Among the four HLA-class II genes, the *HLA-DRB1*, *-DQA1*, and *-DQB1* genes are known to comprise haplotypes carrying diverse non-synonymous SNPs and express polymorphic antigen proteins (HLA-class II alleles) [23]. Therefore, we genotyped the same set of case and control subjects for 24 SNPs in the coding exons of the *DRB1*, *DQA1* and *DQB1* genes that discriminate the HLA-class II alleles by the sequencing-based typing method (**Table III**). These 24 SNPs did not deviate from HWE in either the cases or the controls. These SNPs showed LD with the SNPs in LD block 1 (**Figure 1**), and patterns of LD were quite similar between the cases and controls (**Supplementary Figure 1**), indicating that distribution of 6p21.31 SNPs on chromosome DNA is not significantly different between these two populations. Many HLA-class II alleles, including those for each of the *DRB1*, *DQA1* and *DQB1* genes as well as those for contigs of the three genes (i.e., *DR-DQ* allele), determined by haplotypes for these exonic SNPs showed significantly different distributions between the cases and controls (**Table IV**). Among them, the DQA1*03 allele showed the largest difference with an OR of 1.50 (P=6.6x10^{-6}) and the DQA1*01 allele was the second largest (OR=0.69, P=2.8x10^{-5}). Accordingly, several *DR-DQ* alleles containing the DQA1*03 or DQA1*01 allele as well as several *DRB1* and *DQB1* alleles linked to the DQA1*03 or DQA1*01 allele also showed significantly different distributions (**Table IV**).

Discrimination of HLA alleles using intronic or intergenic SNPs is considered to be appropriate to analyze a large number of samples as an

alternative to conventional methods using exonic SNPs, due to rapidity and cost-effectiveness [23,24]. Two exonic SNPs in the *DQA1* gene, DQA1_2_145 and DQA1_2_150, which were genotyped by sequencing, were responsible for discrimination of the DQA1*01 and *03 alleles. These two SNPs showed high (R^2>0.98) correlation coefficients with two intronic SNPs in *DQA1*, rs17426593 and rs34843907, respectively (**Supplementary Table III**), which were genotyped by the Taqman method (**Supplementary Figure 2**). In fact, DQA1*03 and DQA1*01 alleles deduced by these two intronic SNPs showed high (R^2>0.97) correlation coefficients with those determined by two exonic SNPs (**Supplementary Table IV**). Thus, DQA1*03 and DQA1*01 alleles were discriminated by combined genotypes of two intronic SNPs, rs17426593 and rs34843907, and the association of *DQA1* alleles with lung ADC risk was further examined in a larger number of subjects by genotyping these two SNPs. Genotyping of an additional 1,131 ADC cases and 648 controls in the NCCH set enabled us to calculate combined ORs in 1,656 ADC cases and 1,173 controls (all subjects of the NCCH set in **Table I**), and the ORs of the DQA1*03 and DQA1*01 alleles were 1.36 (P=5.3x10^{-7}) and 0.77 (P=1.4x10^{-5}), respectively (**Figure 2a, Table II and Supplementary Table V**). Associations of these alleles with ADC risk were observed both in smokers and non-smokers; both male and female. A significant increase in OR of DQA1*03 for ADC risk was also observed in another set (NNGH set in **Table I**) of cases and controls **(Figure 2a, Supplementary Table V**), while a decrease in OR of DQA1*01 was insignificant. The DQA1*03 allele comprised the same haplotype with the risk allele of the intronic rs17426593 SNP (R^2=0.988), which showed the largest difference in allelic distribution between the cases and controls (**Table III**). Therefore,

DQA1*03 was defined as a risk allele in the 6p21.31 locus, although it is possible that intronic SNPs rather than exonic SNPs play a causal role in lung ADC susceptibility. Associations of the DQA1*03 allele with risks were further examined for SQC and SCC, two other major histological types of lung cancer to clarify whether the association is specific to ADC or not. Increases in ORs of DQA1*03 for SQC risk were significant both in the NCCH and NNGH sets, while ORs for SCC risk, calculated only for the NCCH set, were marginally increased for DQA1*03 (**Figure 2a, Supplementary Table V**). Therefore, involvements of *HLA-DQA1* not only in ADC risk but also in other histological types of lung cancer were suggested.

The 6p21.31 locus maps 1-Mb proximal to *BAT3-MSH5*, another lung cancer susceptibility locus at 6p21.33 identified by a GWAS on Europeans and Americans [4]. Therefore, we next examined a SNP in this region, rs3117582, which showed a significant association in that study [4], in a set of 525 ADC cases and 525 controls (Subjects for the SNP analysis stage, **Table II**). It was monomorphic for the protective allele in these subjects. We therefore examined 7 SNPs in LD with this SNP in Europeans (i.e., D'=1 in the HapMap data); however, associations of these loci were weaker than those of the 6p21.31 locus, and these SNPs comprised a distinct LD block from the 6p21.31 locus containing four HLA-class II genes (**Supplementary Figure 3**). Therefore, we concluded that 6p21.31 is a novel lung ADC susceptibility locus on chromosome 6p.

Next, we examined associations of SNPs in other lung cancer susceptibility loci [4-9] in 1,656 ADC cases and 1,173 controls (all subjects of the NCCH set in **Table I**). Two SNPs, rs2736100 and rs401681, were examined for

the 5p15.33 locus, and the former located in intron 2 of the *TERT* gene showed a stronger association than the latter. The association was observed only in ADC, but not in SQC and SCC **(Supplementary Table VI)** as recently reported [11,21]. A SNP in the *CHRNA3* gene at 15q25.1, rs1051730, showed a significant association with risks for ADC, SQC and SCC in our previous study [22].

Therefore, combined effects among the *HLA-DQA1*, *TERT* and *CHRNA3* loci with lung ADC risk were further investigated. Genotypes with risk alleles for each locus showed significantly increased ORs of 1.32-2.21, except for homozygotes for the minor allele of *CHRNA3* **(Figure 2b** and **Supplementary Table VII)**. When ORs were calculated according to the number of risk alleles for two of these three genes, *HLA-DQA1* and *TERT*, there was an increasing trend with increasing number of risk alleles (per risk-allele OR=1.43, P=7.8x10^{-16}), reaching 4.76 for carriers of all four risk alleles **(Figure 2b and Supplementary Table VIII)**. These two alleles independently conferred the risk (P for interaction=0.88). The present results indicated that individuals susceptible to ADC can be defined by combined genotypes of *HLA-DQA1* and *TERT*. There was also an increasing trend for the *TERT* and *CHRNA3* combination with a per risk-allele OR of 1.48. OR reached 4.27 for carriers of three or four risk alleles, when heterozygotes and homozygotes for the *CHRNA3* risk were combined due to a small number of homozygotes **(Supplementary Table VIII).** Increases in OR by the combination of *HLA-DQA1* and *CHRNA3* were not evident, and a negative interaction was suggested (P=0.083). However, it might be due to the small number of homozygotes for the *CHRNA3* risk allele. Accordingly, when compared for all three genes, there was also an increasing trend with a per risk-allele OR of 1.45,

however, only ORs for carriers of up to four risk alleles could be calculated.

The present study indicated *HLA-DQA1* at 6p21.31 as a novel locus associated with lung cancer risk and genotypes for this locus are useful for identification of individuals susceptible to lung ADC. It has been considered that immune surveillance systems conferred by HLA class I and II proteins are involved in the elimination of tumor cells *in vivo* [25]. HLA class I proteins are expressed in most nucleated cells and present tumor-specific antigens for cytotoxic $CD8^+$ T cells to recognize and lyse tumor cells. In addition, the immune response requires the presentation of antigenic peptides to T cells by class II molecules expressed on antigen-presenting cells, i.e., the heterodimer of HLA-DQA1 and -DQB1 proteins and of HLA-DRA and -DRB1 proteins. Therefore, it might be that polymorphisms of *HLA-DQA1* (and also those of *HLA-DQB1* and *-DRB1* that are in LD with those of *HLA-DQA1*) gene confers lung cancer susceptibility by causing inter-individual differences in the ability of HLA class II molecule to bind peptides produced in lung cancer cells and to cause immune response. However, we should consider that the present results were obtained by performing a number of association tests against smaller numbers of subjects than those of recent GWASs [4-11]. In addition, control subjects from NNGH used for validation of association had lung diseases, including COPD. A recent GWAS on COPD has shown the same susceptibility loci as lung cancer, such as 15q25.1, suggesting that lung cancer and other lung diseases share the same genetic etiology [26]. Therefore, it remains possible that associations observed in the present study were under- or over-represented. The number of control subjects in the present study was 30% less than that of ADC cases (Combined analysis in **Table II**), although optimal ratios of control

subjects to case subjects have been considered as being 1:1-4:1 [27], and this fact resulted in larger 95% confidence intervals of OR than analyzing optimal number of control subjects. Thus, further case-control studies will be needed to validate the association of the 6p21.31 locus with lung ADC risk. Notably, synonymous SNPs in the 6p21.31 locus, such as rs2187668 and rs1794282, also showed significant differences in allelic distributions in Europeans and Americans (**Figure 1, Supplementary Table IX**). The strength of this association was similar to those for SNPs in the *BAT3-MSH5* locus in those populations (**Supplementary Figure 3**). Therefore, it was strongly indicated that 6p21.31 is a lung ADC susceptibility locus not only in Japanese but also in Europeans and Americans. However, at present it remains unknown whether SNPs/alleles associated with risk are different among populations, since only a few SNPs have been examined for associations in Europeans and Americans (**Figure 1, Supplementary Figure 3 and Supplementary Table IX**). In addition, LD among SNPs in the HLA-class II locus is known to be different among different ethnic populations [23]. Thus, studies on multiple populations will give us more critical information on the roles of polymorphisms in the 6p21.31 locus and their interaction with other lung cancer susceptibility loci in lung ADC susceptibility.

The present GWAS on ADC risk was performed against 23,010 microsatellite loci spaced at approximately 130-kb intervals in the human genome. However, two other lung cancer susceptibility loci, 15q25.1 and 5p15.33, whose SNPs showed associations with risk in the population analyzed in the present study (**Supplementary Table VI**), were not detected in the present GWAS using microsatellites. Therefore, several lung ADC susceptibility

loci were likely to be overlooked in the present GWAS probably due to insufficient statistical power and a sparse marker density. Thus, a GWAS on lung ADC risk, in which hundreds of thousands SNPs are analyzed, is underway in our laboratory to comprehensively identify lung ADC susceptibility loci. Finally, in spite of facts that ADC is the commonest histological type of lung cancer in non-smokers and that ADC of non-smokers is showing an increasing trend [2,28], loci specifically associated with ADC risk of no-smokers have not been identified. Therefore, GWASs focusing on lung ADC risk of non-smokers would be also worth investigating to identify additional lung ADC susceptibility loci. In addition, case-control studies on subjects that were carefully chosen to represent cases and controls in the same population, such as a nested case-control study designated in a large-scale cohort study, will be critical to validate the significance of susceptibility loci on lung carcinogenesis for the application to targeted screening and/or prevention of lung ADC in future.

Disclosure of Potential Conflicts of Interest

The authors disclosed no potential conflicts of interest.

Acknowledgement

Grant support: Grants-in-Aid from the Ministry of Health, Labor and Welfare for the 3rd-term Comprehensive 10-year Strategy for Cancer Control (J. Yokota) and for Cancer Research, 19S-1 (T. Kohno) and the Ministry of Education, Culture, Sports, Science and Technology for Scientific Research on Priority Areas, KAKENHI 20014031 (T. Kohno).

We thank Drs. Takashi Eguchi, Ikuo Saito, Matsuhiko Hayashi, Keiichi Hirao, Teruhiko Yoshida, Hiromi Sakamoto, Kimio Yoshimura and Shunpei Ohnami for their help in collecting blood samples of the control subjects. We thank Drs. Gen Tamiya, Tokuki Sakiyama and Tomohisa Chiku for their help in GWAS. We also thank Ms. Satoko Matsuyama, Sachiyo Mimaki, Rumie Sasaki and Yuki Takei for technical assistance. K.S. is an awardee of a Research Resident Fellowship from the Foundation for Promotion of Cancer Research in Japan.

Figure legends

Figure 1 LD and association with lung cancer risk of polymorphisms in the 6p21.31 locus. The top panel shows association results for polymorphisms and the location of genes. The green square depicts the result for the D6S0067i microsatellite polymorphism, and red lozenges depict those for SNPs in the present study. Circles depict the results of GWASs on European and American populations. Blue circle: results for 1,989 cases and 2,625 controls in European countries [8]; yellow circle: 5,095 cases and 5,200 controls in European countries and USA [4]; purple circle: 2,971 cases and 3,746 controls in European countries, Canada and USA [5]. Results for ten SNPs commonly analyzed in the present and previous GWA studies (indicated in **Supplementary Table IX**) are depicted by bordered lozenges and circles. rs1794282 was monomorphic in the Japanese subjects. The bottom panel shows the LD structure for 55 SNPs in 525 control subjects. Boxes are shaded according to the pair-wise D′ values. Three LD blocks are indicated by bold black lines.

Figure 2 Forest plot representing risk for lung cancer. **(a)** Risk of the DQA1*03 and DQA1*01 alleles for lung cancer. ORs of the alleles adjusted for age, sex, smoking habit and/or hospital, and 95% CI are shown. Detailed data, including the numbers of case and control subjects and variables for adjustments for each test, are summarized in **Supplementary Table V.** **(b)** Risk of combined *HLA-DQA1* and *TERT* genotypes for lung ADC. ORs of the alleles adjusted for age, sex and smoking habit and 95% CI are shown. Detailed data, including the numbers of case and control subjects, are summarized in **Supplementary**

Tables VII and VIII.

References

1. Travis, W.D., Brambilla, E., Muller-Hermelink, H.K. and Harris, C.C. (2004) Pathology and Genetics, Tumours of Lung, Pleura, Thymus and Heart. In Travis, W.D., Brambilla, E., Muller-Hermelink, H.K. and Harris, C.C. (eds.), *World Health Organization Classification of Tumors:* , pp. 1-344.
2. Sun, S., Schiller, J.H. and Gazdar, A.F. (2007) Lung cancer in never smokers--a different disease. *Nat Rev Cancer*, **7**, 778-90.
3. Sobue, T., Yamamoto, S., Hara, M., Sasazuki, S., Sasaki, S. and Tsugane, S. (2002) Cigarette smoking and subsequent risk of lung cancer by histologic type in middle-aged Japanese men and women: the JPHC study. *Int J Cancer*, **99**, 245-51.
4. Wang, Y., Broderick, P., Webb, E., Wu, X., Vijayakrishnan, J., Matakidou, A., Qureshi, M., Dong, Q., Gu, X., Chen, W.V., Spitz, M.R., Eisen, T., Amos, C.I. and Houlston, R.S. (2008) Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat Genet*, **40**, 1407-9.
5. McKay, J.D., Hung, R.J., Gaborieau, V., Boffetta, P., Chabrier, A., Byrnes, G., Zaridze, D., Mukeria, A., Szeszenia-Dabrowska, N., Lissowska, J., Rudnai, P., Fabianova, E., Mates, D., Bencko, V., Foretova, L., Janout, V., McLaughlin, J., Shepherd, F., Montpetit, A., Narod, S., Krokan, H.E., Skorpen, F., Elvestad, M.B., Vatten, L., Njolstad, I., Axelsson, T., Chen, C., Goodman, G., Barnett, M., Loomis, M.M., Lubinski, J., Matyjasik, J., Lener, M., Oszutowska, D., Field, J., Liloglou, T., Xinarianos, G., Cassidy, A., Vineis, P., Clavel-Chapelon, F., Palli, D., Tumino, R., Krogh, V., Panico, S., Gonzalez, C.A., Ramon Quiros, J., Martinez, C., Navarro, C., Ardanaz, E., Larranaga, N., Kham, K.T., Key, T., Bueno-de-Mesquita, H.B., Peeters, P.H., Trichopoulou, A., Linseisen, J., Boeing, H., Hallmans, G., Overvad, K., Tjonneland, A., Kumle, M., Riboli, E., Zelenika, D., Boland, A., Delepine, M., Foglio, M., Lechner, D., Matsuda, F., Blanche, H., Gut, I., Heath, S., Lathrop, M. and Brennan, P. (2008) Lung cancer susceptibility locus at 5p15.33. *Nat Genet*, **40**, 1404-6.
6. Liu, P., Vikis, H.G., Wang, D., Lu, Y., Wang, Y., Schwartz, A.G., Pinney, S.M., Yang, P., de Andrade, M., Petersen, G.M., Wiest, J.S., Fain, P.R., Gazdar, A., Gaba, C., Rothschild, H., Mandal, D., Coons, T., Lee, J., Kupert, E., Seminara, D., Minna, J., Bailey-Wilson, J.E., Wu, X., Spitz, M.R., Eisen, T., Houlston, R.S., Amos, C.I., Anderson, M.W. and You, M. (2008) Familial aggregation of common sequence variants on 15q24-25.1 in lung cancer. *J Natl Cancer Inst*, **100**, 1326-30.
7. Kohno, T., Kunitoh, H., Suzuki, K., Yamamoto, S., Kuchiba, A., Matsuno, Y., Yanagitani, N. and Yokota, J. (2008) Association of KRAS polymorphisms with risk for lung adenocarcinoma accompanied by atypical adenomatous hyperplasias. *Carcinogenesis*, **29**, 957-63.
8. Hung, R.J., McKay, J.D., Gaborieau, V., Boffetta, P., Hashibe, M., Zaridze, D., Mukeria, A., Szeszenia-Dabrowska, N., Lissowska, J., Rudnai, P., Fabianova, E., Mates, D., Bencko, V., Foretova, L., Janout, V., Chen, C., Goodman, G., Field, J.K., Liloglou, T., Xinarianos, G., Cassidy, A., McLaughlin, J., Liu, G., Narod, S., Krokan, H.E., Skorpen, F., Elvestad, M.B., Hveem, K., Vatten, L., Linseisen, J., Clavel-Chapelon, F., Vineis, P., Bueno-de-Mesquita, H.B., Lund, E., Martinez, C., Bingham, S., Rasmuson, T., Hainaut, P., Riboli, E., Ahrens, W.,

Benhamou, S., Lagiou, P., Trichopoulos, D., Holcatova, I., Merletti, F., Kjaerheim, K., Agudo, A., Macfarlane, G., Talamini, R., Simonato, L., Lowry, R., Conway, D.I., Znaor, A., Healy, C., Zelenika, D., Boland, A., Delepine, M., Foglio, M., Lechner, D., Matsuda, F., Blanche, H., Gut, I., Heath, S., Lathrop, M. and Brennan, P. (2008) A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature*, **452**, 633-7.

9. Amos, C.I., Wu, X., Broderick, P., Gorlov, I.P., Gu, J., Eisen, T., Dong, Q., Zhang, Q., Gu, X., Vijayakrishnan, J., Sullivan, K., Matakidou, A., Wang, Y., Mills, G., Doheny, K., Tsai, Y.Y., Chen, W.V., Shete, S., Spitz, M.R. and Houlston, R.S. (2008) Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat Genet*, **40**, 616-22.
10. Broderick, P., Wang, Y., Vijayakrishnan, J., Matakidou, A., Spitz, M.R., Eisen, T., Amos, C.I. and Houlston, R.S. (2009) Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. *Cancer Res*, **69**, 6633-41.
11. Landi, M., Chatterjee, N., Kai Yu, K., Goldin, L., Goldstein, A., Rotunno, M., Mirabello, L., Jacobs, K., Wheeler, W., Yeager, M., Bergen, A., Li, Q., Consonni, D., Pesatori, A., Wacholder, S., Thun, M., Diver, R., Oken, M., Virtamo, J., Albanes, D., Wang, Z., Burdette, L., Doheny, K., Pugh, E., Laurie, C., Brennan, P., Hung, R., Gaborieau, V., McKay, J., Lathrop, M., McLaughlin, J., Wang, Y., Tsao, M., Spitz, M., Wang, Y., Krokan, H., Vatten, L., Skorpen, F., Arnesen, E., Benhamou, S., Bouchard, C., Metsapalu, A., Vooder, T., Nelis, M., Välk, K., Field, J., Chen, C., Goodman, G., Sulem, P., Thorleifsson, G., Rafnar, T., Eisen, T., Sauter, W., Rosenberger, A., Bickeböller, H., Risch, A., Chang-Claude, J., Wichmann, H., Stefansson, K., Houlston, R., Amos, C., Fraumeni, J., Savage, S., Bertazzi, P., Tucker, M., Chanock, S. and Caporaso, N. (2009) A Genome-wide Association Study of Lung Cancer Identifies a Region of Chromosome 5p15 Associated with Risk for Adenocarcinoma *Am J Hum Genet*, **85**, 1-13.
12. Yamaguchi-Kabata, Y., Nakazono, K., Takahashi, A., Saito, S., Hosono, N., Kubo, M., Nakamura, Y. and Kamatani, N. (2008) Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am J Hum Genet*, **83**, 445-56.
13. Sakamoto, H., Yoshimura, K., Saeki, N., Katai, H., Shimoda, T., Matsuno, Y., Saito, D., Sugimura, H., Tanioka, F., Kato, S., Matsukura, N., Matsuda, N., Nakamura, T., Hyodo, I., Nishina, T., Yasui, W., Hirose, H., Hayashi, M., Toshiro, E., Ohnami, S., Sekine, A., Sato, Y., Totsuka, H., Ando, M., Takemura, R., Takahashi, Y., Ohdaira, M., Aoki, K., Honmyo, I., Chiku, S., Aoyagi, K., Sasaki, H., Ohnami, S., Yanagihara, K., Yoon, K.A., Kook, M.C., Lee, Y.S., Park, S.R., Kim, C.G., Choi, I.J., Yoshida, T., Nakamura, Y. and Hirohashi, S. (2008) Genetic variation in PSCA is associated with susceptibility to diffuse-type gastric cancer. *Nat Genet*, **40**, 730-40.
14. Tamiya, G., Shinya, M., Imanishi, T., Ikuta, T., Makino, S., Okamoto, K., Furugaki, K., Matsumoto, T., Mano, S., Ando, S., Nozaki, Y., Yukawa, W., Nakashige, R., Yamaguchi, D., Ishibashi, H., Yonekura, M., Nakami, Y., Takayama, S., Endo, T., Saruwatari, T., Yagura, M., Yoshikawa, Y., Fujimoto, K., Oka, A., Chiku, S., Linsen, S.E., Giphart, M.J., Kulski, J.K., Fukazawa, T., Hashimoto, H., Kimura, M., Hoshina, Y., Suzuki, Y., Hotta, T., Mochida, J.,

Minezaki, T., Komai, K., Shiozawa, S., Taniguchi, A., Yamanaka, H., Kamatani, N., Gojobori, T., Bahram, S. and Inoko, H. (2005) Whole genome association study of rheumatoid arthritis using 27 039 microsatellites. *Hum Mol Genet*, **14**, 2305-21.

15. Sunaga, N., Kohno, T., Yanagitani, N., Sugimura, H., Kunitoh, H., Tamura, T., Takei, Y., Tsuchiya, S., Saito, R. and Yokota, J. (2002) Contribution of the NQO1 and GSTT1 polymorphisms to lung adenocarcinoma susceptibility. *Cancer Epidemiol Biomarkers Prev*, **11**, 730-8.
16. Sakiyama, T., Kohno, T., Mimaki, S., Ohta, T., Yanagitani, N., Sobue, T., Kunitoh, H., Saito, R., Shimizu, K., Hirama, C., Kimura, J., Maeno, G., Hirose, H., Eguchi, T., Saito, D., Ohki, M. and Yokota, J. (2005) Association of amino acid substitution polymorphisms in DNA repair genes TP53, POLI, REV1 and LIG4 with lung cancer risk. *Int J Cancer*, **114**, 730-7.
17. Kohno, T., Kunitoh, H., Toyama, K., Yamamoto, S., Kuchiba, A., Saito, D., Yanagitani, N., Ishihara, S., Saito, R. and Yokota, J. (2006) Association of the OGG1-Ser326Cys polymorphism with lung adenocarcinoma risk. *Cancer Sci*, **97**, 724-8.
18. Kohno, T., Sakiyama, T., Kunitoh, H., Goto, K., Nishiwaki, Y., Saito, D., Hirose, H., Eguchi, T., Yanagitani, N., Saito, R., Sasaki-Matsumura, R., Mimaki, S., Toyama, K., Yamamoto, S., Kuchiba, A., Sobue, T., Ohta, T., Ohki, M. and Yokota, J. (2006) Association of polymorphisms in the MTH1 gene with small cell lung carcinoma risk. *Carcinogenesis*, **27**, 2448-54.
19. Tomlinson, I.P., Webb, E., Carvajal-Carmona, L., Broderick, P., Howarth, K., Pittman, A.M., Spain, S., Lubbe, S., Walther, A., Sullivan, K., Jaeger, E., Fielding, S., Rowan, A., Vijayakrishnan, J., Domingo, E., Chandler, I., Kemp, Z., Qureshi, M., Farrington, S.M., Tenesa, A., Prendergast, J.G., Barnetson, R.A., Penegar, S., Barclay, E., Wood, W., Martin, L., Gorman, M., Thomas, H., Peto, J., Bishop, D.T., Gray, R., Maher, E.R., Lucassen, A., Kerr, D., Evans, D.G., Schafmayer, C., Buch, S., Volzke, H., Hampe, J., Schreiber, S., John, U., Koessler, T., Pharoah, P., van Wezel, T., Morreau, H., Wijnen, J.T., Hopper, J.L., Southey, M.C., Giles, G.G., Severi, G., Castellvi-Bel, S., Ruiz-Ponte, C., Carracedo, A., Castells, A., Forsti, A., Hemminki, K., Vodicka, P., Naccarati, A., Lipton, L., Ho, J.W., Cheng, K.K., Sham, P.C., Luk, J., Agundez, J.A., Ladero, J.M., de la Hoya, M., Caldes, T., Niittymaki, I., Tuupanen, S., Karhu, A., Aaltonen, L., Cazier, J.B., Campbell, H., Dunlop, M.G. and Houlston, R.S. (2008) A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet*, **40**, 623-30.
20. Hauge, X.Y. and Litt, M. (1993) A study of the origin of 'shadow bands' seen when typing dinucleotide repeat polymorphisms by the PCR. *Hum Mol Genet*, **2**, 411-5.
21. Jin, G., Xu, L., Shu, Y., Tian, T., Liang, J., Xu, Y., Wang, F., Chen, J., Dai, J., Hu, Z. and Shen, H. (2009) Common genetic variants on 5p15.33 contribute to risk of lung adenocarcinoma in a Chinese population. *Carcinogenesis*.
22. Shiraishi, K., Kohno, T., Kunitoh, H., Watanabe, S.I., Goto, K., Nishiwaki, Y., Shimada, Y., Hirose, H., Saito, I., Kuchiba, A., Yamamoto, S. and Yokota, J. (2009) Contribution of Nicotine Acetylcholine Receptor Polymorphisms to Lung Cancer Risk in a Smoking -independent Manner in the Japanese. *Carcinogenesis, in press*.

23. de Bakker, P.I., McVean, G., Sabeti, P.C., Miretti, M.M., Green, T., Marchini, J., Ke, X., Monsuur, A.J., Whittaker, P., Delgado, M., Morrison, J., Richardson, A., Walsh, E.C., Gao, X., Galver, L., Hart, J., Hafler, D.A., Pericak-Vance, M., Todd, J.A., Daly, M.J., Trowsdale, J., Wijmenga, C., Vyse, T.J., Beck, S., Murray, S.S., Carrington, M., Gregory, S., Deloukas, P. and Rioux, J.D. (2006) A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet*, **38**, 1166-72.

24. Miretti, M.M., Walsh, E.C., Ke, X., Delgado, M., Griffiths, M., Hunt, S., Morrison, J., Whittaker, P., Lander, E.S., Cardon, L.R., Bentley, D.R., Rioux, J.D., Beck, S. and Deloukas, P. (2005) A high-resolution linkage-disequilibrium map of the human major histocompatibility complex and first generation of tag single-nucleotide polymorphisms. *Am J Hum Genet*, **76**, 634-46.

25. Knutson, K.L. and Disis, M.L. (2005) Tumor antigen-specific T helper cells in cancer immunity and immunotherapy. *Cancer Immunol Immunother*, **54**, 721-8.

26. Pillai, S.G., Ge, D., Zhu, G., Kong, X., Shianna, K.V., Need, A.C., Feng, S., Hersh, C.P., Bakke, P., Gulsvik, A., Ruppert, A., Lodrup Carlsen, K.C., Roses, A., Anderson, W., Rennard, S.I., Lomas, D.A., Silverman, E.K. and Goldstein, D.B. (2009) A genome-wide association study in chronic obstructive pulmonary disease (COPD): identification of two major susceptibility loci. *PLoS Genet*, **5**, e1000421.

27. Breslow, N.E. and Day, N.E. (1980) 1.5 Planning. In Davis, W. (ed.), *Statistical Methods in Cancer Research, Volume 1-The Analysis of Case-Control Studies*. IARC Sci Publ, pp. 23-32.

28. Subramanian, J. and Govindan, R. (2007) Lung cancer in never smokers: a review. *J Clin Oncol*, **25**, 561-70.

Table I. Characteristics of study subjects

Category	Group	No.	Age (Mean±SD)	Sex (% male)	Smoking habit (%)		Pack-years of smokers (Mean±SD)
					Non-smoker	Smoker	
NCCH set	Case	2,343	59±9	65	34	66	51±30
	Adenocarcinoma	1,656	58±9	56	46	54	43±27
	Squamous cell carcinoma	390	62±7	91	3	97	61±29
	Small cell carcinoma	297	62±9	80	5	95	62±32
	Control	1,173	48±14	58	63	37	27±21
NNGH set	Case	136	68±10	74	27	73	55±29
	Adenocarcinoma	84	67±10	64	39	61	48±25
	Squamous cell carcinoma	52	70±9	90	6	94	62±32
	Control	145	64±14	71	33	67	45±35

Table II. GWAS and validation studies to identify loci associated with lung adenocarcinoma risk

Stage		Subject (No.)		Polymorphic loci analyzed	Genotyping method	Result				
		Case	Control			Loci/Allele	Crude OR (95% CI)	*P* value	Adjusted OR[a] (95% CI)	*P* value
GWAS	1st[b]	Patients of NCCH (200)	Volunteers enrolled in Tokai University (200)	23,010 microsatellites	Pooled DNA typing	1,328 loci	ND	*P*2cm < 0.05	ND	ND
	2nd[c]	Patients of NCCH (200)	Volunteers enrolled in Keio University (200)	431 microsatellites	Pooled DNA typing	17 loci	ND	*P*2cm < 0.05	ND	ND
	3rd[d]	Patients of NCCH (576) consisting of 192 1st set, 192 2nd set and 192 another subjects	Tokai and Keio volunteers and non-cancer NCCH patients (576) consisting of 192 1st and 192 2nd set volunteers; and 192 non-cancer NCCH patients	17 microsatellites	Individual DNA typing	D6S0067i and other 5 loci	ND	*P*2cm=2.4 x 10^{-7} at D6S0067i *P*2cm =0.012-0.0011 at 5 other loci	ND	ND
SNP analysis[e]		Randomly selected 3rd stage GWAS subjects (525)	Randomly selected 3rd stage GWAS subjects (525)	56 SNPs	24 SNPs by sequencing, 32 SNPs by Taqman PCR	rs17426593 DQA1*03 DQA1*01	1.51 (1.27-1.80) 1.50 (1.26-1.79) 0.69 (0.58-0.82)	$4.2x10^{-6}$ $6.6x10^{-6}$ $2.8x10^{-5}$	ND	ND
		Same as above	Same as above	rs17426593 and rs34843907	-	DQA1*03 DQA1*01	1.52 (1.27-1.81) 0.69 (0.58-0.82)	$3.4x10^{-6}$ $2.3x10^{-5}$	ND	ND
Validation		Other patients of NCCH (1,131)	Other subjects (648) consisting of 9 Tokai and 478 Keio volunteers; and 161 non-cancer NCCH patients	Same as above	Taqman PCR	DQA1*03 DQA1*01	1.27 (1.11-1.45) 0.86 (0.75-0.99)	$5.6x10^{-4}$ 0.030	ND	ND
(Combined analysis)		Patients of NCCH (1,656)	Subjects (1,173) consisting of 200 Tokai and 635 Keio volunteers and 338 non-cancer NCCH patients	Same as above	-	DQA1*03 DQA1*01	1.35 (1.21-1.51) 0.78 (0.70-0.87)	$5.6x10^{-8}$ $6.0x10^{-6}$	1.36 (1.20-1.54) 0.77 (0.68-0.87)	$5.3x10^{-7}$ $1.4x10^{-5}$
Validation in another set		Patients of NNGH (84)	Non-cancer patients of NNGH (145)	Same as above	Taqman PCR	DQA1*03 DQA1*01	1.57 (1.07-2.30) 0.77 (0.52-1.13)	0.022 0.18	1.70 (1.14-2.53) 0.68 (0.49-1.09)	0.0087 0.12

ND: not determined.

[a]adjusted for sex, age, and smoking.

[b]23,010 microsatellite loci containing repeat units of 2-6-bp were examined, and 1,328 loci showed significant differences in allele distribution.

[c]Among 1,328 loci selected in the 1st GWAS stage, 431 loci that contained repeat units of 3-6-bp were examined, and 17 loci showed significant differences in allele distribution.

[d]17 loci selected in the 2nd GWAS stage were examined, and 6 loci showed significant differences in allele distribution. Only the D6S0067i locus was identified as being significantly different after Bonferroni correction (i.e., $P < 0.05/23,010 = 2.2x10^{-6}$). In this stage, 576 subjects consisted of two sets of 192 subjects which were chosen from two sets of 200 subjects examined in the 1st and 2nd GWAS stages, respectively, and another 192 subjects were examined.

[e]In this stage, 525 cases and 525 controls, which were randomly chosen from the 576 cases and 576 controls examined in the 3rd GWAS stage, were examined for 56 SNPs.

Table III. Differences in the allele distribution of 56 SNPs in the 450-kb region surrounding the D6S0067i locus between 525 lung adenocarcinoma cases and 525 controls

SNP	Genome location	Gene	Position	Allele	Minor allele frequency		*P* value	OR	95%CI	Deviation from HWE (*P*)		LD block
					Control	Case				Control	Case	
First 32 SNPs												
rs7773756	32,510,442			T/C	0.375	0.443	0.0016	1.33	1.11 - 1.58	0.66	0.83	1
rs16822586	32,515,751	*HLA-DRA*	exon 1	G/C	0.076	0.074	0.86	0.97	0.70 - 1.34	0.20	0.001	1
rs2239806	32,519,285	*HLA-DRA*	intron 3	G/A	0.208	0.179	0.089	0.83	0.67 - 1.03	0.09	0.31	1
rs7192*	32,519,624	*HLA-DRA*	exon 4	G/T	0.444	0.376	0.0016	0.76	0.63 - 0.90	1.00	0.19	1
rs3129763[a]	32,698,903			G/A	0.064	0.070	0.60	1.10	0.78 - 1.55	0.91	0.44	1
rs9272346	32,712,350			G/A	0.469	0.564	1.5×10^{-5}	1.46	1.23 - 1.74	0.11	0.95	1
rs2187668[a]	32,713,862	*HLA-DQA1*	intron 1	G/A	0.033	0.040	0.42	1.21	0.77 - 1.91	0.54	0.09	1
rs17426593	32,716,055	*HLA-DQA1*	intron 1	T/C	0.344	0.442	4.2×10^{-6}	1.51	1.27 - 1.80	0.55	0.18	1
rs34843907	32,718,037	*HLA-DQA1*	intron 3	A/C	0.470	0.563	1.3×10^{-5}	1.45	1.22 - 1.72	0.11	0.91	1
rs28584179	32,734,097			C/T	0.106	0.063	4.1×10^{-4}	0.57	0.41 - 0.78	0.68	0.50	1
rs17205373	32,734,188			C/G	0.054	0.046	0.42	0.85	0.57 - 1.26	0.66	0.92	1
rs6906021	32,734,289			T/C	0.377	0.358	0.39	0.92	0.77 - 1.10	0.82	0.28	1
rs28672722	32,734,515			G/T	0.416	0.353	0.0032	0.77	0.64 - 0.91	0.59	0.86	1
rs28746825	32,741,450	*HLA-DQB1*	intron 1	A/G	0.365	0.463	5.4×10^{-6}	1.50	1.26 - 1.79	0.22	0.26	1
rs34692792	32,741,519	*HLA-DQB1*	intron 1	T/C	0.218	0.205	0.47	0.93	0.75 - 1.14	0.47	0.29	1
rs2647012[a]	32,772,436			G/A	0.214	0.143	2.3×10^{-5}	0.61	0.49 - 0.77	0.37	0.22	1
rs1794282[a]	32,774,504			G/A	0.000	0.000	-	-	- - -	-	-	-
rs2856717	32,778,286			C/T	0.219	0.147	1.7×10^{-5}	0.61	0.49 - 0.77	0.19	0.33	1
rs2051600	32,817,287	*HLA-DQA2*	intron 1	C/T	0.188	0.153	0.030	0.78	0.62 - 0.98	0.62	0.33	2
rs2239800[a]	32,821,245	*HLA-DQA2*	intron 2	T/C	0.289	0.292	0.88	1.01	0.84 - 1.23	0.92	0.47	2
rs2071798	32,822,570	*HLA-DQA2*	3'UTR	T/C	0.360	0.302	0.0048	0.77	0.64 - 0.92	0.33	0.92	2
rs9276558	32,832,039			G/A	0.358	0.301	0.0051	0.77	0.64 - 0.92	0.28	0.94	2
rs1573649[a]	32,839,236			T/C	0.426	0.377	0.021	0.81	0.68 - 0.97	0.11	0.77	2
rs2071475	32,890,365	*HLA-DOB*	intron 2	C/T	0.238	0.261	0.21	1.14	0.93 - 1.38	0.53	0.24	2
rs2071469[a]	32,892,761	*HLA-DOB*	5'UTR	G/A	0.424	0.469	0.037	1.20	1.01 - 1.43	0.13	0.66	2
rs241455	32,903,997	*TAP2*	3'UTR	G/T	0.314	0.339	0.21	1.12	0.94 - 1.35	0.35	0.17	2
rs1800454	32,908,390	*TAP2*	exon 6	G/A	0.126	0.150	0.12	1.22	0.95 - 1.56	0.51	0.20	2/3
rs2071552	32,914,439	*TAP2*	5'UTR	T/C	0.418	0.406	0.56	0.95	0.80 - 1.13	0.53	0.22	3
rs2071463	32,920,506	*PSMB8*	5'UTR	G/A	0.380	0.404	0.25	1.11	0.93 - 1.32	0.74	0.18	3
rs1057373[a]	32,921,257	*TAP1*	3'UTR	G/T	0.105	0.121	0.23	1.18	0.90 - 1.55	0.72	0.18	3
rs2071480	32,929,837			G/T	0.339	0.341	0.95	1.01	0.84 - 1.20	0.16	0.23	3
rs17587[a]	32,933,068	*PSMB9*	exon 3	G/A	0.243	0.225	0.33	0.90	0.74 - 1.11	0.47	0.71	3
Additional 24 SNPs for HLA allele discrimination												
DRB1_2_244	32,659,890	*HLA-DRB1*	exon 2	G/T	0.269	0.271	0.88	1.01	0.84 - 1.23	0.49	0.88	1
DRB1_2_160	32,659,974	*HLA-DRB1*	exon 2	C/A	0.029	0.023	0.69	0.79	0.47 - 1.33	0.52	0.56	1
DRB1_2_156	32,659,978	*HLA-DRB1*	exon 2	G/A	0.204	0.259	0.0055	1.37	1.12 - 1.66	0.83	1.00	1
DRB1_2_127	32,660,007	*HLA-DRB1*	exon 2	A/T	0.418	0.325	1.1×10^{-5}	0.67	0.56 - 0.80	0.16	0.28	1
DRB1_2_106	32,660,028	*HLA-DRB1*	exon 2	T/A	0.003	0.004	0.70	1.35	0.30 - 6.03	0.95	0.93	1
DRB1_2_84	32,660,050	*HLA-DRB1*	exon 2	A/C	0.191	0.266	1.2×10^{-4}	1.53	1.26 - 1.87	0.86	0.57	1
DRB1_2_81	32,660,053	*HLA-DRB1*	exon 2	T/C	0.224	0.207	0.30	0.90	0.74 - 1.10	0.47	0.66	1
DRB1_2_64	32,660,070	*HLA-DRB1*	exon 2	T/A	0.130	0.154	0.047	1.22	0.97 - 1.55	0.60	0.79	1
DRB1_2_61	32,660,073	*HLA-DRB1*	exon 2	G/A	0.002	0.000	0.16	0.00	- - -	0.96	1.00	1
DRB1_2_33	32,660,101	*HLA-DRB1*	exon 2	C/T	0.172	0.174	0.91	1.01	0.82 - 1.26	0.10	0.47	1
DQA1_2_136	32,717,200	*HLA-DQA1*	exon 2	A/T	0.464	0.568	8.4×10^{-6}	1.52	1.28 - 1.79	0.16	0.84	1
DQA1_2_141	32,717,205	*HLA-DQA1*	exon 2	A/C	0.348	0.450	9.7×10^{-6}	1.54	1.30 - 1.83	0.91	0.27	1
DQA1_2_145	32,717,209	*HLA-DQA1*	exon 2	A/G	0.348	0.444	6.7×10^{-6}	1.50	1.26 - 1.79	0.51	0.26	1
DQA1_2_150	32,717,304	*HLA-DQA1*	exon 2	G/A	0.470	0.562	2.8×10^{-5}	1.44	1.22 - 1.71	0.11	0.83	1
DQB1_2_156	32,740,667	*HLA-DQB1*	exon 2	G/A	0.029	0.031	0.90	1.07	0.65 - 1.76	0.49	0.47	1
DQB1_2_145	32,740,678	*HLA-DQB1*	exon 2	A/T	0.471	0.560	1.8×10^{-5}	1.43	1.21 - 1.69	0.04	0.87	1
DQB1_2_134	32,740,689	*HLA-DQB1*	exon 2	G/A	0.065	0.059	0.72	0.90	0.64 - 1.27	0.53	0.39	1
DQB1_2_131	32,740,692	*HLA-DQB1*	exon 2	G/C	0.092	0.058	0.0029	0.61	0.44 - 0.84	0.19	0.86	1
DQB1_2_121	32,740,702	*HLA-DQB1*	exon 2	G/A	0.097	0.099	0.81	1.03	0.77 - 1.36	0.73	0.15	1
DQB1_2_101	32,740,722	*HLA-DQB1*	exon 2	G/A	0.315	0.355	0.034	1.20	1.00 - 1.43	0.71	0.96	1
DQB1_2_100	32,740,723	*HLA-DQB1*	exon 2	C/T	0.335	0.323	0.55	0.95	0.79 - 1.13	0.43	0.51	1
DQB1_2_68	32,740,755	*HLA-DQB1*	exon 2	G/A	0.114	0.065	3.2×10^{-5}	0.54	0.40 - 0.73	0.42	0.50	1
DQB1_2_55	32,740,768	*HLA-DQB1*	exon 2	G/T	0.103	0.153	7.3×10^{-5}	1.57	1.22 - 2.03	0.72	0.61	1
DQB1_2_27	32,740,796	*HLA-DQB1*	exon 2	A/C	0.121	0.116	0.80	0.96	0.74 - 1.24	0.57	0.64	1

[a]SNPs examined in other GWASs[4,5,8].

Table IV. Differences in the distribution of the HLA class II alleles between 525 lung adenocarcinoma cases and 525 controls

Gene	No.	Allele[a]	Frequency		OR	(95% CI)	P value by χ^2 test
			Control	Case			
DRB1	1	DRB1*1502	0.148	0.125	0.82	(0.64 - 1.05)	0.12
	2	DRB1*0901	0.117	0.157	1.40	(1.09 - 1.81)	0.0079
	3	DRB1*0405	0.106	0.154	1.53	(1.18 - 1.98)	0.0012
	4	DRB1*1302	0.097	0.058	0.58	(0.42 - 0.81)	0.0011
	5	DRB1*1501	0.091	0.055	0.58	(0.41 - 0.81)	0.0013
	6	DRB1*0803	0.073	0.075	1.02	(0.74 - 1.42)	0.90
	7	DRB1*0101	0.057	0.033	0.56	(0.37 - 0.86)	0.0077
	8	DRB1*1401/1405/1406/1412/1429	0.057	0.069	1.24	(0.87 - 1.77)	0.23
	9	DRB1*0802	0.052	0.037	0.70	(0.46 - 1.07)	0.10
	10	DRB1*0403/0404/0406	0.044	0.073	1.73	(1.19 - 2.52)	0.0040
	11	DRB1*1201/1202	0.040	0.050	1.25	(0.82 - 1.89)	0.30
	12	DRB1*1101	0.026	0.022	0.81	(0.46 - 1.42)	0.47
	13	DRB1*0410	0.021	0.016	0.78	(0.41 - 1.48)	0.44
	Total		0.929	0.924			
DQA1	1	DQA1*01	0.530	0.438	0.69	(0.58 - 0.82)	2.8×10^{-5}
	2	DQA1*03	0.348	0.444	1.50	(1.26 - 1.79)	6.6×10^{-6}
	3	DQA1*04/05/06	0.122	0.118	0.88	(0.68 - 1.14)	0.43
	Total		1.000	1.000			
DQB1	1	DQB1*0601	0.218	0.205	0.92	(0.75 - 1.14)	0.46
	2	DQB1*0303	0.139	0.164	1.22	(0.96 - 1.54)	0.11
	3	DQB1*0401	0.104	0.153	1.54	(1.19 - 2.00)	0.0010
	4	DQB1*0604	0.097	0.058	0.57	(0.41 - 0.79)	7.4×10^{-4}
	5	DQB1*0301	0.093	0.093	1.01	(0.75 - 1.35)	0.97
	6	DQB1*0602	0.086	0.053	0.59	(0.42 - 0.84)	0.0028
	7	DQB1*0302	0.081	0.092	1.15	(0.85 - 1.57)	0.35
	8	DQB1*0501	0.061	0.059	0.96	(0.67 - 1.37)	0.82
	9	DQB1*0402	0.042	0.051	1.21	(0.80 - 1.82)	0.36
	10	DQB1*0502	0.030	0.027	0.90	(0.54 - 1.51)	0.70
	11	DQB1*0503	0.028	0.028	1.03	(0.61 - 1.72)	0.93
	Total		0.979	0.983			
DR-DQ	1	DRB1*1502-DQA1*01-DQB1*0601	0.150	0.130	0.85	(0.66 - 1.09)	0.19
	2	DRB1*0901-DQA1*03-DQB1*0303	0.119	0.153	1.34	(1.05 - 1.73)	0.021
	3	DRB1*0405-DQA1*03-DQB1*0401	0.101	0.144	1.51	(1.16 - 1.96)	0.0022
	4	DRB1*1302-DQA1*01-DQB1*0604	0.094	0.057	0.58	(0.42 - 0.81)	0.0013
	5	DRB1*1501-DQA1*01-DQB1*0602	0.082	0.051	0.60	(0.42 - 0.85)	0.0042
	6	DRB1*0803-DQA1*01-DQB1*0601	0.070	0.072	1.02	(0.73 - 1.43)	0.89
	7	DRB1*0101-DQA1*01-DQB1*0501	0.047	0.034	0.72	(0.46 - 1.12)	0.14
	8	DRB1*0403/0404/0406-DQA1*03-DQB1*0302	0.040	0.068	1.72	(1.17 - 2.54)	0.0058
	9	DRB1*1201/1202-DQA1*04/05/06-DQB1*0301	0.032	0.032	0.99	(0.61 - 1.61)	0.97
	10	DRB1*1401/1405/1406/1412/1429-DQA1*01-DQB1*0503	0.028	0.031	1.13	(0.68 - 1.88)	0.64
	11	DRB1*1101-DQA1*04/05/06-DQB1*0301	0.025	0.017	0.68	(0.37 - 1.25)	0.21
	12	DRB1*0802-DQA1*04/05/06-DQB1*0402	0.024	0.026	1.10	(0.64 - 1.91)	0.73
	Total		0.812	0.815			

[a] *DRB1* and *DQB1* alleles linked to the DQA1*03 or DQA1*01 alleles and DR-DQ alleles containing the DQA1*03 or DQA1*01 alleles, which were significantly associated with lung ADC risk, are underlined.

Figure 1

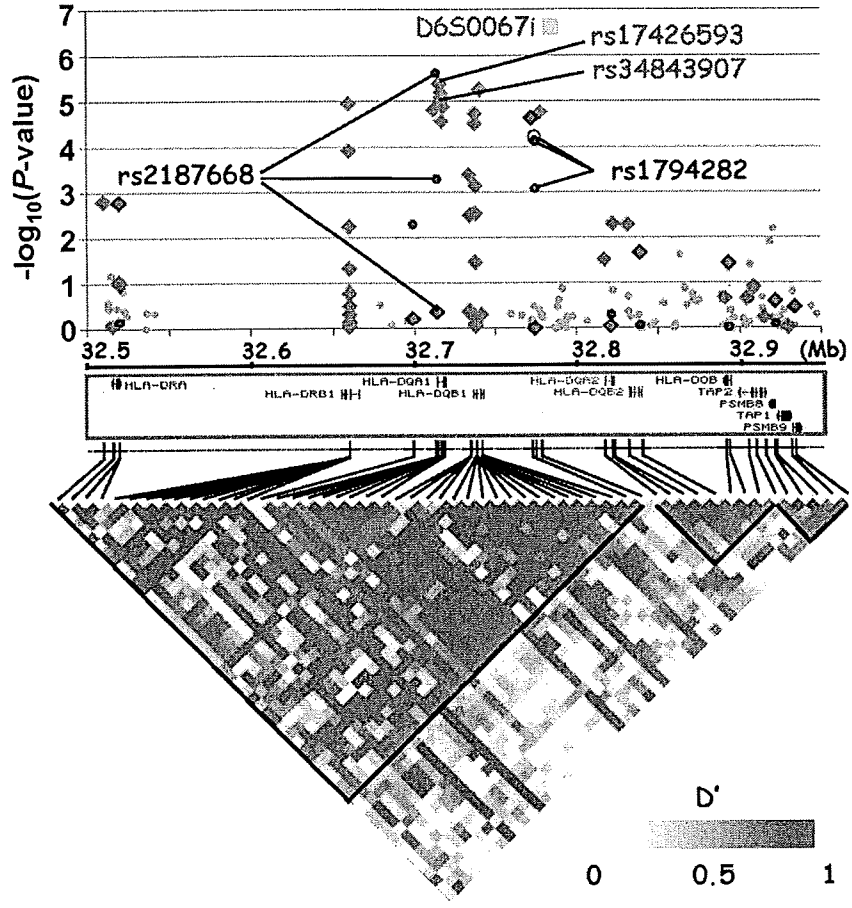
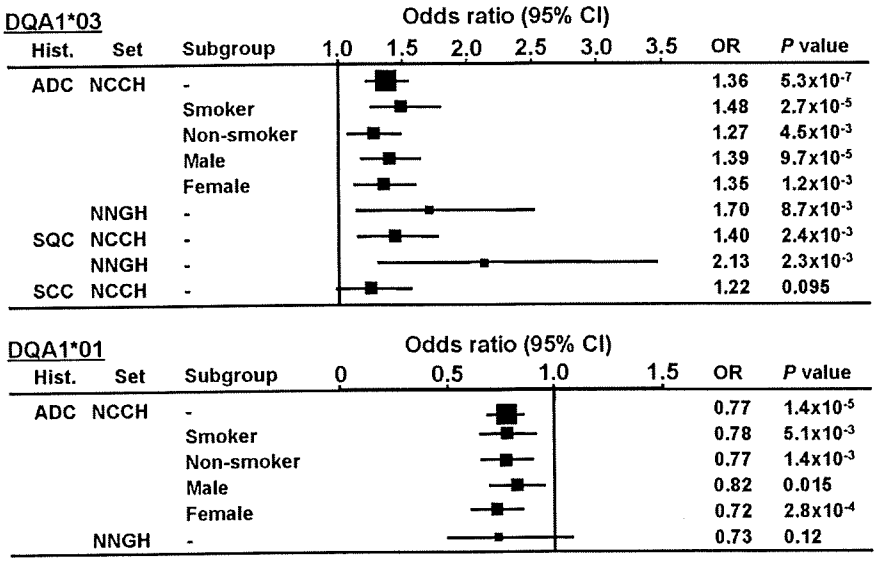

D6S0067i
rs17426593
rs34843907
rs2187668
rs1794282
-log10(P-value)
7
6
5
4
3
2
1
0
32.5
32.6
32.7
32.8
32.9
(Mb)
HLA-DRA
HLA-DRB1
HLA-DQA1
HLA-DQB1
HLA-DQA2
HLA-DQB2
HLA-DOB
TAP2
PSMB8
TAP1
PSMB9
D'
0
0.5
1

Figure 2a


DQA1*03
Odds ratio (95% CI)
Hist. Set Subgroup 1.0 1.5 2.0 2.5 3.0 3.5 OR P value
ADC NCCH - 1.36 5.3x10-7
Smoker 1.48 2.7x10-5
Non-smoker 1.27 4.5x10-3
Male 1.39 9.7x10-5
Female 1.36 1.2x10-3
NNGH - 1.70 8.7x10-3
SQC NCCH - 1.40 2.4x10-3
NNGH - 2.13 2.3x10-3
SCC NCCH - 1.22 0.095
DQA1*01
Odds ratio (95% CI)
Hist. Set Subgroup 0 0.5 1.0 1.5 OR P value
ADC NCCH - 0.77 1.4x10-5
Smoker 0.78 5.1x10-3
Non-smoker 0.77 1.4x10-3
Male 0.82 0.015
Female 0.72 2.8x10-4
NNGH - 0.73 0.12