (2) reported an odds ratio of 1.30 ($P = 10^{-20}$) for rs8034191 and rs16969968, and the deCODE study reported an odds ratio of 1.31 ($P = 10^{-8}$) for rs1051730 (3). All of these three SNPs (rs8034191, rs16969968, and rs1051730) are in strong linkage disequilibrium.

Subsequent meta-analyses identified another putative causative region at 5p15.33 (4,5). This region contains two genes: the human telomerase reverse transcriptase gene (*TERT*) and cleft lip and palate transmembrane 1–like gene (*CLPTM1L*). The IARC (4) and the Institute of Cancer Research (ICR) and MD Anderson groups (5) reported that two different SNPs at 5p15.33 (rs402710 and rs401681, respectively), which are in strong linkage disequilibrium ($D' = 1.00$, $r^2 = .66$), are associated with the risk of lung cancer (IARC study: $P = 2 \times 10^{-7}$; ICR and MD Anderson groups: $P = 8 \times 10^{-9}$). The IARC group also identified a second SNP, rs2736100, that was associated with the risk of lung cancer ($P = 4 \times 10^{-6}$). A report from the deCODE group (6) provided evidence that the 5p15.33 region may be a susceptibility locus for multiple cancer types and also reported associations between risk of lung cancer and two potential susceptibility alleles.

The third region that has been implicated by genome-wide association studies in susceptibility to lung cancer is the HLA region at chromosome 6p21. Hung et al. (2) presented evidence for an association between the SNP rs4324798 at 6p21 and the risk of lung cancer ($P = 4 \times 10^{-7}$). Wang et al. (5) identified two other SNPs that were statistically significantly associated with risk of lung cancer and that mapped to this region: rs3117582 ($P = 5 \times 10^{-10}$) and rs9295740 ($P = 4 \times 10^{-7}$).

We aimed to replicate these findings in a large sample size dataset because there is still no consensus about the relative impact with respect to risk of lung cancer of the chromosome 15q25 locus on smoking behavior vs a direct lung carcinogenic effect. In addition, the newly identified susceptibility loci on 5p15 and 6p21 require further investigation in a larger sample size and in different ethnic groups. It is also important to evaluate effect modification by sex, age at cancer diagnosis, and family history, as well as by histological classification.

The International Lung Cancer Consortium (ILCCO) was established in 2004 with the aim of sharing comparable data from ongoing case–control and cohort studies of lung cancer. The overall objectives of the consortium are to share the data to increase statistical power, especially for subgroup analyses, reduce duplication of research efforts, replicate novel findings, and realize substantial cost savings through large collaborative efforts. Details of how the consortium was established have been published previously (7). With the aim of replicating association findings concerning these variants and the risk of lung cancer with sufficient statistical power for analysis of specific subgroups, we invited the principle investigators of all case–control studies from ILCCO to conduct genotyping of their lung cancer case patients and control subjects of European and Asian ancestry for two variants at the 15q25 locus (rs8034191 and rs16969968), two variants at 5p15 (rs402710 and rs2736100), and two variants at 6p21 (rs4324798 and rs2256543, the latter of which was the second most statistically significant SNP on chromosome 6 from the IARC genome-wide association study). For studies that were conducted in Asian populations, we selected three additional variants in the 15q25 region for genotyping (rs12914385, rs1317286, and rs931794) and the variants in 6p21 were not genotyped because of their low prevalence in these populations (according to the HapMap genome browser, www.hapmap.org).

Materials and Methods

Study Population

Twenty-one of the 52 case–control studies from the ILCCO participated in this pooled analysis. Of these studies, nine were conducted in North America, eight in Europe, and four in Asia. The study designs are briefly outlined in Table 1, and some of them have been described in more detail previously (6,8–19). The studies are referred to here either by the study location or the name of the coordinating institution.

The Singapore study included only women; the MD Anderson, Norwegian, and French studies included only ever-smokers. All

Table 1. Summary of the participating studies from the International Lung Cancer Consortium*

Study reference	Coordinating institution	Study name	Study location	Recruitment period	Eligibility	Source of control subjects	No. of case patients†	No. of control subjects†
			Studies that enrolled white subjects					
—	MD Anderson Cancer Center	MD Anderson	TX, United States	1992–2004	Only ever-smokers	Hospital	709	629
—	Karmanos Cancer Institute (KCI), Wayne State University	KCI	MI, United States	1988–2007	No restriction	Population	575	860
—	University of Hawaii	Hawaii	HI, United States	1992–1997	26–79 years old	Population	138	175
—	Mayo Clinic	Mayo	MN, United States	1997–2006	18 years or older	Hospital	1644	1021
8	Norris Cotton Cancer Center, Dartmouth Medical School	NELCS	NH, United States	2005–2008	30–74 years old	Population	228	162
9	Penn State University College of Medicine	Penn State University	FL, United States	2000–2003	18–79 years old, within 1 year of diagnosis, no previous cancer diagnosis	Community	447	733
10	University of California Los Angeles (UCLA)	UCLA	CA, United States	1999–2004	18–65 years old	Population	319	581
11	University of California, San Francisco (UCSF)	UCSF	San Francisco, CA, United States	1998–2003, 2005–2009	18 years or older	Population and community	1804	558
12	National Institute of Occupational Health	Norway	Norway	1986–2005	Current smokers or quit smoking within past 5 years	Population	443	436
—	University of Sheffield	Sheffield	United Kingdom	2005–2009	Diagnosed before age 61 years or reported family history of lung cancer	Population recruited through family	114	133
—	INSERM U794	France	France	1987–1992	Only ever-smokers		135	146
13	Helmholtz Center Munich	German multicenter	Munich, Göttingen, Germany	2000–2008	LUCY: 18–51 years old	Population	1839	3336
14	University of Göttingen Medical School		Munich, Germany	1990–1998	KORA: all ages			
15	German Cancer Research Center (DKFZ)		Heidelberg, Germany	1997–2007	DKFZ: 18 years or older			
16	German Cancer Research Center (DKFZ)		Heidelberg, Germany	1994–1998	EPIC: 18 years or older			
17	German Cancer Research Center (DKFZ)	Saarland	Saarland, Germany	2000–2003	50–75 years old	Population	198	203
—	University Hospital of Cologne	Cologne	Cologne, Germany	2005–2008	No restriction	Hospital	450	327
6	Division of Medical Oncology, University Hospital	Spain	Zaragoza, Spain	2006–2008	No restriction	Hospital	350	1227
6	Radboud University Nijmegen Medical Centre	The Netherlands	The Netherlands	2008	18–75 years old	Population	396	2068
—	CHS National Cancer Control Center at Carmel Medical Center and Technion	Israel	Haifa, Israel	2007–2009	No restriction	Population	212	197
			Studies that enrolled Asian subjects					
10	UCLA	UCLA	California, United States	1999–2004	18–65 years old	Population	58	53
—	University of Hawaii	Hawaii	HI, United States	1992–1997	26–79 years old	Population	100	170

(Table continues)

Table 1 (Continued).

Study reference	Coordinating institution	Study name	Study location	Recruitment period	Eligibility	Source of control subjects	No. of case patients†	No. of control subjects†
—	Samuel Lunenfeld Research Institute	Toronto	ON, Canada	1997–2002	Residents of greater Toronto area	Population and hospital	65	98
18	Seoul National University	Seoul	Korea	2001–2008	No restriction	Hospital	271	276
—	National University of Singapore	Singapore	Singapore	2005–2007	Only women	Hospital	484	813
19	Aichi Cancer Center	Aichi	Japan	2000–2005	20–79 years old	Hospital	716	716
Total							11645	14954

* — = Unpublished; DZFZ = German Cancer Research Center; EPIC = European Prospective Investigation into Cancer and Nutrition; INSERM = National Institute of Health and Medical Research; LUCY = Lung Cancer in the Young; KORA = Cooperative Health Research in the Augsburg Region; NELCS = New England Lung Cancer Study.

† Maximum number of case patients and control subjects of white and Asian ethnic groups with DNA samples.

studies had detailed information on histology that was based on International Classification of Diseases codes or pathology reports. All studies included incident cases of lung cancer. In most of the studies, the control subjects were frequency matched to the case patients on age and sex; some studies also matched on ethnicity (Hawaii and Canadian studies), place of residence (UCLA study), or smoking status (Norway and MD Anderson studies). Written informed consent was obtained from all study subjects, and the investigations were approved by the institutional review boards at each study center. Only individuals who reported white or Asian ethnicity were included in this analysis of 11645 lung cancer case patients and 14954 control subjects, of whom 85% were white and 15% were Asian (Table 1).

Genotyping and Quality Control

Genotyping from genomic DNA isolated from blood sample or saliva (the extraction technique for each center is available upon request) was performed locally at the participating centers using TaqMan probes (Applied Biosystems, Foster City, CA) (probe and primer sequences are provided in Supplementary Table 1 [available online]) that were supplied by IARC with the following exceptions: Two studies (Toronto and France) used genotyping data that were obtained from HumanHap300 BeadChips (Illumina, San Diego, CA), the German multicenter and Saarland studies used the iPLEX assay (Sequenom, San Diego, CA), and two studies (Spain and the Netherlands) performed genotyping with the use of the Centaurus platform (Nanogen, San Diego, CA). All genotyping assays were performed according to the manufacturer's protocol. The quality of the Centaurus assays was evaluated by genotyping each assay in the HapMap CEU sample, which comprises Utah residents with ancestry from Northern and Western Europe (www.hapmap.org), and comparing the results with the HapMap publicly released data. Assays with a mismatch rate greater than 1.5% were not included in the statistical analysis. Standardized quality control procedures were applied in all centers that used TaqMan or iPLEX assays: Each center genotyped a generic series of 90 standard DNA samples (from either the HapMap CEU sample or a generic series from IARC) in their local genotyping facility. The genotype concordance across studies was subsequently computed for each genotyping assay. When more than one discrepancy between the genotypes obtained from the local genotyping technique and the HapMap publicly available genotypes or the IARC generic series genotypes for a variant was found in a study, that study was excluded from the analysis of that variant (Supplementary Table 2, available online). The average genotype completion rate per SNP varied from 97.1% to 99.6% in the pooled data, and all genotype completion rates per study were greater than 90% for each variant.

We used a χ^2 test with 1 *df* to verify that the allele distributions for each SNP were in Hardy–Weinberg equilibrium within each study and separately among white control subjects and Asian control subjects. A Bonferroni correction for multiple tests was applied for the Hardy–Weinberg equilibrium test and gave a *P* value of .0005 as the cutoff for statistical significance (based on approximately 100 independent tests carried out). No deviation from Hardy–Weinberg equilibrium was observed (Supplementary Table 3, available online).

Statistical Analysis

We used unconditional logistic regression to estimate odds ratios and 95% confidence intervals (CIs). The heterozygous and homozygous carriers of the risk allele were each compared with the homozygous carriers of the nonrisk allele. Odds ratios per allele or *P* values for trend were calculated by assuming a log-additive genetic model with 1 *df*. Pooled odd ratios were calculated using individual-level data. Information on demographic variables (age, sex, ethnicity), tobacco exposures, family history of cancer, and histology classification (for case patients) was available. Mean numbers of cigarettes smoked per day were derived from analysis of variance and are adjusted for age, sex, study, and case–control status when appropriate. Ethnicity was self-declared, and only subjects who declared themselves to be white or Asian were included in the analysis. Whites and Asians were analyzed separately. Models were adjusted for potential confounders, including age, sex, study center, and, where appropriate, cumulative tobacco consumption (expressed as pack-years). To evaluate effect modification, we conducted analyses stratified by smoking status (never, former, current), smoking quantity (by 10-pack-year categories), sex, age at lung cancer diagnosis (by 10-year age groups), or family history of cancer among first-degree relatives. We also analyzed associations between genetic variants and the risk of lung cancer by major histological subtypes (squamous cell carcinoma, adenocarcinoma, small-cell carcinoma, and large-cell carcinoma). Heterogeneity of odds ratios across the studies and across the stratification groups was assessed by using the Cochran Q test.

All analyses were conducted with SAS software (version 9.1; SAS Institute, Cary, NC). All statistical tests were two-sided, and statistical significance required a *P* value of .05 or less.

Results

Among the white subjects, 57% were male, whereas among the Asian subjects, a slight majority was female (50% of the case subjects and 58% of the control subjects) (Table 2). We observed a higher prevalence of never-smoking lung cancer case patients among Asians (40%) than among whites (10%). This difference is mainly because of the Singapore study, which included only women, of whom 79% were never-smokers.

Table 3 summarizes the pooled estimates of the main effects for each variant. In whites, both of the variants at 15q25 were strongly associated with the risk of lung cancer and exhibited similar odds ratios in heterozygotes, in homozygotes, and per allele. The strongest association of the two variants at 15q25 was for rs16969968 (OR = 1.26, 95% CI = 1.21 to 1.32, $P_{trend} = 2 \times 10^{-26}$). We also noted associations between the two variants located on chromosome 5p15 and the risk of lung cancer (rs2736100: OR = 1.15, 95% CI = 1.10 to 1.20, $P_{trend} = 1 \times 10^{-10}$; rs402710: OR = 1.14, 95% CI = 1.09 to 1.19, $P_{trend} = 5 \times 10^{-8}$). Among the two variants at 6p21, we observed a statistically significant association between the wild-type allele of rs4324798 and the risk of lung cancer among homozygotes (OR = 1.39, 95% CI = 1.04 to 1.87); however, in the log-additive model, neither variant on this chromosome was associated with the risk of lung cancer. In Asians, the minor allele frequencies of rs16969968 and rs8034191 on chromosome 15q25 were lower than 5% and no association with the risk of lung cancer was observed. None of the other variants selected from this region was associated with risk of lung cancer in this ethnic group. However, for chromosome 5p15, there were statistically significant associations between rs2736100 (OR = 1.23, 95% CI = 1.12 to 1.35, $P_{trend} = 2 \times 10^{-5}$) and rs402710 (OR = 1.15, 95% CI = 1.04 to 1.27, $P_{trend} = .007$) and the risk of lung cancer. No statistically

Table 2. Distribution of selected demographic variables by ethnic group*

	Whites		Asians	
Variable	**Case patients, No. (%)**	**Control subjects, No. (%)**	**Case patients, No. (%)**	**Control subjects, No. (%)**
Sex				
Male	5741 (58)	7325 (57)	838 (50)	902 (42)
Female	4210 (42)	5503 (43)	856 (50)	1224 (58)
Age, y				
<50	1252 (13)	2969 (23)	182 (11)	243 (11)
50–59	2499 (25)	3443 (27)	426 (25)	492 (23)
60–69	3273 (33)	3859 (30)	565 (33)	679 (32)
70–79	2451 (25)	2310 (18)	443 (26)	612 (29)
≥80	476 (5)	247 (2)	78 (5)	100 (5)
Smoking status				
Never	962 (10)	4136 (32)	674 (40)	1270 (60)
Former smoker	4125 (41)	4491 (35)	461 (27)	470 (22)
Current smoker	4644 (47)	3173 (25)	526 (31)	308 (15)
Former or current	134 (1)	455 (4)	23 (1)	20 (1)
Missing	86 (1)	573 (5)	10 (1)	58 (3)
Histology				
Adenocarcinoma	3892 (39)	—	929 (55)	—
Squamous cell	2370 (24)	—	317 (19)	—
Large cell	413 (4)	—	96 (6)	—
Small cell	1235 (12)	—	109 (6)	—
Other or not specified	2041 (21)	—	243 (14)	—
Total	9951	12828	1694	2126

* Numbers of case and control subjects are those included in the analysis. — = not applicable.

Table 3. Summary estimates of the main effects of the selected variants in whites and Asians*

Chromosomal locus and variant	Risk allele	Minor allele frequency†	No. of homozygotes for the common allele‡/No. of heterozygotes for the risk allele/No. of homozygotes for the risk allele		OR§ (95% CI)			P_{trend}‖	$P_{heterogeneity}$¶
			Case patients	Control subjects	Heterozygotes	Homozygotes	Per allele		
Whites									
15q25									
rs16969968	A	0.35	3371/4523/1484	4827/5019/1373	1.33 (1.24 to 1.41)	1.54 (1.41 to 1.69)	1.26 (1.21 to 1.32)	2×10^{-26}	.09
rs8034191	G	0.35	2586/3488/1185	4036/4256/1171	1.33 (1.24 to 1.43)	1.62 (1.47 to 1.79)	1.29 (1.23 to 1.35)	6×10^{-25}	.15
5p15									
rs2736100	C	0.51	1878/4526/2722	2853/5817/3142	1.16 (1.07 to 1.25)	1.32 (1.21 to 1.43)	1.15 (1.10 to 1.20)	1×10^{-10}	.60
rs402710	G	0.65	873/3847/4140	1115/4178/3905	1.16 (1.04 to 1.28)	1.30 (1.18 to 1.45)	1.14 (1.09 to 1.19)	5×10^{-8}	.73
6p									
rs2256543	A	0.43	2898/4519/1803	3860/5813/2260	1.03 (0.96 to 1.10)	1.07 (0.98 to 1.16)	1.03 (0.99 to 1.08)	.14	.92
rs4324798	A	0.08	8066/1630/111	10580/1911/99	1.04 (0.96 to 1.12)	1.39 (1.04 to 1.87)	1.07 (0.99 to 1.14)	.07	.11
Asians									
15q25									
rs16969968	A	0.03	1591/98/2	1986/125/5	0.98 (0.75 to 1.30)	0.44 (0.08 to 2.31)	0.94 (0.73 to 1.23)	.67	.07
rs8034191	G	0.03	1583/104/3	1992/122/3	1.06 (0.81 to 1.40)	1.06 (0.21 to 5.36)	1.06 (0.82 to 1.37)	.66	.09
rs12914385	T	0.30	728/647/148	584/762/177	1.05 (0.91 to 1.21)	1.04 (0.81 to 1.32)	1.03 (0.93 to 1.14)	.58	.10
rs1317286	G	0.10	1223/291/13	1521/313/22	1.18 (0.99 to 1.41)	0.73 (0.36 to 1.47)	1.10 (0.94 to 1.30)	.23	.10
rs931794	G	0.37	591/721/213	764/828/264	1.12 (0.96 to 1.29)	1.01 (0.82 to 1.25)	1.03 (0.93 to 1.14)	.54	.10
5p15									
rs2736100	C	0.39	538/836/312	775/1014/312	1.24 (1.07 to 1.43)	1.51 (1.24 to 1.83)	1.23 (1.12 to 1.35)	2×10^{-5}	.32
rs402710	G	0.68	144/694/842	219/917/981	1.15 (0.91 to 1.46)	1.32 (1.05 to 1.66)	1.15 (1.04 to 1.27)	.007	.22

* CI = confidence interval; OR = odds ratio.

† Risk allele frequencies were calculated among control subjects.

‡ Referent group for the logistic regression model.

§ Adjusted for age, sex, and study.

‖ P_{trend} (two-sided) was derived from a log-additive model.

¶ *P* for heterogeneity by study (two-sided) was derived from the Cochran Q test.

significant heterogeneity by study was observed. The study-specific odds ratios are presented in Supplementary Figures 1 and 2 (available online).

We conducted stratified analyses of the chromosome 15q25 variants in whites (Figure 1). Because linkage disequilibrium between rs16969968 and rs8034191 was high (D' = .95, r^2 = .88), we reported the results only for rs16969968 (results for rs8034191 were similar and are not shown). Among never-smokers, there was no association between rs16969968 and the risk of lung cancer (OR = 1.02, 95% CI = 0.91 to 1.14). Among ever-smokers, this association was statistically significantly stronger in current smokers (OR = 1.39, 95% CI = 1.29 to 1.50) than in former smokers (OR = 1.27, 95% CI = 1.18 to 1.37, $P_{heterogeneity} = 1 \times 10^{-5}$). We also noted a higher odds ratio in subjects younger than 50 years compared with older subjects ($P_{heterogeneity} = 9 \times 10^{-4}$). The mean age at lung cancer diagnosis was statistically significantly higher among homozygous carriers of the common allele than among homozygous carriers of the rare allele (62.8 vs 60.7 years, difference = 2.1 years, 95% CI = 1.2 to 3.3 years). There were no statistically significant differences in the odds ratio estimates by histology, pack-years of cumulative tobacco consumption in ever-smokers, or sex. Among subjects with no missing data for pack-years and smoking status, the overall odds ratio adjusted for these two variables was slightly lower than the unadjusted odds ratio (adjusted OR = 1.25, 95% CI = 1.18 to 1.32; unadjusted OR = 1.33, 95% CI = 1.26 to 1.40). A stratified analysis of rs16969968 by family history of lung cancer among first-degree relatives revealed no heterogeneity between subjects with and without a family history of cancer (data not shown).

We also investigated the association between rs16969968 and the risk of lung cancer in the context of smoking intensity in whites and observed a gene dosage effect with the mean number of cigarettes smoked per day (Table 4). Overall, the mean number of cigarettes smoked per day was 20.74 (95% CI = 20.36 to 21.12) among homozygous carriers of the common allele and 23.48 (95% CI = 22.92 to 24.04) among homozygous carriers of the risk allele. Similar trends were observed in case patients and control subjects when analyzed separately.

Figure 2 shows the stratified estimates for rs2736100 and rs402710 at chromosome 5p15. We observed statistically significant heterogeneity by histology for rs2736100 for both whites ($P < .001$) and Asians ($P = .01$), and the risks of adenocarcinomas and large-cell carcinomas were higher than the risks of squamous and small-cell carcinomas. We also observed heterogeneity by histology for rs402710, with stronger associations for adenocarcinomas and large-cell and squamous cell carcinomas than for small-cell

CHRNA5 (rs16969968)	Case subjects	Control subjects	OR	(95% CI)
Heterozygous	4523	5019	1.33	(1.24 to 1.41)
Homozygous	1484	1373	1.54	(1.41 to 1.69)
Co-dominant	9378	11219	1.26	(1.21 to 1.32)
By histology ($P_{heterogeneity}$=.38)				
Adenocarcinomas	3776	11219	1.22	(1.15 to 1.29)
Squamous	2128	11219	1.31	(1.23 to 1.41)
Large cell	384	10896	1.23	(1.06 to 1.43)
Small cell	1106	10597	1.21	(1.10 to 1.33)
By smoking status ($P_{heterogeneity}$<.001)				
Never smokers	922	3706	1.02	(0.91 to 1.14)
Former smokers	3994	4181	1.27	(1.18 to 1.37)
Current smokers	4277	2875	1.39	(1.29 to 1.50)
By pack-years of smoking ($P_{heterogeneity}$=.75)				
>0-<10	453	1591	1.19	(0.99 to 1.43)
10-<20	771	1271	1.22	(1.05 to 1.42)
20-<30	1076	1165	1.22	(1.07 to 1.39)
30-<40	1373	1057	1.28	(1.13 to 1.46)
40-<50	1257	691	1.11	(0.97 to 1.29)
50-<60	919	400	1.29	(1.07 to 1.56)
≥60	2069	693	1.14	(1.00 to 1.30)
By age (P_{trend}=.002)				
<50	1177	2323	1.49	(1.34 to 1.66)
50-60	2356	3084	1.31	(1.21 to 1.42)
60-70	3095	3530	1.19	(1.11 to 1.29)
≥70	2750	2282	1.16	(1.06 to 1.27)
By sex ($P_{heterogeneity}$=.88)				
Men	5264	6445	1.26	(1.19 to 1.33)
Women	4114	4774	1.27	(1.19 to 1.35)

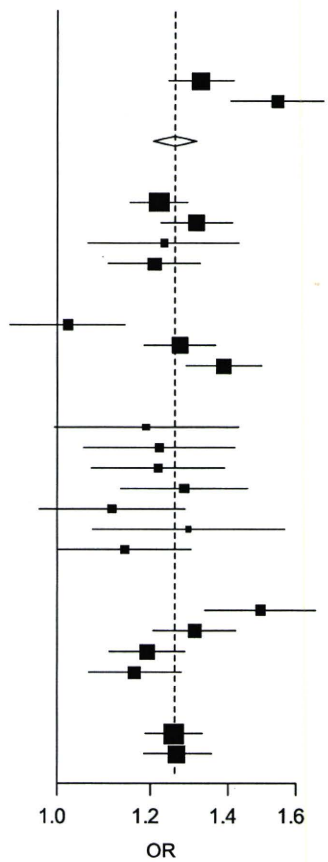

Figure 1. Stratified analysis of the association between rs16969968 and the risk of lung cancer in whites. Except for the odds ratios (ORs) for heterozygous and homozygous effect, odds ratios and 95% confidence intervals (CIs) were derived from the per-allele model. All models are adjusted for age, sex, and study. *P* for heterogeneity was derived from the Cochran Q test. **Squares** represent odds ratios; **size of the square** represents inverse of the variance of the log odds ratio; **horizontal lines** represent 95% confidence intervals; **diamonds** represent summary estimate combining the study-specific estimates with a random-effects model; **solid vertical line** represents an odds ratio of 1; **dashed vertical line** represents the overall odds ratio.

Table 4. Association between rs16969968 and smoking intensity expressed in cigarettes smoked per day in whites*

Genotype	All subjects N	All subjects Mean no. of cigarettes per day (95% CI)	Control subjects N	Control subjects Mean no. of cigarettes per day (95% CI)	Case patients N	Case patients Mean no. of cigarettes per day (95% CI)
rs16969968 (*CHRNA5*)						
GG	5425	20.74 (20.36 to 21.12)	2610	17.99 (17.45 to 18.53)	2815	22.68 (22.14 to 23.22)
GA	6597	21.85 (21.49 to 22.20)	2701	19.20 (18.67 to 19.78)	3896	23.70 (23.22 to 24.18)
AA	2039	23.48 (22.92 to 24.04)	746	20.56 (19.68 to 21.44)	1293	25.56 (24.84 to 26.28)
P_{trend}	7×10^{-19}		6×10^{-9}		5×10^{-12}	

* Mean numbers of cigarettes smoked per day were derived from analysis of variance and are adjusted for age, sex, study, and case–control status when appropriate. P_{trend} was derived from a linear regression with log(number of cigarettes per day) as an outcome. All statistical tests were two-sided.
CI = confidence interval; N = number of subjects included in the analysis.

carcinomas; however, this heterogeneity was statistically significant only in whites ($P_{heterogeneity}$ = .03). We also observed a sex difference for rs2736100, with a stronger association in women than in men ($P_{heterogeneity}$ = .02 for whites and $P_{heterogeneity}$ = .03 for Asians).

Because a higher proportion of adenocarcinomas is usually more frequent in women than in men (in this study, adenocarcinomas were diagnosed in 21% of the female case patients vs 15% of the male case patients), we stratified the analysis of rs2736100 by histology and by sex. For both men and women, the association between the risk of lung cancer and rs2736100 remained stronger for adenocarcinomas than for other histologies (data not shown). Conversely, when the analysis of rs2736100 was restricted to adenocarcinomas, no heterogeneity by sex was observed (data not shown).

We also compared patients with a family history of lung cancer with patients without a family history of lung cancer and observed a statistically significant association between having a family history of lung cancer and carrying the rare allele of rs2736100 (OR = 1.16, 95% CI = 1.03 to 1.32, P_{trend} = .02). Likewise, the risk of lung cancer associated with the variant genotype was higher among subjects with a family history (OR = 1.31, 95% CI = 1.16 to 1.48) than among those without a family history (OR = 1.14, 95% CI = 1.06 to 1.23); however, the difference was not statistically significant ($P_{heterogeneity}$ = .06). No evidence of heterogeneity by family history was observed for rs402710, and no heterogeneity by age group was observed for either chromosome 5p15 variant (data not shown). These results did not change after adjustment for smoking intensity (pack-years) or smoking status.

Although the physical distance between rs2736100 and rs402710 on chromosome 5 is approximately 34 kb, the linkage disequilibrium between these two variants is low (whites: r^2 = .03, D' = .23; Asians: r^2 = .04, D' = .38). We therefore examined the independent associations of rs402710 and rs2736100 with the risk of lung cancer by adjusting the effect of each variant for the other and found that the association remained statistically significant for both variants (data not shown). Moreover, we calculated the association between rs2736100 (C allele) and the risk of lung cancer among those who were homozygous for the common allele of rs402710 (GG genotype) and found an allelic odds ratio of 1.15 (95% CI = 1.08 to 1.22) in whites and 1.14 (95% CI = 1.00 to 1.30) in Asians. Conversely, the allelic odds ratio for rs402710 (G allele) among homozygous carriers of rs2736100 (AA genotype) was 1.09 (95% CI = 1.00 to 1.19) in whites and 1.15 (95% CI = 0.98 to 1.36) in Asians.

When we summed the number of risk alleles for the rs2736100 and rs402710 genotypes, we found a statistically significant odds ratio per risk allele (whites: OR = 1.12, 95% CI = 1.09 to 1.16; Asians: OR = 1.15, 95% CI = 1.08 to 1.23) (Table 5).

We also analyzed the risk of lung cancer associated with the combined genotypes of rs16969968, rs2736100, and rs402710 in whites (Table 6). The odds ratio of lung cancer for homozygous carriers of the three risk variants compared with individuals with no risk alleles was 2.64 (95% CI = 1.86 to 3.74, $P = 4 \times 10^{-8}$).

Discussion

We replicated the results of previous genome-wide association studies for associations between the 15p25 locus, which includes the α5α3β4 family of nicotinic receptor genes, and the risk of lung cancer in whites and obtained an odds ratio of similar magnitude to that previously reported. We also confirmed previous reports (1,21) of no association between this locus and the risk of lung cancer in white lifetime never-smokers. We also observed no association between 15q25 and the risk of lung cancer in Asians. For the chromosome 5p15 region, we confirmed the statistically significant association with risk of lung cancer in whites reported previously and now report an association of similar magnitude in Asians. We also noted a stronger association between the two variants in 5p15 and adenocarcinoma vs other histology groups for both whites and Asians. We did not replicate the association between chromosome 6p21 and the risk of lung cancer that was reported by Hung et al. (2).

This replication study in two distinct ethnic groups represents, to our knowledge, the largest international effort in lung cancer based on independent studies that were not included in previous genome-wide association studies. In addition to replicating associations from the genome-wide association studies in whites, we expanded our analysis to Asian populations because we hypothesized that the different genetic architecture and linkage disequilibrium structure of Asians might elucidate associations with the putative causal variants. The sample size allowed us to analyze individual-level data and ensured that we had adequate statistical power for stratified analyses of associations between variants and the risk of lung cancer based on histology, age at lung cancer diagnosis, smoking status, smoking quantity, family history of lung cancer, and ethnicity.

There is unequivocal evidence that the 15q25 locus is associated with smoking status and nicotine dependence in whites.

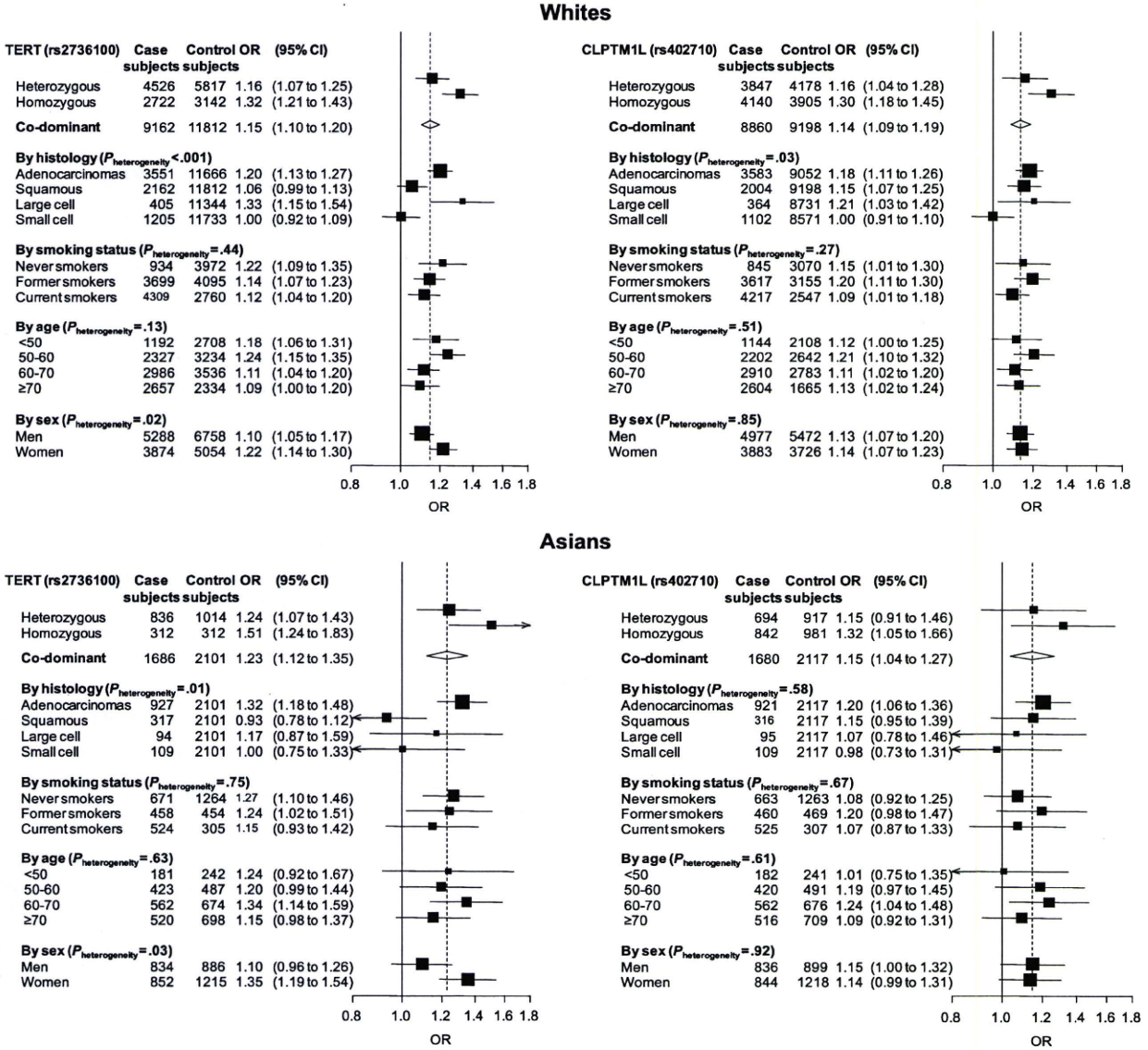

Figure 2. Stratified analysis of associations between rs2736100 and rs402710 and the risk of lung cancer in whites and Asians. Except for the odds ratios (ORs) for heterozygous and homozygous effect, odds ratios and 95% confidence intervals (CIs) were derived from the per-allele model. All models are adjusted for age, sex, and study. *P* for heterogeneity was derived from the Cochran Q test. **Squares** represent odds ratios; **size of the square** represents inverse of the variance of the log odds ratio; **horizontal lines** represent 95% confidence intervals; **diamonds** represent summary estimate combining the study-specific estimates with a random-effects model; **solid vertical lines** represent an odds ratio of 1; **dashed vertical lines** represent the overall odds ratio.

Saccone et al. (20) identified this region in a candidate gene study that compared nicotine-dependent smokers with nondependent smokers who were categorized according to measures derived from the Fagerström Test of Nicotine Dependence. Subsequently, Berrettini et al. (21) identified the same genetic association between chromosome 15q25 and smoking intensity in a comparison of heavy vs light smokers. Finally, in a genome-wide association study on smoking quantity and nicotine dependence, Thorgeirsson et al. (3) found a statistically significant association between the number of cigarettes smoked per day and the 15q25 locus. Nicotine dependence was also statistically significantly associated with the same genetic markers at 15q25. These findings were subsequently confirmed by other groups that used correlated clinical characteristics, such as smoking quantity and heavy vs light smoking groups, in different populations, including community-based populations and alcohol-dependent subjects (22,23). In a genome-wide association study, Caporaso et al. (24) identified multiple SNPs that were statistically significantly associated with the number of cigarettes consumed per day at a *P* value less than .001. They also combined their 15q25 results with data from the three published lung cancer genome-wide association studies and found that rs1051730 was highly statistically significantly associated with the number of cigarettes smoked per day ($P = 5 \times 10^{-32}$), as was rs8034191 ($P = 2 \times 10^{-29}$).

Le Marchand et al. (25) also reported that smokers who carried either the rs1051730 or the rs16969968 variant had higher internal

Table 5. Association between the risk of lung cancer and combined genotypes of rs402710 and rs2736100*

Number of risk alleles	*CLPTM1L* rs402710	*TERT* rs2736100	Whites: Case patients (n = 8140)	Whites: Control subjects (n = 8437)	Whites: OR (95% CI)	Whites: *P*	Asians: Case patients (n = 1674)	Asians: Control subjects (n = 2093)	Asians: OR (95% CI)	Asians: *P*
0	AA	AA	273	368	1.00 (referent)		80	118	1.00 (referent)	
1	AA	AC	371	472	1.07 (0.86 to 1.33)	.55	54	80	1.08 (0.69 to 1.70)	.72
2	AA	CC	141	153	1.25 (0.93 to 1.67)	.13	10	18	0.95 (0.42 to 2.19)	.91
1	AG	AA	837	1040	1.07 (0.89 to 1.30)	.47	241	385	0.95 (0.68 to 1.32)	.75
2	AG	AC	1837	1943	1.25 (1.05 to 1.50)	0.01	354	434	1.27 (0.92 to 1.75)	.14
3	AG	CC	860	867	1.34 (1.10 to 1.62)	.003	96	88	1.73 (1.15 to 2.60)	.009
2	GG	AA	598	656	1.17 (0.96 to 1.43)	.13	213	268	1.2 (0.86 to 1.69)	.29
3	GG	AC	1825	1742	1.36 (1.14 to 1.62)	8×10^{-4}	421	496	1.34 (0.98 to 1.84)	.07
4	GG	CC	1398	1196	1.58 (1.29 to 1.86)	3×10^{-6}	205	206	1.56 (1.10 to 2.20)	.01
0	—	—	273	368	1.00 (referent)		80	118	1.00 (referent)	
1	—	—	1208	1512	1.07 (0.89 to 1.29)	.46	295	465	0.97 (0.70 to 1.34)	.85
2	—	—	2576	2752	1.23 (1.04 to 1.47)	.02	577	720	1.24 (0.91 to 1.68)	.17
3	—	—	2685	2609	1.35 (1.13 to 1.60)	7×10^{-4}	517	584	1.40 (1.02 to 1.91)	.03
4	—	—	1398	1196	1.55 (1.29 to 1.86)	3×10^{-6}	205	206	1.56 (1.10 to 2.20)	.01
Per risk allele					1.12 (1.09 to 1.16)	3×10^{-13}			1.15 (1.08 to 1.23)	7×10^{-6}

* The upper section of the table presents data for the association between different combinations of genotypes for rs402710 and r2736100 and the risk of lung cancer. The lower section of the table presents data for the association between the total number of risk alleles for rs402710 and rs2736100 combined and the risk of lung cancer. *P* values (two-sided) were derived from logistic regression. — = Not applicable; CI = confidence interval; OR = odds ratio.

Table 6. Association between the risk of lung cancer and the number of risk alleles combining genotypes of rs402710, rs2736100, and rs16969968 in whites*

Number of risk alleles	Case patients (n = 7964)	Control subjects (n = 8212)	OR (95% CI)	*P*
0	95	153	1.00 (referent)	
1	551	765	1.16 (0.87 to 1.55)	.32
2	1538	1856	1.30 (0.98 to 1.71)	.06
3	2364	2458	1.53 (1.16 to 2.01)	.003
4	2097	1955	1.72 (1.30 to 2.26)	1×10^{-4}
5	1099	883	1.98 (1.49 to 2.63)	2×10^{-6}
6	220	142	2.64 (1.86 to 3.74)	4×10^{-8}
Per risk allele			1.15 (1.12 to 1.18)	1×10^{-26}

* *P* values (two-sided) were from logistic regression. CI = confidence interval; OR = odds ratio.

doses of nicotine (nicotine equivalents) and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (a tobacco-specific carcinogen) per cigarette smoked compared with smokers who did not carry either variant, indicating that carriers of these variants not only smoke more cigarettes but also smoke more intensely, extracting a greater amount of nicotine and carcinogens per cigarette, compared with noncarriers. rs16969968 causes an amino acid substitution in the neuronal acetylcholine receptor subunit alpha-5. In vitro studies by Bierut et al. (26) have shown that in carriers of the risk allele, the $\alpha4\beta2\alpha5$ receptors exhibit a lower response to an agonist. In a sample of 1050 nicotine-dependent case patients and 879 non–nicotine dependent control subjects, Saccone et al. (27) reported two distinct loci in *CHRNA5–CHRNA3–CHRNB4* gene cluster to be associated with nicotine dependence.

In this study, we report different results in whites and in Asians for the selected variants at chromosome 15q25. In whites, we identified a statistically significant gene dosage effect with the highest reported number of cigarettes smoked per day in both case and control subjects who were carriers of the rs16969968 homozygous mutant genotype. We found no difference in the association between this variant and the risk of lung cancer by histological subtype, sex, or number of pack-years smoked, although current smokers exhibited a slightly higher risk of lung cancer compared with former smokers. However, we did note that the highest overall risk between rs16969968 and the risk of lung cancer was among patients who were diagnosed before the age of 50 years. This finding confirm previous observations from the MD Anderson genome-wide association study that risk estimates for subjects who carried the variant genotype were higher for younger patients and that carriers of the variant genotype exhibited earlier age at lung cancer diagnosis than noncarriers (22). This inverse trend with age may argue for a direct role of this region in lung carcinogenesis. However, in this study, among the 922 white patients who had never smoked, there was no evidence of any association between rs16968869 and the risk of lung cancer, suggesting that active smoking is a necessary cofactor for lung carcinogenesis.

In Asians, we observed no association between the five variants at 15q25 and the risk of lung cancer or the number of cigarettes smoked per day. The lower minor allele frequency of rs16969968 in Asians compared with whites (0.03 vs 0.3; Table 3) and the high proportion of Asian never-smokers (40% of the case patients and

60% of the control subjects) may partially explain these negative findings. We also lack any a priori evidence that the genetic markers at 15q25 that we chose to study are relevant in Asians. However, other studies in Asian populations have reported lung cancer susceptibility loci at chromosome 15q25. For example, a Japanese case–control study (28) reported an association between rs16969968 and the risk of lung cancer (OR = 2.2, 95% CI = 1.5 to 3.4, $P = 1.5 \times 10^{-4}$), and similar associations were observed among never-smokers (OR = 2.4, 95% CI = 1.2 to 4.7, $P = .013$) and ever-smokers (OR = 2.2, 95% CI = 1.1 to 4.1, $P = .016$). Although the minor allele frequency of rs16969968 in the Japanese study population was, as expected, very low (0.015), the proportion of never-smokers among the case patients was lower than that in our study (21% vs 40%). Another large study conducted in China (29) identified statistically significant associations between four novel SNPs in the 15q25 region and age at lung cancer diagnosis and smoking behavior, whereas none of these associations was reported for the 15q25 SNPs that were previously reported in whites. The associations between the four novel SNPs and the risk of lung cancer were substantially stronger in younger age at diagnosis patients, similar to our findings in white populations, and were consistent among never-smokers and smokers (29). The latter observation argues strongly for a role of the 15q25 locus in lung cancer that is independent of smoking behavior in Asians.

This replication analysis provided conclusive evidence for associations between rs2736100 and rs402710 at chromosome 5p15 and susceptibility to lung cancer in both whites and Asians. These associations appeared to be independent because restricting the analysis to homozygous carriers of common alleles of one variant did not alter the association of the other variant. This finding may also suggest the existence of an unknown variant that is in linkage disequilibrium with rs2736100 and rs402710 that captures the effect of both rs2736100 and rs402710. In contrast with the chromosome 15q25 findings, both SNPs at chromosome 5p15 were associated with statistically significant increased risks of lung cancer in never-smokers as well as in ever-smokers, and there were no patterns of association by age at diagnosis or duration of smoking for either ethnic group.

These results were in accordance with those reported by previous genome-wide association studies in a white population. McKay et al. (4) reported odds ratios of 1.14 for rs2736100 and 1.18 for rs402710, whereas Wang et al. (5) reported an odds ratio of 1.14 for rs401681 (in strong linkage disequilibrium with rs402710). Neither of these studies reported heterogeneity by histology, smoking status, age at diagnosis, or sex. The magnitudes of these associations are consistent with our findings. However, associations between these variants and the risk of lung cancer differed by histology, and this finding was consistent in both ethnic groups. In particular, we identified an increased risk for adenocarcinomas for both variants in both whites and Asians and an absence of any risk for small-cell carcinomas. Squamous and large-cell carcinomas gave intermediate results. Another study conducted in Iceland (6) found associations between rs401681, located in the *CLPTM1L* gene, and several smoking-related cancers, including lung cancer ($P = 7 \times 10^{-8}$), as well as cancers of the bladder, prostate, skin, and cervix. They also noted that rs2736098, which is located in the *TERT* gene at 5p15, showed a stronger association with lung cancer ($P = 3 \times 10^{-5}$), bladder and prostate cancers, and basal cell carcinomas.

In a case–control study conducted in a Chinese population, Jin et al. (30) reported that rs2736100 was associated with an increased risk of non–small cell lung carcinomas, with odds ratios of 1.26 (95% CI = 1.05 to 1.51) and 1.31 (95% CI = 1.04 to 1.66) for one and two copies of the variant C allele, respectively. They noted that the association was more prominent among women ($P_{heterogeneity}$ = .044), nonsmokers ($P_{heterogeneity}$ = .054), and patients with adenocarcinoma ($P_{heterogeneity}$ = .058). Our data suggested that the association between rs2736100 and the risk of lung cancer was strongest for adenocarcinoma and that the difference between males and females was at least partially explained by the higher proportion of adenocarcinoma among women.

rs2736100 and rs402710 were also previously found to be associated with other diseases. rs2736100 was found to be associated with glioma susceptibility in two recent genome-wide association studies (31,32) conducted in whites. Another genome-wide association study conducted in a Japanese population found that this SNP was associated with idiopathic pulmonary fibrosis (33).

TERT encodes the enzyme telomerase reverse transcriptase, which is the catalytic subunit of telomerase that adds telomeric repeat sequences onto chromosome ends (34). High expression of telomerase is commonly observed in lung cancer, which suggests that *TERT* may have a critical role in lung tumorigenesis (35–37). An association between short telomeres and an increased risk of cancer has been reported for several types of cancers, including basal cell carcinoma, cancers of the lung, head and neck, bladder, kidney, esophagus, and breast, as well as lymphoma (6,38–42). Rafnar et al. (6) reported that rs401681 and rs2736098 were associated with shorter telomere length in older healthy women but not in younger healthy women and suggested that these variants may lead to an increase in the gradual shortening of telomeres over time. Zhu et al. (43) reported that *TERT* gene amplification is more commonly seen in adenocarcinoma than in squamous cell carcinoma and that overexpression of TERT mRNA is correlated with *TERT* gene amplification in adenocarcinoma but not in squamous cell carcinoma. Zhu et al. hypothesized that overexpression of TERT mRNA in adenocarcinomas is largely due to *TERT* amplification, whereas in other lung tumor types, it is mainly controlled by epigenetic factors. Our data are in accordance with these findings.

Concerning the variants at chromosome 6p21, the associations with rs2256543 and rs4324798 were not replicated in whites in this study. It should be noted that for this replication analysis, we selected the two most statistically significant variants from the IARC genome-wide association study (4). The borderline statistically significant associations we reported may indicate that other SNPs in this region may be better candidates than the ones we selected. For example, Wang et al. (5) reported associations between two 6p21 variants—rs3117582 in *BAT3* and rs3131379 in *MSH5*—that are located approximately 3 Mb away from the 6p21 variants analyzed in this study. These two SNPs are highly correlated (r^2 = .99), and the genes in which they reside are strong candidates for lung cancer susceptibility loci: *BAT3* is implicated in apoptosis (44), and *MSH5* is involved in DNA mismatch repair (45). Further investigation of this region is warranted.

This study has several limitations. First, we selected variants that were found to be associated with lung cancer in genome-wide association studies that were conducted in whites. Replication of these variants in Asians may not be relevant. As we reported in this study, the variants we selected at chromosome 15q25 (rs16969968 and rs8034191) had a very low minor allele frequency in Asians. Three other variants at 15q25 were therefore selected for this population, albeit with a lack of a priori evidence, and were also not replicated. Second, many different studies with different genotyping protocols were included in this study, which could have lead to heterogeneity. However, we implemented stringent interlaboratory quality control procedures in all centers and found no evidence of any such heterogeneity by study.

In conclusion, this analysis exemplifies the timely and cost-effective contributions that international consortia can provide to genome-wide association replication studies. Our observations of heterogeneity by histology of associations between variants at 5p15 and the risk of lung cancer are particularly notable and indicate that further study of the role of this locus in lung cancer development is warranted. Future lung cancer genome-wide association studies should routinely include histology-specific analyses.

Supplementary Data

Supplementary data can be found at http://www.jnci.oxfordjournals.org/.

References

1. 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(5):616–622.
2. 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(7187):633–637.
3. Thorgeirsson TE, Geller F, Sulem P, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature.* 2008;452(7187):638–642.
4. McKay JD, Hung RJ, Gaborieau V, et al. Lung cancer susceptibility locus at 5p15.33. *Nat Genet.* 2008;40(12):1404–1406.
5. Wang Y, Broderick P, Webb E, et al. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat Genet.* 2008;40(12):1407–1409.
6. Rafnar T, Sulem P, Stacey SN, et al. Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. *Nat Genet.* 2009;41(2):221–227.
7. Hung RJ, Christiani DC, Risch A, et al. International Lung Cancer Consortium: pooled analysis of sequence variants in DNA repair and cell cycle pathways. *Cancer Epidemiol Biomarkers Prev.* 2008;17(11):3081–3089.
8. Heck J, Andrew A, Onega T, et al. Lung cancer in a US population with low to moderate arsenic exposure. *Environ Health Perspect.* 2009;117(11):1718–1723.
9. Gallagher CJ, Muscat JE, Hicks AN, et al. The UDP-glucuronosyltransferase 2B17 gene deletion polymorphism: sex-specific association with urinary 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol glucuronidation phenotype and risk for lung cancer. *Cancer Epidemiol Biomarkers Prev.* 2007;16(4):823–828.
10. Cui Y, Morgenstern H, Greenland S, et al. Polymorphism of Xeroderma Pigmentosum group G and the risk of lung cancer and squamous cell carcinomas of the oropharynx, larynx and esophagus. *Int J Cancer.* 2006;118(3):714–720.
11. Wrensch MR, Miike R, Sison JD, et al. CYP1A1 variants and smoking-related lung cancer in San Francisco Bay area Latinos and African Americans. *Int J Cancer.* 2005;113(1):141–147.
12. Zienolddiny S, Campa D, Lind H, et al. Polymorphisms of DNA repair genes and risk of non-small cell lung cancer. *Carcinogenesis.* 2006;27(3):560–567.
13. Sauter W, Rosenberger A, Beckmann L, et al. Matrix metalloproteinase 1 (MMP1) is associated with early-onset lung cancer. *Cancer Epidemiol Biomarkers Prev.* 2008;17(5):1127–1135.
14. Wichmann HE, Gieger C, Illig T. KORA-gen—resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen.* 2005;67(suppl 1):S26–S30.
15. Wichmann HE, Rosario AS, Heid IM, Kreuzer M, Heinrich J, Kreienbrock L. Increased lung cancer risk due to residential radon in a pooled and extended analysis of studies in Germany. *Health Phys.* 2005;88(1):71–79.
16. Boeing H, Wahrendorf J, Becker N. EPIC-Germany—a source for studies into diet and risk of chronic diseases. European Investigation into Cancer and Nutrition. *Ann Nutr Metab.* 1999;43(4):195–204.
17. Breitling LP, Raum E, Muller H, Rothenbacher D, Brenner H. Synergism between smoking and alcohol consumption with respect to serum gamma-glutamyltransferase. *Hepatology.* 2009;49(3):802–808.
18. Kim JH, Kim H, Lee KY, et al. Genetic polymorphisms of ataxia telangiectasia mutated affect lung cancer risk. *Hum Mol Genet.* 2006;15(7):1181–1186.
19. Suzuki T, Matsuo K, Hiraki A, et al. Impact of one-carbon metabolism-related gene polymorphisms on risk of lung cancer in Japan: a case control study. *Carcinogenesis.* 2007;28(8):1718–1725.
20. Saccone SF, Hinrichs AL, Saccone NL, et al. Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. *Hum Mol Genet.* 2007;16(1):36–49.
21. Berrettini W, Yuan X, Tozzi F, et al. Alpha-5/alpha-3 nicotinic receptor subunit alleles increase risk for heavy smoking. *Mol Psychiatry.* 2008;13(4):368–373.
22. Spitz MR, Amos CI, Dong Q, Lin J, Wu X. The CHRNA5-A3 region on chromosome 15q24-25.1 is a risk factor both for nicotine dependence and for lung cancer. *J Natl Cancer Inst.* 2008;100(21):1552–1556.
23. Weiss RB, Baker TB, Cannon DS, et al. A candidate gene approach identifies the CHRNA5-A3-B4 region as a risk factor for age-dependent nicotine addiction. *PLoS Genet.* 2008;4(7):e1000125.
24. Caporaso N, Gu F, Chatterjee N, et al. Genome-wide and candidate gene association study of cigarette smoking behaviors. *PLoS ONE.* 2009;4(2):e4653.
25. Le Marchand ML, Derby KS, Murphy SE, et al. Smokers with the CHRNA lung cancer-associated variants are exposed to higher levels of nicotine equivalents and a carcinogenic tobacco-specific nitrosamine. *Cancer Res.* 2008;68(22):9137–9140.
26. Bierut LJ, Madden PA, Breslau N, et al. Novel genes identified in a high-density genome wide association study for nicotine. *Hum Mol Genet.* 2007;16(1):24–35.
27. Saccone NL, Saccone SF, Hinrichs AL, et al. Multiple distinct risk loci for nicotine dependence identified by dense coverage of the complete family of nicotinic receptor subunit (CHRN) genes. *Am J Med Genet B Neuropsychiatr Genet.* 2009;150B(4):453–466.
28. Shiraishi K, Kohno T, Kunitoh H, et al. Contribution of nicotine acetylcholine receptor polymorphisms to lung cancer risk in a smoking-independent manner in the Japanese. *Carcinogenesis.* 2009;30(1):65–70.
29. Wu C, Hu Z, Yu D, et al. Genetic variants on chromosome 15q25 associated with lung cancer risk in Chinese populations. *Cancer Res.* 2009;69(12):5065–5072.
30. Jin G, Xu L, Shu Y, et al. Common genetic variants on 5p15.33 contribute to risk of lung adenocarcinoma in a Chinese population. *Carcinogenesis.* 2009;30(6):987–990.
31. Shete S, Hosking FJ, Robertson LB, et al. Genome-wide association study identifies five susceptibility loci for glioma. *Nat Genet.* 2009;41(8):899–904.
32. Wrensch M, Jenkins RB, Chang JS, et al. Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. *Nat Genet.* 2009;41(8):905–908.
33. Mushiroda T, Wattanapokayakit S, Takahashi A, et al. A genome-wide association study identifies an association of a common variant in TERT

with susceptibility to idiopathic pulmonary fibrosis. *J Med Genet.* 2008; 45(10):654–656.
34. Janknecht R. On the road to immortality: hTERT upregulation in cancer cells. *FEBS Lett.* 2004;564(1–2):9–13.
35. Lantuejoul S, Salon C, Soria JC, Brambilla E. Telomerase expression in lung preneoplasia and neoplasia. *Int J Cancer.* 2007;120(9):1835–1841.
36. Mavrogiannou E, Strati A, Stathopoulou A, Tsaroucha EG, Kaklamanis L, Lianidou ES. Real-time RT-PCR quantification of human telomerase reverse transcriptase splice variants in tumor cell lines and non-small cell lung cancer. *Clin Chem.* 2007;53(1):53–61.
37. Wu TC, Lin P, Hsu CP, et al. Loss of telomerase activity may be a potential favorable prognostic marker in lung carcinomas. *Lung Cancer.* 2003;41(2):163–169.
38. Han J, Qureshi AA, Prescott J, et al. A prospective study of telomere length and the risk of skin cancer. *J Invest Dermatol.* 2009;129(2): 415–421.
39. Risques RA, Vaughan TL, Li X, et al. Leukocyte telomere length predicts cancer risk in Barrett's esophagus. *Cancer Epidemiol Biomarkers Prev.* 2007;16(12):2649–2655.
40. Shen J, Terry MB, Gurvich I, Liao Y, Senie RT, Santella RM. Short telomere length and breast cancer risk: a study in sister sets. *Cancer Res.* 2007;67(11):5538–5544.
41. Widmann TA, Herrmann M, Taha N, Konig J, Pfreundschuh M. Short telomeres in aggressive non-Hodgkin's lymphoma as a risk factor in lymphomagenesis. *Exp Hematol.* 2007;35(6):939–946.
42. Wu X, Amos CI, Zhu Y, et al. Telomere dysfunction: a potential cancer predisposition factor. *J Natl Cancer Inst.* 2003;95(16):1211–1218.
43. Zhu CQ, Cutz JC, Liu N, et al. Amplification of telomerase (hTERT) gene is a poor prognostic marker in non-small-cell lung cancer. *Br J Cancer.* 2006;94(10):1452–1459.
44. 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(7):848–861.
45. Her C, Zhao N, Wu X, Tompkins JD. MutS homologues hMSH4 and hMSH5: diverse functional implications in humans. *Front Biosci.* 2007;12:905–911.

Funding

US National Institutes of Health (NIH); National Cancer Institute (NCI) (R03 CA133939-01 and R01 CA092039).

The individual studies were funded by the following sources: MD Anderson study: NCI (CA55769, CA127219, CA121197, and CA133996); German multicenter study: Lung Cancer in the Young study was partly funded by the Deutsche Forschungsgemeinschaft (DFG; BI 576/2-1 and BI 576/2-2), the genome-wide study was funded by the Helmholtz Association, Germany; the Heidelberg lung cancer study was supported by the Deutsche Krebshilfe; the European Prospective Investigation into Cancer and Nutrition-Heidelberg study was funded by "Europe Against Cancer" Programme of the European Commission (Santé et protection des consommateurs [SANCO]); German Cancer Aid; German Cancer Research Center; German Federal Ministry of Education and Research; Mayo Clinic study: NCI (CA77118, CA80127, and CA84354); University of California, San Francisco study: National Institute of Environmental Health Sciences (ES06717); Aichi study: Scientific Research grant from the Ministry of Education, Science, Sports, Culture and Technology of Japan (17015052) and grant for the Third-Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labor and Welfare of Japan (H20-002); Singapore study: National Medical Research Council, Singapore (NMRC/1075/2006); Norwegian study: The Norwegian Cancer Society; UCLA study: NIH (DA11386, CA90833, CA09142, and ES011667) and the Alper Research Center for Environmental Genomics of the UCLA's Jonsson Comprehensive Cancer Center; New England Lung Cancer study: National Center for Research Resources, a component of the NIH (P20RR018787); Israel study: United States–Israel Binational Science Foundation (BSF) (2003159); Penn State study: NCI (PO1 CA68384, K07 CA104231, and K99CA131477) and PA-DOH 4100038714 and PA-DOH 4100038715 from the Pennsylvania Department of Health; Saarland study: Baden-Württemberg Ministry of Research, Science and Arts; Cologne study: this study was funded by the Helmholtz Association, Germany, through VH.VI-143 and by the Monika Kutzner Stiftung; University of Hawaii Study: NCI (CA55874 and CA85997); Toronto study: Canadian Cancer Society (20214).

Notes

K. Stefansson owns stock in deCODE Genetics, which produces genetic risk assessment tests for common diseases and traits. Neither deCODE Genetics nor the study sponsors played a role in the design of the study, analysis or interpretation of the data, the writing of the manuscript, or the decision to submit the manuscript for publication.

Affiliations of authors: Genetic Epidemiology Group, International Agency for Research on Cancer, Lyon, France (TT, JDM, PB); Environmental Epidemiology of Cancer, CESP Centre for Research in Epidemiology and Population Health, INSERM, U1018, Villejuif, France (TT); Samuel Lunenfeld Research Institute, Toronto, ON, Canada (RJH); Department of Epidemiology, University of Texas, MD Anderson Cancer Center, Houston, TX (CIA, XW, MS); Department of Genetic Epidemiology, University of Göttingen, Göttingen, Germany (HBi, ARo); Institute of Epidemiology, Helmholtz Center Munich, Neuherberg, Germany (WS, H-EW); Institute of Epidemiology, German Research Center for Environmental Health, Neuherberg, Germany (TI); Division of Epigenomics and Cancer Risk Factors (ARi) and Division of Clinical Epidemiology and Aging Research (HM, HBr), German Cancer Research Center (DKFZ), Heidelberg, Germany; Thoraxklinik Heidelberg, University of Heidelberg, Heidelberg, Germany (HD); Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany (RK); Department of Health Sciences Research, Mayo Clinic Cancer Center, Rochester, MN (PY, RJ); Departments of Neurological Surgery and Epidemiology and Biostatistics, University of California, San Francisco, CA (JKW, MW, HH); Departments of Community Health and Pathology and Laboratory Medicine, Brown University, Providence, RI (KTK); Aichi Cancer Center Research Institute, Nagoya, Japan (KM, KT); Karmanos Cancer Institute, Wayne State University School of Medicine, Detroit, MI (AGS, AW); Department of Community, Occupational and Family Medicine, National University of Singapore, Singapore, Singapore (AS, CY); LIMES (Life and Medical Sciences Bonn), Genomics and Immunoregulation, University of Bonn, Bonn, Germany (AS-J); Cologne Center for Genomics and the Institute for Genetics, University of Cologne, Cologne, Germany (PN); Lung Clinic, Hospital of Cologne, Cologne, Germany (ES); Department of Internal Medicine, Center for Integrated Oncology, University Hospital Cologne, Cologne, Germany (JW); Penn State College of Medicine, Hershey, PA (PL, JEM, CG, Y-CH); Department for the Chemical and Biological Work Environment, National Institute of Occupational Health, Oslo, Norway (SZ, AH); Department of Pulmonary Diseases (HFMvdH) and Department of Epidemiology, Biostatistics and HTA (LAK), Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands; Division of Medical Oncology, University Hospital, Zaragoza, Spain (DI); Division of Medical Oncology, University Hospital, University of Zaragoza, and Health Sciences Institute, Nanotechnology Institute of Aragon, Zaragoza, Spain (JIM); deCODE Genetics, Reykjavik, Iceland (TR, KS); Faculty of Medicine, University of Iceland, Reykjavik, Iceland (KS); Department of Epidemiology, University of California Los Angeles, Los Angeles, CA (Z-FZ, S-CC); Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Korea (JHK, Y-CH); Institute of Environmental Medicine, Seoul National University Medical Research Center, Seoul, Korea (JHK, Y-CH); Unit of Nutrition, Environment and Cancer, Cancer Epidemiology Research Program, Catalan Institute of Oncology, Barcelona, Spain (EJD); Section of Biostatistics and Epidemiology, Department of Community and Family Medicine, Norris Cotton Cancer Center, Dartmouth Medical School, Lebanon, NH (EJD, ASA); CHS National Cancer Control Center and Department of Community Medicine and Epidemiology, Carmel Medical Center and Technion, Haifa, Israel (FL, GR); Cancer Research Center of Hawaii, University of Hawaii, Honolulu, HI (LLM); INSERM, U946, Fondation Jean Dausset/CEPH, Paris, France (SB); CNRS FRE2939, Institut Gustave Roussy, Villejuif, France (SB); Geneva Cancer Registry, Geneva, Switzerland (CB); Department of Oncology (XX) and School of Health and Related Research (MDT), Faculty of Medicine, Dentistry and Health, University of Sheffield, Sheffield, UK; Cancer Care Ontario, Toronto, ON, Canada (JM); Ontario Cancer Institute, Princess Margaret Hospital, Toronto, ON, Canada (GL).

DOI:10.1158/1055-9965.EPI-07-2786

CYP1A1, *GSTM1*, and *GSTT1* Polymorphisms, Smoking, and Lung Cancer Risk in a Pooled Analysis among Asian Populations

Kyoung-Mu Lee,[1,7] Daehee Kang,[1,2] Margie L. Clapper,[3] Magnus Ingelman-Sundberg,[4] Masko Ono-Kihara,[5] Chikako Kiyohara,[6] Shen Min,[7] Qing Lan,[7] Loic Le Marchand,[8] Pinpin Lin,[9] Maria Li Lung,[10] Hatice Pinarbasi,[11] Paola Pisani,[12] Petcharin Srivatanakul,[13] Adeline Seow,[14] Haruhiko Sugimura,[15] Shinkan Tokudome,[16] Jun Yokota,[17] and Emanuela Taioli[18]

[1]Department of Preventive Medicine, Seoul National University College of Medicine; [2]Cancer Research Institute, Seoul National University, Seoul, Korea; [3]Division of Population Science, Fox Chase Cancer Center, Philadelphia, Pennsylvania; [4]Department of Physiological Chemistry, Karolinska Institutet, Stockholm, Sweden; [5]Department of Global Health and Socio-Epidemiology, Kyoto University School of Public Health, Kyoto, Japan; [6]Graduate School of Medical Sciences, Kyushu University, Fukuoka, Japan; [7]Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland; [8]University of Hawaii-Cancer Research Center of Hawaii, Honolulu, Hawaii; [9]National Health Research Institutes, Zhunan, Taiwan; [10]Department of Biology and Center for Cancer Research, The Hong Kong University of Science and Technology, Kowloon, Hong Kong; [11]Department of Biochemistry, Faculty of Medicine, Cumhuriyet University, Sivas, Turkey; [12]Cancer Epidemiology Unit, The University of Oxford, Oxford, United Kingdom; [13]Cancer Control Unit, National Cancer Institute, Bangkok, Thailand; [14]Department of Community, Occupational and Family Medicine, National University of Singapore, Singapore, Singapore; [15]Department of Pathology, Hamamatsu University School of Medicine, Sizuoka, Japan; [16]Department of Health Promotion and Preventive Medicine, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan; [17]Biology Division, National Cancer Center Research Institute, Tokyo, Japan; and [18]University of Pittsburgh Cancer Institute and the University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania

Abstract

To evaluate the roles of *CYP1A1* polymorphisms [*Ile*^{462}*Val* and *T*^{6235}*C* (*MspI*)] and deletion of *GSTM1* and and *GSTT1* in lung cancer development in Asian populations, a pooled analysis was conducted on 13 existing studies included in Genetic Susceptibility to Environmental Carcinogenesis database. This pooled analysis included 1,971 cases and 2,130 controls. Lung cancer risk was estimated as odds ratios (OR) and 95% confidence intervals (95% CI) using unconditional logistic regression model adjusting for age, sex, and pack-year. The *CYP1A1* ^{6235}*C* variant was associated with squamous cell lung cancer (TC versus *TT*: OR, 1.42; 95% CI, 0.96-2.09; *CC* versus *TT*: OR, 1.97; 95% CI, 1.26-3.07; P_{trend} = 0.003). In haplotype analysis, ^{462}*Val*-^{6235}*T* and *Ile-C* haplotypes were associated with lung cancer risk with reference to the *Ile-T* haplotype (OR, 3.41; 95% CI, 1.78-6.53 and OR, 1.39; 95% CI, 1.12-1.71, respectively). The *GSTM1*-null genotype increased squamous cell lung cancer risk (OR, 1.36; 95% CI, 1.05-1.77). When the interaction was evaluated with smoking, increasing trend of lung cancer risk as pack-year increased was stronger among those with the *CYP1A1* 6235 *TC/CC* genotype compared with those with *TT* genotype ($P_{interaction}$ = 0.001) and with the *GSTM1*-null genotype compared with the present type ($P_{interaction}$ = 0.08, when no genotype effect with no exposure was assumed). These results suggest that genetic polymorphisms in *CYP1A1* and *GSTM1* are associated with lung cancer risk in Asian populations. However, further investigation is warranted considering the relatively small sample size when subgroup analyses were done and the lack of environmental exposure data other than smoking. (Cancer Epidemiol Biomarkers Prev 2008;17(5):1120–6)

Introduction

Lung cancer mortality has increased rapidly during recent years in Asian countries. Cigarette smoking is the strongest established risk factor for lung cancer, but genetically determined variations in metabolism of tobacco-derived carcinogens may affect individual susceptibility to lung cancer. Cigarette smoke contains a variety of carcinogens, such as polycyclic aromatic hydrocarbons, *N*-nitrosoamines, and aromatic heterocyclic amines (1). These carcinogens undergo biotransformation by several enzymatic pathways, including P450s (CYP), glutathione *S*-transferase (GST), and *N*-acetyltransferase.

CYP1A1 plays an important role in the metabolism of polycyclic aromatic hydrocarbons, including benzo(*a*)pyrene, as a phase I enzyme and two variants (i.e., *Ile*^{462}*Val* and *T*^{6235}*C*), which are potentially functional (2-4), have been evaluated as susceptibility factors for lung cancer by a number of investigators. An increased risk of lung cancer has been observed with the ^{6235}*C* variant among smokers (5) and with ^{462}*Val* among nonsmokers (6) in

Received 11/13/07; revised 2/6/08; accepted 2/20/08.

Grant support: This work was funded in part by grants from European Commision (number 96/CAN/33919) and NIH (2P50CA090440-06). Additionally, we thank Barbara M. Stadlerman, Genetic Susceptibility to Environmental Carcinogens (GSEC) administrator, for her support (NIH grant P30CA047904-19S1).

Requests for reprints: Daehee Kang, Department of Preventive Medicine, Seoul National University College of Medicine, 28 Yongon-Dong, Chongno-Gu, Seoul 110-799, Korea. Phone: 82-2-740-8326; Fax: 82-2-747-4830. E-mail: dhkang@snu.ac.kr

doi:10.1158/1055-9965.EPI-07-2786

DOI:10.1158/1055-9965.EPI-07-2786

previous pooled analyses using the Genetic Susceptibility to Environmental Carcinogenesis (GSEC) database, whereas a separate meta-analysis did not find a significant association with lung cancer risk (7).

GSTM1 catalyzes reactive electrophilic intermediates derived from cigarette smoking, such as benzo(*a*)pyrene-7,8-diol-9,10-epoxides (BPDE), to less reactive and more easily excreted glutathione conjugates (8). Deletion of *GSTM1* has most widely been evaluated for the association with lung cancer risk and a significant association was found in several studies. Although three meta-analyses concluded that the *GSTM1*-null genotype is associated with an increased lung cancer risk (9-11), a GSEC pooled analysis indicated that there is no strong evidence for increased risk of lung cancer among those with the *GSTM1*-null genotype (12). Another isoform of GST (GSTT1) is also involved in carcinogen detoxification and its deletion polymorphism has been suggested to be associated with lung cancer in several studies. In a recent GSEC pooled analysis, the association was not significant for either Asians or Caucasians and no interaction was observed between *GSTT1*-null genotype and smoking on lung cancer (13).

Pooled analyses based on the GSEC data suggest that the effects of these variants tend to differ according to ethnicity possibly because of differences in linkage disequilibrium and environmental exposures. Consequently, gene-environment or gene-gene interactions might differ by ethnic group. Thus, we focused on Asian populations and evaluated the potential role of four selected polymorphisms in the three aforementioned genes (*CYP1A1* $Ile^{462}Val$ and $T^{6235}C$, and null genotypes for *GSTM1* and *GSTT1*) in the development of lung cancer and its specific cell types.

Materials and Methods

Study Population. Subjects were recruited from the International Collaborative Study on GSEC. The design of this collaborative project is explained in detail elsewhere (14). We obtained the original data of 15 case-control studies on genetic polymorphisms in *CYP1A1*, *GSTM1*, or *GSTT1* and risk of lung cancer conducted in Asian populations (15-30). Two studies were excluded due to a sample size of <10 subjects (29) or Caucasian ethnicity (Turkish; Table 1; ref. 30). The participation in GSEC was voluntary, and therefore, some relevant studies were not included in our analysis. The number of subjects included in this pooled analysis was 1,971 cases and 2,130 controls.

Statistical Analysis. All statistical procedures were conducted using Statistical Analysis System version 9.1.3 (SAS Institute) unless otherwise indicated. We estimated the study-specific odds ratios (OR) of lung cancer for each polymorphism using unconditional logistic regression. Results might vary slightly from those reported for some of the published studies because of differences in the inclusion criteria of cases and controls and in the statistical analyses. Heterogeneity among the studies was evaluated by means of the Cochrane Q test and publication bias was assessed by Begg's and Egger's test using STATA version 9. In the pooled analysis, lung cancer risk was estimated with the ORs and 95% confidence intervals (95% CI) by unconditional logistic regression, adjusting for age, sex, and pack-year.

In addition to conducting analyses of all lung cancer, we calculated cell type–specific ORs for the three most prevalent histologic subtypes of lung cancer: adenocarcinoma (n = 905), squamous cell carcinoma (n = 542), and small cell carcinoma (n = 181). Subgroup analyses for other histologic subtypes were not conducted due to small numbers of cases.

Hardy-Weinberg equilibrium for each single nucleotide polymorphism of *CYP1A1* was tested among controls with a Pearson χ^2 and linkage disequilibrium was assessed with D' and r^2. Individual haplotypes for two *CYP1A1* polymorphisms ($Ile^{462}Val$ and $T^{6235}C$) were estimated by expectation-maximization method and the overall difference in haplotype frequency profiles between cases and controls was assessed using the likelihood ratio test. The subjects missing both polymorphisms were excluded in haplotype analysis. The program uses a weighting scheme based on expectation-maximization–derived haplotype frequency estimates. Thus, every haplotype is weighted by the probability of carrying each pair of haplotypes rather than assigning a most likely haplotype to an individual. Missing genotypes result in more low-probability haplotype pairs and

Table 1. Selected characteristics of case-control studies pooled

Author	Ethnicity	Cases (n)	Controls (n)	Reference no.
Kihara et al. (1995)	Japanese	179	259	(15)
Ge et al. (1996)	Chinese	98 (39)*	27 (12)	(16)
Sugimura et al. (1998)	Japanese	260	209	(17)
Persson et al. (1999)	Chinese	80 (35)	123 (45)	(18)
Le Marchand et al. (1998)	Japanese	112 (42)	174 (50)	(19)
Kiyohara et al. (1998, 2000)	Japanese	132 (49)	84	(20, 21)
Lan et al. (2000)	Chinese	122 (43)	122 (43)	(22)
Yin et al. (2001)	Chinese	63 (9)	62 (9)	(23)
Zhao et al. (2001)	Chinese	233 (233)	190 (190)	(24)
Sunaga et al. (2002)	Japanese	198	152	(25)
Wang et al. (2003)	Chinese	112 (40)	119 (40)	(26)
Lee et al. (2006)	Korean	171	196	(27)
Pisani et al. (2006)	Thai	211 (71)	413 (158)	(28)
Total		1,971 (635)	2,130 (591)	

NOTE: One study with <10 subjects [Dresler et al. (29)] and Caucasian subjects [Pinarbasi et al. (30)] was excluded.
*Number of female subjects.

DOI:10.1158/1055-9965.EPI-07-2786

Table 2. Characteristics of subjects (1,971 cases and 2,130 controls)

	Cases, *n* (%)	Controls, *n* (%)	*P*	OR (95% CI)*
Age (y)				
<50	219 (11.1)	444 (20.9)	0.0001	
50-59	501 (25.4)	638 (30.0)		
60-69	718 (36.5)	599 (28.2)		
70-79	447 (22.7)	376 (17.7)		
≥80	85 (4.3)	70 (3.3)		
Mean (±SD)	62.6 (±10.7)	58.4 (±13.2)	0.0001	
Sex				
Male	1,336 (67.8)	1,537 (72.2)	0.002	
Female	635 (32.2)	591 (27.8)		
Smoking status				
Never	462 (24.9)	764 (38.3)	0.0001	Reference
Ever	1,396 (75.1)	1,230 (61.7)		2.29 (1.94-2.70)
Missing	113	136		
Pack-years in ever smokers				
0 < pack-year <35	468 (42.4)	640 (64.6)	0.0001	1.54 (1.28-1.36)
Pack-year ≥35	636 (57.6)	351 (35.4)		4.36 (3.51-5.35)
Missing	292	239		
Mean (±SD)	66.8 (±146.5)	49.4 (±107.9)	0.002	
Pathologic type				
AD	905 (50.2)			
SQ	542 (30.1)			
SM	181 (10.0)			
Other cell types	174 (9.7)			
Missing	169			

Abbreviations: AD, adenocarcinoma; SQ, squamous cell carcinoma; SM, small cell carcinoma.
*ORs were adjusted for age and sex.

each haplotype is weighted as such. An unconditional logistic regression model was used to estimate the effect of individual haplotypes by fitting an additive model, adjusting for sex, age, and pack-year.

Gene-smoking interactions (i.e., the modification of increasing pattern of lung cancer risk as the pack-year increases by different genotype) were evaluated by the significance of the coefficient of product term

Table 3. *CYP1A1* genotypes and lung cancer risk by histologic types

	Controls, *n* (%)	All cases, *n* (%)	OR (95% CI)*	AD, *n* (%)	OR (95% CI)*	SQ, *n* (%)	OR (95% CI)*	SM, *n* (%)	OR (95% CI)*
Ile^{462}Val	*n* = 1,096	*n* = 910		*n* = 337		*n* = 343		*n* = 121	
Ile/Ile	609 (55.6)	502 (55.2)	Reference	188 (55.8)	Reference	180 (52.5)	Reference	72 (59.5)	Reference
Ile/Val	421 (38.4)	329 (36.2)	0.88 (0.71-1.08)	117 (34.7)	0.94 (0.69-1.27)	132 (38.5)	1.06 (0.78-1.45)	41 (33.9)	0.80 (0.50-1.28)
Val/Val	66 (6.0)	79 (8.7)	1.06 (0.71-1.56)	32 (9.5)	1.53 (0.92-2.56)	31 (9.0)	1.01 (0.55-1.85)	8 (6.6)	0.60 (0.22-1.67)
P_{trend}			0.57		0.37		0.78		0.21
Ile/Ile or Ile/Val	1,030 (94.0)	831 (91.3)	Reference	305 (90.5)	Reference	312 (91.0)	Reference	113 (92.4)	Reference
Val/Val	66 (6.0)	79 (8.7)	1.14 (0.76-1.72)	32 (9.5)	1.57 (0.96-2.59)	31 (9.0)	1.14 (0.76-1.72)	8 (6.6)	0.65 (0.24-1.79)
T^{6235}C (MspI)	*n* = 953	*n* = 729		*n* = 284		*n* = 261		*n* = 95	
TT	333 (34.9)	241 (33.1)	Reference	106 (37.3)	Reference	75 (28.7)	Reference	36 (37.9)	Reference
TC	449 (47.1)	341 (46.8)	1.08 (0.84-1.39)	125 (44.0)	1.08 (0.84-1.39)	120 (46.0)	1.42 (0.96-2.09)	45 (47.4)	1.10 (0.65-1.86)
CC	171 (17.9)	147 (20.2)	1.13 (0.82-1.56)	53 (18.7)	1.13 (0.82-1.56)	66 (25.3)	1.97 (1.26-3.07)	14 (14.7)	0.73 (0.36-1.51)
P_{trend}			0.43		0.43		0.003		0.52
TC or CC	620 (65.1)	488 (67.0)	1.10 (0.86-1.39)	178 (62.7)	1.10 (0.86-1.39)	186 (71.3)	1.58 (1.10-2.27)	50 (52.6)	0.98 (0.60-1.62)
Haplotype†	*n* = 1,172 %	*n* = 979 %		*n* = 361 %		*n* = 385 %		*n* = 123 %	
Ile-T	56	52	Reference	55	Reference	49	Reference	57	Reference
Ile-C	19	21	1.39 (1.12-1.71)	18	0.99 (0.73-1.34)	24	2.10 (1.58-2.80)	19	1.29 (0.83-2.01)
Val-T	2	4	3.41 (1.78-6.53)	4	4.84 (2.32-10.1)	4	3.75 (1.70-8.27)	1	0.37 (0.02-8.06)
Val-C	23	23	0.96 (0.79-1.15)	23	0.94 (0.73-1.12)	24	1.06 (0.81-1.38)	23	0.89 (0.60-1.31)
$P_{omnibus}$‡			0.0001		0.0003		0.0001		0.40

*ORs were adjusted for age (<50, 50-59, 60-69, 70-79, and ≥80 y), sex, and pack-year.
†Subjects missing for both *CYP1A1 Ile^{462}Val* and *T^{6235}C (MspI)* data were excluded.
‡*P* value from the test of overall difference of haplotype distribution between cases and controls.

DOI:10.1158/1055-9965.EPI-07-2786

Table 4. *GSTM1* and *GSTT1* genotypes and lung cancer risk by histologic types

	Controls, *n* (%)	All cases, *n* (%)	OR (95% CI)*	AD, *n* (%)	OR (95% CI)*	SQ, *n* (%)	OR (95% CI)*	SM, *n* (%)	OR (95% CI)*
GSTM1	*n* = 1,604	*n* = 1,419		*n* = 760		*n* = 333		*n* = 169	
Present	713 (44.5)	589 (41.5)	Reference	332 (43.7)	Reference	124 (37.2)	Reference	59 (41.3)	Reference
Null	891 (55.6)	830 (58.5)	1.11 (0.95-1.29)	428 (56.3)	0.99 (0.82-1.19)	209 (62.8)	1.36 (1.05-1.77)	84 (58.7)	1.27 (0.88-1.83)
GSTT1	*n* = 1,024	*n* = 1,135		*n* = 579		*n* = 248		*n* = 71	
Present	538 (52.5)	579 (51.0)	Reference	300 (51.8)	Reference	141 (56.9)	Reference	25 (35.2)	Reference
Null	486 (47.5)	556 (49.0)	1.02 (0.84-1.24)	279 (48.2)	1.00 (0.80-1.26)	107 (43.2)	0.87 (0.62-1.21)	46 (64.8)	1.36 (0.99-1.86)

*ORs were adjusted for age (<50, 50-59, 60-69, 70-79, and ≥80 y), sex, and pack-year.

genotype*pack-year in the model. The test was equal to evaluate the difference of the slopes of two fitted lines stratified by categorized genotypes. Additionally, we tested the significance of the product term in the model without main effect term of genotype, which assumes that if there is no exposure to cigarette smoking, there is no difference in the risk of lung cancer between genotypes (27, 31). The assumption of no genotype effect when there is no smoking exposure was equal to common intercept assumption for two fitted lines by genotypes.

Results

The distributions by age, sex, smoking status, and cell types of the 1,971 lung cancer cases and 2,130 controls are presented in Table 2. The mean age was 62.6 (±10.7 years) in cases and 58.4 (±13.2 years) in controls ($P = 0.0001$). The proportion of ever smokers was much greater in cases (75.1%) than in controls (61.7%; $P = 0.0001$). In terms of cell types, adenocarcinoma (50.2%) and squamous cell carcinoma (30.1%) were the most common.

Genotype frequencies of *CYP1A1* $Ile^{462}Val$ and $T^{6235}C$ were consistent with Hardy-Weinberg equilibrium in the control group ($P > 0.35$) and the two polymorphisms were in moderate linkage disequilibrium ($D' = 0.86$ and $r^2 = 0.35$). The variant allele frequencies of the three polymorphisms (*CYP1A1* ^{462}Val, 0.25; ^{6235}C, 0.42; and *GSTT1* null, 0.48) in the controls were higher compared with those of Caucasian or African populations (13, 32). The frequency of the *GSTM1* null (0.56) was similar to that of Caucasians but higher compared with Africans (32). The *CYP1A1* ^{6235}C variant was associated with squamous cell lung cancer (*TC* versus *TT*: OR, 1.42; 95% CI, 0.96-2.09; *CC* versus *TT*: OR, 1.97; 95% CI, 1.26-3.07; $P_{trend} = 0.003$; Table 3). The *CYP1A1* ^{462}Val variant was moderately associated with adenocarcinoma (*Val/Val* versus *Ile/Ile* or *Ile/Val*: OR, 1.57; 95% CI, 0.96-2.59).

In haplotype analysis, ^{462}Val-^{6235}T and *Ile-C* haplotypes were associated with lung cancer risk with reference to the *Ile-T* haplotype (OR, 3.41; 95% CI, 1.78-6.53 and OR, 1.39; 95% CI, 1.12-1.71, respectively). An omnibus test showed that the distribution of the *CYP1A1* haplotypes was significantly different between all lung cancer cases and controls ($P = 0.0001$). In subgroup analysis, the difference was also significant for adenocarcinoma ($P = 0.0003$) and squamous cell carcinoma ($P = 0.0001$) and not for small cell carcinoma ($P = 0.40$).

The *GSTM1*-null genotype significantly increased squamous cell lung cancer risk (OR, 1.36; 95% CI, 1.05-1.77), and the *GSTT1*-null genotype was moderately associated only with small cell lung cancer risk (OR, 1.36; 95% CI, 0.99-1.86; Table 4). Analysis of combined genotypes did not reveal associations beyond what was apparent in the single polymorphism analyses (data not shown).

When the interaction was evaluated with smoking, increasing trend of lung cancer risk as pack-year increased was much stronger among those with the *CYP1A1* 6235 *TC/CC* genotype compared with those with *TT* genotype ($P_{interaction} = 0.001$; Fig. 1). Although the association between smoking and lung cancer was stronger among those with the *GSTM1*-null genotype compared with the present type, it was only marginally significant with the assumption of no genotype effect in the absence of the smoking exposure ($P_{interaction} = 0.08$). Significant interactive effect with smoking has not been observed for *GSTT1*.

There was no evidence of significant heterogeneity among studies or of publication bias for all four polymorphisms investigated in our study; we found only moderate heterogeneity for the effect of *CYP1A1* $^{462}Val/Val$ compared with *Ile/Ile* ($P = 0.08$), and all Begg's and Egger's tests were not significant ($P \geq 0.2$ and 0.3, respectively).

Discussion

Our results suggest that the *CYP1A1* polymorphisms ($Ile^{462}Val$ and $T^{6235}C$) and the *GSTM1*-null genotype are associated with lung cancer risk, especially for squamous cell carcinoma, in Asian populations. In addition, the association of smoking with lung cancer was significantly modified by the *CYP1A1* $T^{6235}C$ polymorphism in our study.

A significant interactive effect between the *CYP1A1* ^{6235}C allele and smoking is consistent with the results of previous pooled analysis that the stronger association between the ^{6235}C allele and lung cancer was found among ever smokers (5). The previous pooled analysis for the *GSTM1*-null genotype conducted by Benhamou et al. (12) found a nonsignificant elevated lung cancer risk among Asians, especially among heavy smoker (>40 pack-years). Likewise, our extended analysis with additional Asian populations also observed a moderate elevation of overall lung cancer risk by the *GSTM1* deletion and moderate interaction with smoking. On the