# Influence of common genetic variation on lung cancer risk: meta-analysis of 14 900 cases and 29 485 controls

Maria N. Timofeeva[1], Rayjean J. Hung[2], Thorunn Rafnar[3], David C. Christiani[4], John K. Field[5], Heike Bickeböller[6], Angela Risch[7,41], James D. McKay[1], Yufei Wang[8], Juncheng Dai[9], Valerie Gaborieau[1], John McLaughlin[2], Darren Brenner[2], Steven A. Narod[10], Neil E. Caporaso[11], Demetrius Albanes[11], Michael Thun[12], Timothy Eisen[13], H.-Erich Wichmann[14,15], Albert Rosenberger[6], Younghun Han[16], Wei Chen[16], Dakai Zhu[16], Margaret Spitz[18], Xifeng Wu[17], Mala Pande[17], Yang Zhao[4], David Zaridze[19], Neonilia Szeszenia-Dabrowska[20], Jolanta Lissowska[21], Peter Rudnai[22], Eleonora Fabianova[23], Dana Mates[24], Vladimir Bencko[25], Lenka Foretova[26], Vladimir Janout[27], Hans E. Krokan[28], Maiken Elvestad Gabrielsen[28], Frank Skorpen[29], Lars Vatten[30], Inger Njølstad[31], Chu Chen[32], Gary Goodman[32], Mark Lathrop[33], Simone Benhamou[34], Tõnu Vooder[35], Kristjan Välk[36], Mari Nelis[37,38], Andres Metspalu[37], Olaide Raji[5], Ying Chen[5], John Gosney[5], Triantafillos Liloglou[5], Thomas Muley[39,41], Hendrik Dienemann[40,41], Gudmar Thorleifsson[3], Hongbing Shen[9], Kari Stefansson[3], Paul Brennan[1], Christopher I. Amos[16,42,*†], Richard Houlston[8,†], Maria Teresa Landi[11,†] and for TRICL Research Team[‡]

## *Affiliations*

[1]International Agency for Research on Cancer, Lyon 69008, France, [2]Samuel Lunenfeld Research Institute of Mount Sinai Hospital, Toronto, ON, Canada M5G 1X5, [3]DeCODE Genetics, Reykjavik, Iceland, [4]Harvard University School of Public Health, Boston, MA 02115, USA, [5]Roy Castle Lung Cancer Research Programme, Department of Molecular and Clinical Cancer Medicine, The University of Liverpool Cancer Research Centre, Institute of Translational Medicine, The University of Liverpool, Liverpool, UK, [6]Department of Genetic Epidemiology, University Medical Center, Georg-August-University Göttingen, 37073 Göttingen, Germany, [7]DKFZ - German Cancer Research Center, Translational Lung Research Centre Heidelberg (TLRC-H), Member of the German Center for Lung Research, 69120 Heidelberg, Germany, [8]Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, Surrey SM2 5NG, UK, [9]Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, Nanjing, China, [10]Women's College Research Institute, Toronto, Ontario, Canada, [11]Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, U.S. Public Health Service, Bethesda, MD 20892, USA, [12]American Cancer Society, Epidemiology and Surveillance Research, Atlanta, GA 30301, USA, [13]Department of Oncology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK, [14]Helmholtz Center Munich, Institute of Epidemiology I, 85764 Neuherberg, Germany, [15]Chair of Epidemiology, Institute of Medical Informatics, Biometry and Epidemiology, Ludwig-Maximilians-Universität Munich, 81377 Munich, Germany, [16]Department of Genetics and [17]Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA, [18]Baylor

College of Medicine, Houston, TX 77030, USA, [19]Institute of Carcinogenesis, Russian N.N. Blokhin Cancer Research Centre, 115478 Moscow, Russia, [20]Department of Epidemiology, Institute of Occupational Medicine, 91348 Lodz, Poland, [21]The M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw 02781, Poland, [22]National Institute of Environmental Health, Budapest 1097, Hungary, [23]Regional Authority of Public Health, Banska´ Bystrica 97556, Slovak Republic, [24]National Institute of Public Health, Bucharest 050463, Romania, [25]1st Faculty of Medicine, Institute of Hygiene and Epidemiology, Charles University in Prague, 128 00 Prague 2, Czech Republic, [26]Department of Cancer Epidemiology and Genetics, Masaryk Memorial Cancer Institute, Brno 65653, Czech Republic, [27]Palacky University, Olomouc 77515, Czech Republic, [28]Department of Cancer Research and Molecular Medicine, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim 7489, Norway, [29]Department of Laboratory Medicine, Children's and Women's Health, Faculty of Medicine and [30]Department of Public Health and General Practice, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim 7489, Norway, [31]Department of Community Medicine, University of Tromso, Tromso 9037, Norway, [32]Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA, [33]Centre d'Etude du Polymorphisme Humain (CEPH), Paris 75010, France, [34]INSERM U946, Paris 75010, France, [35]Institute of Molecular and Cell Biology, University of Tartu, Tartu 51010, Estonia, [36]Competence Centre on Reproductive Medicine and Biology, 50410 Tartu, Estonia, [37]Estonian Genome Center, Institute of Molecular and Cell Biology, Tartu 51010, Estonia, [38]Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland, [39]Translational Research Unit and [40]Department of Thoracic Surgery, Thoraxklinik am Universitätsklinikum Heidelberg, Germany, [41]Translational Lung Research Centre Heidelberg (TLRC-H), Member of the German Center for Lung Research, 69126 Heidelberg, Germany and [42]Department of Community and Family Medicine, Geisel School of Medicine, Dartmouth College, Hanover, NH 03750, USA

[†]These authors contributed equally to this work.
[‡]Transdisciplinary Research in Cancer of the Lung.


*Corresponding author:*


Christopher I. Amos


Department of Community and Family Medicine,
Geisel School of Medicine,
Dartmouth College,
46 Centerra Parkway Suite 330,
Lebanon, NH 03766.
Tel: +603-653-6696;
Email: Christopher.I.Amos@Dartmouth.edu

*Participating study descriptions*

**Studies under NCI umbrella**

Detailed characteristic of NCI studies is presented elsewhere (1). Briefly it includes samples from studies:

1.  The Environment and Genetics in Lung cancer Etiology (EAGLE) is a population-based case-control study including 2,100 lung cancer cases and 2,120 healthy controls enrolled in Italy between 2002 and 2005 (2).

2.  The Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study (ATBC) is a randomized primary prevention trial including 29,133 male smokers enrolled in Finland between 1985 and 1993.

3.  The Prostate, Lung, Colon, Ovary Screening Trial (PLCO) is a randomized trial including 150,000 individuals enrolled in ten U.S. study centers between 1992 and 2001 (3).

4.  The Cancer Prevention Study II Nutrition Cohort (CPS-II) is a cohort study including over 183,000 subjects enrolled by the American Cancer Society between 1992 and 2001 across all U.S. states (4).

The lung cancer diagnosis in the NCI GWAS was based on clinical criteria and confirmed by pathology reports from surgery, biopsy or cytology samples in approximately 95% of cases and on clinical history and imaging for the remaining 5%. Tumor histology was coded according to the International Classification of Disease fir Oncology (ICD-O). Overall, between 10% and 50% of all diagnosis from the NCI GWAS were centrally reviewed by expert lung pathologies from NCI. For the histology analysis the following codes were used - non small cell carcinoma: adenocarcinoma (8140, 8251, 8255, 8260, 8310, 8323, 8480, 8481, and 8570), bronchioloalveolar carcinoma (8250, 8252, and 8253), large cell carcinoma (8012 and 8031), squamous cell carcinoma (8052, 8070, 8071, 8072, 8073, and 8074), and other non-small cell carcinoma (ONSCLC) (8010, 8020, 8022, 8032, 8033, 8046, 8050, 8490, 8550, and 8560); small cell carcinoma: (NOS 804 8041/3): 8042/3 oat cell carcinoma, 8043/3 small cell carcinoma, fusiform cell, 8044/3 Small cell carcinoma, intermediate cell, 8045/3, combined small cell carcinoma.

Overall 5,739 cases and 5,848 controls were included after initial quality control as previously described (1). Additionally 10 cases and 107 controls were excluded due to updated diagnoses in cancer cases, controls later identified with lung cancer in follow-up of cohort studies and quality control related issues. The final series used for the current meta-analysis included 5,729 cases and 5,741 controls (Table 1, Table S1). For the imputed data, additional quality control procedures were performed (Table S2), which led to a total of 5,719 cases and 5,735 controls.

**Studies under IARC umbrella**

The IARC lung cancer GWAS includes individuals from the following studies described previously: (i) Carotene and Retinol Efficacy Trial (CARET) cohort (5), (ii) the Central Europe multicenter hospital-based case-control (6,7), (iii) the hospital-based case-control study from France (8), (iv) the hospital based case-control lung cancer study from Estonia (9,10) and (v) the population-based HUNT2 and Tromsø 4 lung cancer studies (7,11,12). In total the combined

dataset included 3,062 cases and 4,455 controls. The diagnosis of lung cancer for all studies was established based on clinical criteria and confirmed by pathology reports from surgery, biopsy or cytology samples. The International Classification of Diseases for Oncology (ICD-O) was used. No information on histology was available for the case-control lung cancer study from Estonia. Therefore Estonia study did not contribute to the histology subgroup analysis. After applying standardized quality control procedures, 2,894 cases and 4,233 controls entered the data analysis for the TRICL meta-analysis project (Table S1, Table 1).

**Toronto/Samuel Lunenfeld Research Institute (SLRI) study**

The Toronto study was conducted in the Great Toronto Area between 1997 and 2002. Cases were recruited at the hospitals in the network of University of Toronto and Samuel Lunenfeld Research Institute. At the time of recruitment in the clinical setting, provisional diagnoses of lung carcinoma were first assigned based on clinical criteria. Diagnoses for all cases included were histologically confirmed by the reference pathologist who is a specialist in pulmonary pathology, based on review of pathology reports from surgery, biopsy or cytology samples in 100% of cases. Diagnostic classification was done initially according to ICD-9, ICD-10, and ICD for oncology-2, and subsequently converted to ICD-O-3. Tumors were grouped into the major categories included in this analysis according to primary cancer type based on the ICD-3 definitions. Controls were randomly selected from individual visiting family medicine clinics and Ministry of Finance Municipal Tax Tapes. All subjects were interviewed using a standard questionnaire and information on lifestyle risk factors, occupational history, medical and family history was collected. Blood samples were collected from more than 85% of the subjects. After applying the standardized quality control procedures and restricting to the study participants with European ancestry, 331 cases and 499 controls were included in the TRICL meta-analysis (7).

**deCODE Genetics**

The Icelandic lung cancer study population has been described previously (13). Briefly, according to the population-based Icelandic Cancer Registry (www.krabbameinsskra.is), a total of 4,252 lung cancer patients were diagnosed from January 1, 1955, to December 31, 2010. The registry receives information from all pathology and cytology laboratories in Iceland, in addition to hematology laboratories, hospital wards, private medical practitioners and other individual health care workers. Approximately 94.5% of diagnoses in the ICR have histological confirmation. Histological subtype is registered in the Icelandic Cancer Registry using the ICD-03 classification. For the histology analysis the following codes were used - adenocarcinoma: 8140/3, 8250/3, 8260/3, 8310/3, 8480/3 8560/3; large cell carcinoma: 8012/3, 8031/3; squamous cell carcinoma: 8070/3, 8071/3, 8072/3, 8074/3; other non-small cell carcinomas: 8010/3, 8020/3, 8021/3, 8032/3, 8230/3; small cell carcinoma: 8041/3, 8042/3, 8043/3, 8044/3, 8045/3. Recruitment of both prevalent and incident cases was initiated in the year 1998, the recruitment is ongoing and DNA samples from lung cancer cases are subjected to whole-genome genotyping as they are collected.

The controls used in this study consisted of individuals from other ongoing genome-wide association studies at deCODE, age and sex matched to the group of cases. No individual disease group accounts for more than 10% of the total control group.

Quality control for chip genotyping: SNPs were excluded if they had (i) yield lower than 95%, (ii) minor allele frequency less than 1% in the population or (iii) significant deviation from Hardy-Weinberg equilibrium in the controls ($P < 0.001$), (iv) if they produced an excessive inheritance error rate (*i.e.* >0.001), (v) if there was substantial difference in allele frequency between chip types (from just a single chip if that resolved all differences, but from all chips otherwise). All samples having a call rate <97% were excluded from the analysis.

Genotype data from 830 lung cancer cases and 11,228 controls were included in the present study.

## Institute of Cancer Research (ICR) lung cancer study

Cases (1,182 male; mean age at diagnosis 57 years, SD 6) with pathologically confirmed lung cancer were ascertained through the Genetic Lung Cancer Predisposition Study (GELCAPS) (14). Diagnostic classification was done initially according to ICD-9. All cases were British residents and self-reported to be of European Ancestry. Genotype frequencies were compared with publicly accessible genotype data generated by the UK Wellcome Trust Case-Control Consortium 2 (WTCCC2) study of 2,699 individuals from the 1958 British Birth Cohort (58C) (15) that had been genotyped using Illumina Human1.2M-Duo Custom_v1 Array BeadChips. Comprehensive details of on genotyping and quality control have been previously published (16,17).

Ethical approval was obtained from the London Multi-Centre Research Ethics Committee (MREC/98/2/67) in accordance with the tenets of the Declaration of Helsinki. All participants provided written informed consent.

## Liverpool Lung Project

Incident cases were ascertained through the Liverpool Lung Project, an ongoing molecular epidemiological study conducted in Liverpool, United Kingdom (18). All cases had histologically or cytologically confirmed lung cancer and were aged 20-80 years at diagnosis. Histological type is recorded according to the classification based on the WHO typing of tumours (World Health Organization. (19) along with the updated classification of adenocarcinomas (20). SNOMED (**S**ystematized **No**menclature of **Med**icine) classification was used for coding the site and histological diagnosis. The details are outlined in the Royal College of Pathology UK publication; Standards and Datasets for Reporting Cancers Dataset for lung cancer histopathology reports (3rd edition) April 2011, (http://www.rcpath.org/Resources/RCPath/Migrated%20Resources/Documents/G/g048datasetlu ngapril11.pdf).

A standardized lifestyle questionnaire was used to collect detailed information on demographic characteristics, medical history, family history of cancer, history of tobacco consumption and lifetime occupational history. Genotype frequencies were compared with publicly accessible genotype data generated by the UK Wellcome Trust Case-Control Consortium 2 (WTCCC2) study of 2,501 individuals from the UK Blood Service collections (UKBS) that had been genotyped using Illumina Human1.2M-Duo Custom_v1 Array BeadChips.

The study protocol was approved by the Liverpool Research Ethic Committee and all research participants provided written, informed consent in accordance with the Declaration of Helsinki.

**Harvard Lung Cancer Study**

For the Harvard Lung Cancer Susceptibility Study, details of participant recruitment have been described previously (21). The genotyped dataset includes 1,000 cases and 1,000 controls. Cases were patients over the age of 18 years, with newly diagnosed and histologically confirmed primary non-small cell lung cancer; controls were healthy non-blood-related family members and friends of patients with cancer or with cardiothoracic conditions undergoing surgery. The histological classification was done by two staff pulmonary pathologists at the Massachusetts General Hospital according to the International Classification of Diseases for Oncology (ICD-O3). For the histology analysis the following codes were used - adenocarcinoma: 8140/3, 8250/3, 8260/3, 8310/3, 8480/3 8560/3; large cell carcinoma: 8012/3, 8031/3; squamous cell carcinoma: 8070/3, 8071/3, 8072/3, 8074/3; other non-small cell carcinomas: 8010/3, 8020/3, 8021/3, 8032/3, 8230/3; and small cell carcinoma: 8041/3, 8042/3, 8043/3, 8044/3, 8045/3.
A general quality control procedure for GWAS data was performed on the genotyped dataset. The procedure included assessing the failure rate per individual and per SNP, identifying individuals with discordant gender information, assessing the degree of the relatedness and identifying population outliers. After all quality control procedures were done, the finalized dataset included 984 cases and 970 controls with 543,697 SNPs.

**MDACC**

Cases and controls were ascertained from a case-control study that has been ongoing at the U.T. M.D. Anderson Cancer Center since 1997. Cases are newly diagnosed, histologically-confirmed patients presenting at M.D. Anderson Cancer and who had not previously received treatment other than surgery. Histological information about tumors was derived by chart review from medical records. Tumors were histologically classified as adenocarcinoma, squamous cell carcinoma, large cell carcinoma, mixed histologies, or unclassified NSCLC (22). Small cell carcinomas were excluded from the genome-wide association analysis conducted at M.D. Anderson Cancer Center. Controls are healthy individuals seen for routine care at Kelsey-Seybold Clinics; the largest physician group-practice plan in the Houston Metropolitan area. Controls were frequency matched to cases according to their smoking behavior, age in 5 year categories, ethnicity, and sex. Former smoking controls were further frequency matched to former smoking cases according to the number of years since smoking cessation (in 5 year categories).

In addition to quality control procedures previously reported by Amos et al. (2008) (23), 7 samples were removed in EIGENSTRAT analysis. After all quality control procedures were completed, 2284 samples including 1150 cases and 1134 controls were available for genotyping.

**HGF Germany**

The HGF GWA study was made up of three independent German studies as detailed below: in total 506 incident lung cancer cases (LUCY-study: n=305, Heidelberg lung cancer case-control study: n=201) were compared to 480 population controls (KORA surveys KORA). After

excluding individuals with missing values and potentially related individuals 487 cases and 480 controls entered the data analysis for the TRICL meta-analysis project.

LUCY (LUng Cancer in the Young) is a multicenter study with 31 recruiting hospitals in Germany (24,25). The study is conducted by the Institute of Epidemiology, Helmholtz Zentrum Muenchen, and the Department of Genetic Epidemiology, Medical School, University of Göttingen). The LUCY-study provides access to a nationwide, population based family and a case-control sample (control population KORA, described below) of lung cancer patients aged 50 years or younger at diagnosis. Detailed epidemiologic data have been collected including data on medical history, education, family history of cancer and smoking exposure by phase assessment. Blood samples are taken and DNA and lymphoblastoid cell lines are prepared of all cases and controls and of parts of the relatives. Phenotype data of 847 young patients with primary lung cancer and 5524 relatives have been collected. All cases were histologically confirmed and classified by a pathologist as adenocarcinoma, squamous cell carcinoma, large cell carcinoma, small cell carcinoma, other or unclassified histology.

Heidelberg lung cancer case-control study is an ongoing hospital based case-control study (25) (26). The German Cancer Research Center (DKFZ) has recruited over 2000 lung cancer cases at and in collaboration with the Thoraxklinik Heidelberg, including 300 lung cancer cases with onset of disease at the age of ≤ 50. The histological classification was done by a pathologist, according to the WHO International Histological Classification of Tumours (ICD-O) current at the time of recruitment. Approximately 750 hospital-based controls have also been recruited. Data on occupational exposure, tobacco smoking, educational status, and for a subgroup also on family history of lung cancer, assessed by a self-administered questionnaire is available. Blood samples have been taken, and DNA has been extracted.

KORA (Cooperative health research in the Region of Augsburg) survey is a population-based KORA platform established by the Helmholtz Center Munich (27). In total, four population based health surveys have been conducted during 1984/85-1999/2001. Overall 18000 participants in the age range between of 25 and 74 years at first interview were recruited. Detailed information on demographic characteristics, medical history, history of tobacco consumption and lifetime occupation together with biological materials were collected for more than 16000 proband

**Han Chinese GWAS**

Han Chinese GWAS includes 2,338 cases and 3,077 controls from the Nanjing study and Beijing study as described previously (28). All these case-control sample sets were hospital based and the subjects were unrelated ethnic Han Chinese. The case patients were histopathologically or cytologically confirmed lung cancer by at least two local pathologists. Tumors were histologically classified as adenocarcinoma, squamous cell carcinoma, large cell carcinoma, mixed histologies, or unclassified NSCLC. All cancer-free control subjects were selected from those receiving routine physical examinations in local hospitals or those participating in the community screening of non-communicable diseases and frequency-matched for age, gender, and geographic regions to each set of the lung cancer patients. At recruitment, informed consent was obtained from each subject, and this study was approved by the Institutional Review Boards of each participating institution.

*Quality control*

Samples were excluded if (i) the average call rate was < 90%, (ii) there was sex discordance (threshold of heterozygosity > 10% for males and < 20% for females), (iii) unexpected duplicates and evidence of first-degree relative relationships from identity-by-descent values, (iv) heterozygosity rates for autosomal chromosomes exceeded 6 standard deviation of the mean, (v) based on STRUCTURE analysis (29) subjects were less than 80% European ancestry and (vi) detected as outliers based on principal component analysis (PCA) using EIGENSTRAT (30). SNPs were excluded if (i) genotyped call rate was less than 95%. Test for Hardy-Weinberg equilibrium (HWE) were also performed, but no SNPs were excluded based on this test. However, each study center was requested to report this information together with results of the analysis for the specific study.

Reference List

1. Landi,M.T., Chatterjee,N., Yu,K., Goldin,L.R., Goldstein,A.M., Rotunno,M., Mirabello,L., Jacobs,K., Wheeler,W., Yeager,M.*, et al.* (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**, 679-691.

2. Landi,M.T., Consonni,D., Rotunno,M., Bergen,A.W., Goldstein,A.M., Lubin,J.H., Goldin,L., Alavanja,M., Morgan,G., Subar,A.F.*, et al.* (2008) Environment And Genetics in Lung cancer Etiology (EAGLE) study: an integrative population-based case-control study of lung cancer. *BMC. Public Health*, **8**, 203.

3. Hayes,R.B., Sigurdson,A., Moore,L., Peters,U., Huang,W.Y., Pinsky,P., Reding,D., Gelmann,E.P., Rothman,N., Pfeiffer,R.M.*, et al.* (2005) Methods for etiologic and early marker investigations in the PLCO trial. *Mutat. Res.*, **592**, 147-154.

4. Calle,E.E., Rodriguez,C., Jacobs,E.J., Almon,M.L., Chao,A., McCullough,M.L., Feigelson,H.S., Thun,M.J. (2002) The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. *Cancer*, **94**, 2490-2501.

5. Omenn,G.S., Goodman,G., Thornquist,M., Grizzle,J., Rosenstock,L., Barnhart,S., Balmes,J., Cherniack,M.G., Cullen,M.R., Glass,A., . (1994) The beta-carotene and retinol efficacy trial (CARET) for chemoprevention of lung cancer in high risk populations: smokers and asbestos-exposed workers. *Cancer Res.*, **54**, 2038s-2043s.

6. Scelo,G., Constantinescu,V., Csiki,I., Zaridze,D., Szeszenia-Dabrowska,N., Rudnai,P., Lissowska,J., Fabianova,E., Cassidy,A., Slamova,A.*, et al.* (2004) Occupational exposure to vinyl chloride, acrylonitrile and styrene and lung cancer risk (europe). *Cancer Causes Control*, **15**, 445-452.

7. Hung,R.J., McKay,J.D., Gaborieau,V., Boffetta,P., Hashibe,M., Zaridze,D., Mukeria,A., Szeszenia-Dabrowska,N., Lissowska,J., Rudnai,P.*, et al.* (2008) A susceptibility locus for

lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature.*, **452**, 633-637.

8. Feyler,A., Voho,A., Bouchardy,C., Kuokkanen,K., Dayer,P., Hirvonen,A., Benhamou,S. (2002) Point: myeloperoxidase -463G --> a polymorphism and lung cancer risk. *Cancer Epidemiol. Biomarkers Prev.*, **11**, 1550-1554.

9. Nelis,M., Esko,T., Magi,R., Zimprich,F., Zimprich,A., Toncheva,D., Karachanak,S., Piskackova,T., Balascak,I., Peltonen,L.*, et al.* (2009) Genetic structure of Europeans: a view from the North-East. *PLoS One.*, **4**, e5472.

10. Valk,K., Vooder,T., Kolde,R., Reintam,M.A., Petzold,C., Vilo,J., Metspalu,A. (2010) Gene expression profiles of non-small cell lung cancer: survival prediction and new biomarkers. *Oncology*, **79**, 283-292.

11. Holmen,J., Midthjell,K., Krüger,O., Langhammer,A., Holmen,T.L., Bratberg,G.H., Vatten,L., Lund-Larsen,P.G. (2003) The Nord-Trøndelag Health Study 1995-97 (HUNT 2):Objectives, contents, methods and participation. *Norsk Epidemiologi*, **13**, 19-32.

12. Jacobsen,B.K., Eggen,A.E., Mathiesen,E.B., Wilsgaard,T., Njolstad,I. (2012) Cohort profile: The Tromso Study. *Int. J. Epidemiol.*, **41**, 961-967.

13. Thorgeirsson,T.E., Geller,F., Sulem,P., Rafnar,T., Wiste,A., Magnusson,K.P., Manolescu,A., Thorleifsson,G., Stefansson,H., Ingason,A.*, et al.* (2008) A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature.*, **452**, 638-642.

14. Eisen,T., Matakidou,A., Houlston,R. (2008) Identification of low penetrance alleles for lung cancer: the GEnetic Lung CAncer Predisposition Study (GELCAPS). *BMC. Cancer*, **8**, 244.

15. Power,C., Elliott,J. (2006) Cohort profile: 1958 British birth cohort (National Child Development Study). *Int. J. Epidemiol.*, **35**, 34-41.

16. Wang,Y., Broderick,P., Webb,E., Wu,X., Vijayakrishnan,J., Matakidou,A., Qureshi,M., Dong,Q., Gu,X., Chen,W.V.*, et al.* (2008) Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat. Genet.*, **40**, 1407-1409.

17. Broderick,P., Wang,Y., Vijayakrishnan,J., Matakidou,A., Spitz,M.R., Eisen,T., Amos,C.I., Houlston,R.S. (2009) Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. *Cancer Res.*, **69**, 6633-6641.

18. Field,J.K., Smith,D.L., Duffy,S., Cassidy,A. (2005) The Liverpool Lung Project research protocol. *Int. J. Oncol.*, **27**, 1633-1645.

19. Travis, W. D., Brambilla, E., Muller-Hermelink, K H, and Harris, C. C.(2004) *Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart*. IARC*Press*, Lyon.

20. Travis,W.D., Brambilla,E., Noguchi,M., Nicholson,A.G., Geisinger,K.R., Yatabe,Y., Beer,D.G., Powell,C.A., Riely,G.J., Van Schil,P.E.*, et al.* (2011) International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J. Thorac. Oncol.*, **6**, 244-285.

21. Asomaning,K., Miller,D.P., Liu,G., Wain,J.C., Lynch,T.J., Su,L., Christiani,D.C. (2008) Second hand smoke, age of exposure and lung cancer risk. *Lung Cancer*, **61**, 13-20.

22. Bosken,C.H., Wei,Q., Amos,C.I., Spitz,M.R. (2002) An analysis of DNA repair as a determinant of survival in patients with non-small-cell lung cancer. *J. Natl. Cancer Inst.*, **94**, 1091-1099.

23. Amos,C.I., Wu,X., Broderick,P., Gorlov,I.P., Gu,J., Eisen,T., Dong,Q., Zhang,Q., Gu,X., Vijayakrishnan,J.*, et al.* (2008) Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat. Genet.*, **40**, 616-622.

24. Rosenberger,A., Illig,T., Korb,K., Klopp,N., Zietemann,V., Wolke,G., Meese,E., Sybrecht,G., Kronenberg,F., Cebulla,M.*, et al.* (2008) Do genetic factors protect for early onset lung cancer? A case control study before the age of 50 years. *BMC. Cancer.*, **8:60.**, 60.

25. Sauter,W., Rosenberger,A., Beckmann,L., Kropp,S., Mittelstrass,K., Timofeeva,M., Wolke,G., Steinwachs,A., Scheiner,D., Meese,E.*, et al.* (2008) Matrix metalloproteinase 1 (MMP1) is associated with early-onset lung cancer. *Cancer Epidemiol. Biomarkers Prev.*, **17**, 1127-1135.

26. Dally,H., Edler,L., Jager,B., Schmezer,P., Spiegelhalder,B., Dienemann,H., Drings,P., Schulz,V., Kayser,K., Bartsch,H., Risch,A. (2003) The CYP3A4*1B allele increases risk for small cell lung cancer: effect of gender and smoking dose. *Pharmacogenetics.*, **13**, 607-618.

27. Wichmann,H.E., Gieger,C., Illig,T. (2005) KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen*, **67 Suppl 1**, S26-S30.

28. Hu,Z., Wu,C., Shi,Y., Guo,H., Zhao,X., Yin,Z., Yang,L., Dai,J., Hu,L., Tan,W.*, et al.* (2011) A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. *Nat. Genet.*, **43**, 792-796.

29. Falush,D., Stephens,M., Pritchard,J.K. (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567-1587.

30. Price,A.L., Patterson,N.J., Plenge,R.M., Weinblatt,M.E., Shadick,N.A., Reich,D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*, **38**, 904-909.

**Table S1. Characteristic of the participating studies by subgroups**

| Subgroups | | MDACC Cases | MDACC Controls | Liverpool Cases | Liverpool Controls | ICR Cases | ICR Controls | SLRI Cases | SLRI Controls | IARC Cases | IARC Controls | deCODE Cases | deCODE Controls | HGF Cases | HGF Controls | Harvard Cases | Harvard Controls | NCI Cases | NCI Controls | Han Chinese Cases | Han Chinese Controls | Overall Cases | Overall Controls |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ALL | | 1,150 | 1,134 | 543 | 2,501 | 1,952 | 2,699 | 331 | 499 | 2,894 | 4,233 | 830 | 11,228 | 487 | 480 | 984 | 970 | 5,729 | 5,741 | 2,338 | 3,077 | 17,238 | 32,562 |
| Gender | Men | 655 | 644 | 314 | n/a | 1,166 | n/a | 159 | 190 | 2,176 | 2,843 | 414 | 5,721 | 303 | 243 | 507 | 442 | 4,476 | 4,841 | 1,717 | 2,084 | 11,887 | 17,008 |
| | Women | 495 | 490 | 229 | n/a | 786 | n/a | 172 | 309 | 718 | 1,390 | 416 | 5,508 | 184 | 237 | 477 | 528 | 1,253 | 900 | 621 | 993 | 5,351 | 10,355 |
| Smoking | Never | n/a | n/a | 32 | 2,501 | 109 | 2,699 | 95 | 217 | 178 | 1,401 | 34 | 634 | 35 | 220 | 92 | 161 | 352 | 1,376 | 829 | 1,266 | 1,756 | 10,475 |
| | Former | 572 | 626 | 239 | 2,501 | 762 | 2,699 | 95 | 139 | 521 | 993 | n/a | n/a | 49 | 122 | 502 | 555 | 1,933 | 2,006 | 254 | 226 | 5,928 | 11,226 |
| | Current | 578 | 508 | 251 | 2,501 | 1,065 | 2,699 | 90 | 89 | 2,073 | 1,699 | n/a | n/a | 399 | 138 | 390 | 254 | 3,423 | 2,338 | 1,255 | 1,585 | 8,269 | 10,226 |
| | Ever | 1,150 | 1,134 | 490 | 2,501 | 1,843 | 2,699 | 236 | 272 | 2,709 | 2,797 | 701 | 7,586 | 452 | 260 | 892 | 809 | 5,356 | 4,344 | n/a | n/a | 13,829 | 22,402 |
| Age | ≤ 50 (YLC) | 193 | 137 | n/a | n/a | n/a | n/a | 47 | 248 | 381 | 1,146 | 67 | 2,720 | 487 | 480 | 81 | 214 | 100 | 124 | 399 | 442 | 1,755 | 5,511 |
| | > 50 (OLC) | 957 | 997 | n/a | n/a | n/a | n/a | 284 | 251 | 2,513 | 3,087 | 763 | 8,509 | n/a | n/a | 903 | 756 | 5,628 | 5,617 | 1,939 | 2,635 | 12,987 | 21,852 |
| Histology | Adenocarcinoma | 619 | 1,134 | 169 | 2,501 | 465 | 2,699 | 90 | 499 | 595 | 3,237 | 346 | 11,228 | 192 | 480 | 488 | 970 | 1,849 | 5,741 | 1,307 | 3,077 | 6,120 | 31,566 |
| | Squamous cell | 306 | 1,134 | 167 | 2,501 | 611 | 2,699 | 50 | 499 | 1,031 | 3,382 | 179 | 11,228 | 100 | 480 | 215 | 970 | 1,451 | 5,741 | 825 | 3,077 | 4,935 | 31,711 |
| | Small cell | n/a | n/a | 57 | 2,501 | 530 | 2,699 | 22 | 499 | 451 | 3,382 | 98 | 11,228 | 108 | 480 | n/a | n/a | 706 | 5,741 | 179 | 3,077 | 2,151 | 29,607 |
| | Large cell | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | 32 | 11,228 | n/a | n/a | 81 | 970 | 170 | 5,741 | n/a | n/a | 283 | 17,939 |
| | Non Small Cell | | | 411 | 2,501 | 1,420 | 2,699 | | | | | | | | | | | | | | | 1,831 | 5,200 |
| | other | 224 | 1,134 | | | | | | | | | | | | | | | | | | | 224 | 1,134 |
| Stage | I/II stage for NSCLC | 378 | 1,134 | n/a | n/a | 92 | 2,699 | 48 | 499 | n/a | n/a | n/a | n/a | n/a | n/a | 455 | 970 | 1,434 | 5,741 | n/a | n/a | 2,407 | 11,043 |
| | III/IV stage for NSCLC | 707 | 1,134 | n/a | n/a | 320 | 2,699 | 103 | 499 | n/a | n/a | n/a | n/a | n/a | n/a | 529 | 970 | 2,066 | 5,741 | n/a | n/a | 3,725 | 11,043 |
| | limited stage | n/a | n/a | n/a | n/a | 56 | 2,699 | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | 56 | 2,699 |
| | extended | n/a | n/a | n/a | n/a | 104 | 2,699 | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | n/a | 104 | 2,699 |
| Family history | Yes | 241 | 167 | 96 | 2,501 | 285 | 2,699 | n/a | n/a | 148 | 113 | 208 | 1,410 | n/a | n/a | 166 | 110 | 783 | 556 | n/a | n/a | 1,927 | 7,556 |
| | No | 904 | 964 | 427 | 2,501 | 1,667 | 2,699 | 212 | 380 | 1,716 | 3,153 | 622 | 9,818 | n/a | n/a | 817 | 860 | 4,265 | 4,728 | n/a | n/a | 10,630 | 25,103 |

Supplementary Tables

**Table S2 Genome-Wide Significant Result (p<5 x 10$^{-8}$) for the Overall Meta-Analysis**

| rs number | Chr | Position (NCBI Build36) | Allele Reference | Effect | Number of studies | N cases | N controls | MAF cases | MAF controls | Heterogeneity P$^1$ | I$^2$ (%) | Fixed effect model OR | P | Random effect model OR | P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| rs1051730 | 15 | 76681394 | G | A | 9 | 14879 | 29427 | 0.40 | 0.34 | 0.99 | 0 | 1.32 | 2.2E-63 | 1.32 | 2.2E-63 |
| rs8034191 | 15 | 76593078 | T | C | 9 | 14887 | 29446 | 0.40 | 0.35 | 0.98 | 0 | 1.31 | 9.5E-59 | 1.31 | 9.5E-59 |
| rs6495309 | 15 | 76702300 | C | T | 9 | 14893 | 29465 | 0.18 | 0.22 | 0.73 | 0 | 0.79 | 1.1E-32 | 0.79 | 1.1E-32 |
| rs6495314 | 15 | 76747584 | A | C | 9 | 14879 | 29433 | 0.44 | 0.40 | 0.85 | 0 | 1.21 | 2.2E-32 | 1.21 | 2.2E-32 |
| rs11638372 | 15 | 76770614 | C | T | 9 | 14887 | 29435 | 0.44 | 0.40 | 0.84 | 0 | 1.21 | 5.0E-31 | 1.21 | 5.0E-31 |
| rs1394371 | 15 | 76511524 | C | T | 9 | 14889 | 29442 | 0.33 | 0.29 | 0.85 | 0 | 1.21 | 1.2E-28 | 1.21 | 1.2E-28 |
| rs12914385 | 15 | 76685778 | C | T | 3 | 7186 | 7166 | 0.46 | 0.40 | 0.74 | 0 | 1.32 | 2.8E-28 | 1.32 | 2.8E-28 |
| rs2036534 | 15 | 76614003 | T | C | 7 | 12369 | 24213 | 0.19 | 0.22 | 0.72 | 0 | 0.80 | 3.3E-25 | 0.80 | 3.3E-25 |
| rs13180 | 15 | 76576543 | T | C | 9 | 14872 | 29409 | 0.35 | 0.39 | 0.35 | 10.3 | 0.85 | 6.2E-23 | 0.85 | 9.3E-20 |
| rs4887053 | 15 | 76499754 | C | A | 9 | 14895 | 29458 | 0.18 | 0.21 | 0.59 | 0 | 0.83 | 2.6E-20 | 0.83 | 2.6E-20 |
| rs4887077 | 15 | 76765419 | C | T | 7 | 8180 | 22616 | 0.44 | 0.40 | 0.81 | 0 | 1.20 | 2.4E-17 | 1.20 | 2.4E-17 |
| rs3813565 | 15 | 76806665 | C | A | 3 | 7110 | 7124 | 0.44 | 0.39 | 0.66 | 0 | 1.24 | 3.4E-17 | 1.24 | 3.4E-17 |
| rs938682 | 15 | 76683602 | T | C | 3 | 7197 | 7178 | 0.20 | 0.23 | 0.77 | 0 | 0.78 | 1.0E-16 | 0.78 | 1.0E-16 |
| rs12910984 | 15 | 76678682 | A | G | 3 | 7190 | 7176 | 0.20 | 0.24 | 0.68 | 0 | 0.78 | 1.1E-16 | 0.78 | 1.1E-16 |
| rs8042374 | 15 | 76695087 | A | G | 3 | 7186 | 7161 | 0.20 | 0.24 | 0.64 | 0 | 0.78 | 1.1E-16 | 0.78 | 1.1E-16 |
| rs578776 | 15 | 76675455 | C | T | 3 | 7193 | 7182 | 0.23 | 0.28 | 0.27 | 22.9 | 0.79 | 1.1E-16 | 0.78 | 1.9E-09 |
| rs401681 | 5 | 1375087 | C | T | 9 | 14872 | 29433 | 0.40 | 0.44 | 0.02 | 55.0 | 0.88 | 7.2E-16 | 0.87 | 1.7E-07 |
| rs31489 | 5 | 1395714 | C | A | 9 | 14883 | 29440 | 0.37 | 0.41 | 0.01 | 59.7 | 0.88 | 1.7E-15 | 0.87 | 1.1E-06 |
| rs12441998 | 15 | 76716427 | A | G | 3 | 7194 | 7181 | 0.19 | 0.22 | 0.79 | 0 | 0.79 | 2.6E-15 | 0.79 | 2.6E-15 |
| rs12594247 | 15 | 76733688 | C | T | 9 | 14895 | 29457 | 0.14 | 0.16 | 0.94 | 0 | 0.84 | 1.8E-14 | 0.84 | 1.8E-14 |
| rs1316971 | 15 | 76717565 | G | A | 3 | 7184 | 7172 | 0.19 | 0.22 | 0.68 | 0 | 0.79 | 1.9E-14 | 0.79 | 1.9E-14 |
| rs3117582 | 6 | 31728499 | T | G | 9 | 14845 | 29389 | 0.13 | 0.10 | 0.49 | 0 | 1.22 | 2.3E-14 | 1.22 | 2.3E-14 |
| rs3130350 | 6 | 30435818 | G | T | 9 | 14853 | 29414 | 0.12 | 0.10 | 0.64 | 0 | 1.22 | 2.7E-14 | 1.22 | 2.7E-14 |
| rs8321 | 6 | 30140501 | A | C | 9 | 14871 | 29434 | 0.12 | 0.10 | 0.44 | 0 | 1.23 | 3.9E-14 | 1.23 | 3.9E-14 |
| rs558702 | 6 | 31978305 | G | A | 9 | 14888 | 29468 | 0.13 | 0.10 | 0.38 | 6.3 | 1.22 | 5.8E-14 | 1.22 | 9.9E-13 |
| rs3094054 | 6 | 30441484 | G | T | 9 | 14847 | 29341 | 0.13 | 0.11 | 0.91 | 0 | 1.22 | 7.7E-14 | 1.22 | 7.7E-14 |
| rs2736100 | 5 | 1339516 | A | C | 9 | 14870 | 29386 | 0.47 | 0.49 | 0.08 | 42.6 | 1.13 | 7.9E-14 | 1.14 | 5.0E-08 |
| rs9261290 | 6 | 30146626 | T | C | 9 | 14887 | 29409 | 0.12 | 0.10 | 0.36 | 9.0 | 1.22 | 9.6E-14 | 1.23 | 2.4E-12 |
| rs3885951 | 15 | 76612972 | A | G | 9 | 14878 | 29376 | 0.11 | 0.10 | 0.01 | 59.0 | 1.22 | 1.1E-13 | 1.22 | 2.2E-05 |
| rs3130380 | 6 | 30387109 | G | A | 9 | 14869 | 29458 | 0.12 | 0.10 | 0.67 | 0 | 1.22 | 1.2E-13 | 1.22 | 1.2E-13 |
| rs3132610 | 6 | 30652380 | A | G | 9 | 14890 | 29429 | 0.14 | 0.11 | 0.76 | 0 | 1.21 | 1.2E-13 | 1.21 | 1.2E-13 |
| rs389884 | 6 | 32048876 | A | G | 9 | 14884 | 29405 | 0.13 | 0.10 | 0.44 | 0 | 1.21 | 1.3E-13 | 1.21 | 1.3E-13 |
| rs1270942 | 6 | 32026839 | A | G | 9 | 14872 | 29420 | 0.13 | 0.10 | 0.45 | 0 | 1.21 | 1.4E-13 | 1.21 | 1.4E-13 |

Supplementary Tables

| rs number | Chr | Position (NCBI Build36) | Allele | | Number of studies | N cases | N controls | MAF | | Heterogeneity | | Fixed effect model | | Random effect model | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Reference | Effect | | | | cases | controls | $P^1$ | $I^2$ (%) | OR | P | OR | P |
| rs3131379 | 6 | 31829012 | G | A | 9 | 14871 | 29423 | 0.13 | 0.10 | 0.36 | 8.5 | 1.21 | 1.5E-13 | 1.21 | 6.8E-12 |
| rs1235162 | 6 | 29645203 | A | G | 9 | 14869 | 29408 | 0.12 | 0.09 | 0.58 | 0 | 1.22 | 2.3E-13 | 1.22 | 2.3E-13 |
| rs2187668 | 6 | 32713862 | C | T | 9 | 14863 | 29378 | 0.14 | 0.12 | 0.56 | 0 | 1.19 | 2.7E-13 | 1.19 | 2.7E-13 |
| rs7775397 | 6 | 32369230 | T | G | 9 | 14856 | 29389 | 0.12 | 0.10 | 0.25 | 22.1 | 1.21 | 6.6E-13 | 1.21 | 3.1E-09 |
| rs1794282 | 6 | 32774504 | C | T | 9 | 14889 | 29471 | 0.12 | 0.10 | 0.54 | 0 | 1.21 | 7.1E-13 | 1.21 | 7.1E-13 |
| rs9262143 | 6 | 30760760 | C | T | 9 | 14875 | 29424 | 0.13 | 0.11 | 0.88 | 0 | 1.20 | 1.0E-12 | 1.20 | 1.0E-12 |
| rs4635969 | 5 | 1361552 | G | A | 9 | 14810 | 29318 | 0.17 | 0.20 | 0.07 | 44.3 | 0.86 | 1.3E-12 | 0.86 | 4.9E-07 |
| rs2734986 | 6 | 29926547 | T | C | 9 | 14887 | 29451 | 0.19 | 0.16 | 0.30 | 15.7 | 1.17 | 1.5E-12 | 1.17 | 2.2E-10 |
| rs3130544 | 6 | 31166319 | C | A | 9 | 14883 | 29450 | 0.14 | 0.12 | 0.68 | 0 | 1.19 | 2.1E-12 | 1.19 | 2.1E-12 |
| rs8038920 | 15 | 76761600 | G | A | 9 | 14873 | 29402 | 0.33 | 0.35 | 0.54 | 0 | 0.89 | 4.0E-12 | 0.89 | 4.0E-12 |
| rs2746150 | 6 | 29550680 | C | T | 9 | 14893 | 29422 | 0.11 | 0.09 | 0.27 | 19.8 | 1.21 | 6.5E-12 | 1.22 | 3.6E-09 |
| rs1996371 | 15 | 76743861 | T | C | 5 | 5672 | 17543 | 0.43 | 0.39 | 0.62 | 0 | 1.19 | 1.1E-11 | 1.19 | 1.1E-11 |
| rs3132685 | 6 | 30053928 | G | A | 9 | 14832 | 29372 | 0.13 | 0.11 | 0.63 | 0 | 1.19 | 1.6E-11 | 1.19 | 1.6E-11 |
| rs7750641 | 6 | 31237289 | C | T | 9 | 14858 | 29385 | 0.14 | 0.12 | 0.66 | 0 | 1.18 | 2.4E-11 | 1.18 | 2.4E-11 |
| rs3132580 | 6 | 31028103 | G | A | 9 | 14897 | 29467 | 0.16 | 0.13 | 0.41 | 3.3 | 1.17 | 4.3E-11 | 1.17 | 1.4E-10 |
| rs886424 | 6 | 30889981 | C | T | 9 | 14866 | 29434 | 0.15 | 0.12 | 0.73 | 0 | 1.18 | 6.6E-11 | 1.18 | 6.6E-11 |
| rs3099844 | 6 | 31556955 | C | A | 9 | 14869 | 29439 | 0.14 | 0.12 | 0.42 | 1.8 | 1.17 | 1.0E-10 | 1.18 | 1.8E-10 |
| rs3129763 | 6 | 32698903 | G | A | 9 | 14893 | 29441 | 0.26 | 0.24 | 0.30 | 16.0 | 1.13 | 1.2E-10 | 1.13 | 1.5E-08 |
| rs1551821 | 18 | 41120468 | A | C | 6 | 7317 | 11489 | 0.11 | 0.14 | 0.00 | 85.1 | 0.81 | 1.4E-10 | 0.81 | 3.0E-02 |
| rs17487514 | 15 | 76740840 | C | T | 3 | 7186 | 7168 | 0.32 | 0.29 | 0.83 | 0.0 | 1.19 | 1.9E-10 | 1.19 | 1.9E-10 |
| rs11072793 | 15 | 76793497 | A | G | 9 | 14867 | 29406 | 0.24 | 0.26 | 0.33 | 12.6 | 0.89 | 2.3E-10 | 0.89 | 7.8E-09 |
| rs259919 | 6 | 30133482 | G | A | 9 | 14884 | 29404 | 0.34 | 0.32 | 0.59 | 0 | 1.11 | 6.2E-10 | 1.11 | 6.2E-10 |
| rs3129791 | 6 | 29062272 | G | A | 9 | 14896 | 29452 | 0.11 | 0.09 | 0.28 | 18.0 | 1.19 | 7.0E-10 | 1.19 | 7.1E-08 |
| rs680244 | 15 | 76658343 | C | T | 9 | 14876 | 29452 | 0.40 | 0.42 | 0.83 | 0 | 0.90 | 7.2E-10 | 0.90 | 7.2E-10 |
| rs3094694 | 6 | 30559883 | T | C | 9 | 14840 | 29400 | 0.20 | 0.18 | 0.33 | 12.7 | 1.14 | 9.6E-10 | 1.13 | 9.2E-08 |
| rs3131093 | 6 | 28945416 | C | T | 9 | 14892 | 29457 | 0.11 | 0.09 | 0.40 | 3.6 | 1.19 | 1.1E-09 | 1.19 | 3.1E-09 |
| rs3130893 | 6 | 29088686 | A | G | 9 | 14887 | 29393 | 0.11 | 0.09 | 0.31 | 14.9 | 1.19 | 1.5E-09 | 1.19 | 5.8E-08 |
| rs3749971 | 6 | 29450754 | G | A | 9 | 14884 | 29440 | 0.11 | 0.10 | 0.27 | 19.6 | 1.18 | 1.5E-09 | 1.19 | 1.4E-07 |
| rs2844773 | 6 | 30315474 | C | A | 9 | 14894 | 29470 | 0.17 | 0.15 | 0.87 | 0 | 1.15 | 1.5E-09 | 1.15 | 1.5E-09 |
| rs2523987 | 6 | 30187972 | A | C | 9 | 14875 | 29407 | 0.15 | 0.13 | 0.20 | 27.1 | 1.15 | 1.7E-09 | 1.16 | 5.4E-07 |
| rs4324798 | 6 | 28884096 | G | A | 9 | 14832 | 29386 | 0.10 | 0.09 | 0.28 | 18.6 | 1.19 | 1.7E-09 | 1.19 | 1.6E-07 |
| rs4539564 | 15 | 76915554 | A | G | 9 | 14860 | 29397 | 0.49 | 0.49 | 0.45 | 0 | 1.10 | 1.8E-09 | 1.10 | 1.8E-09 |
| rs6495306 | 15 | 76652948 | A | G | 9 | 14885 | 29443 | 0.40 | 0.42 | 0.86 | 0 | 0.91 | 1.8E-09 | 0.91 | 1.8E-09 |
| rs2524005 | 6 | 30007656 | G | A | 9 | 14883 | 29454 | 0.22 | 0.20 | 0.17 | 31.1 | 1.13 | 2.0E-09 | 1.14 | 1.3E-06 |

Supplementary Tables

| rs number | Chr | Position (NCBI Build36) | Reference | Effect | Number of studies | N cases | N controls | MAF cases | MAF controls | $P^1$ | $I^2$ (%) | OR | P | OR | P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Allele | | | | | | | Heterogeneity | | Fixed effect model | | Random effect model | |
| rs13194504 | 6 | 28738670 | G | A | 9 | 14894 | 29471 | 0.10 | 0.09 | 0.14 | 35.3 | 1.19 | 2.2E-09 | 1.20 | 3.6E-06 |
| rs2535238 | 6 | 29753017 | C | A | 9 | 14849 | 29388 | 0.24 | 0.22 | 0.62 | 0 | 1.12 | 2.7E-09 | 1.12 | 2.7E-09 |
| rs13211507 | 6 | 28365356 | T | C | 9 | 14894 | 29456 | 0.10 | 0.09 | 0.20 | 27.3 | 1.19 | 3.3E-09 | 1.19 | 8.2E-07 |
| rs3094073 | 6 | 30339203 | G | A | 9 | 14885 | 29438 | 0.17 | 0.15 | 0.80 | 0 | 1.14 | 3.9E-09 | 1.14 | 3.9E-09 |
| rs1233579 | 6 | 28820642 | A | G | 9 | 14838 | 29402 | 0.10 | 0.09 | 0.37 | 7.6 | 1.18 | 4.1E-09 | 1.19 | 2.7E-08 |
| rs2747457 | 6 | 29764396 | T | G | 9 | 14873 | 29432 | 0.25 | 0.23 | 0.78 | 0 | 1.12 | 5.5E-09 | 1.12 | 5.5E-09 |
| rs259940 | 6 | 30119913 | A | G | 9 | 14876 | 29439 | 0.30 | 0.28 | 0.31 | 14.9 | 1.11 | 5.9E-09 | 1.11 | 2.2E-07 |
| rs4380028 | 15 | 76898148 | C | T | 9 | 14874 | 29418 | 0.42 | 0.39 | 0.43 | 0.1 | 1.10 | 6.0E-09 | 1.10 | 6.2E-09 |
| rs3134942 | 6 | 32276749 | G | T | 9 | 14814 | 29334 | 0.15 | 0.13 | 0.32 | 13.5 | 1.15 | 6.9E-09 | 1.15 | 3.0E-07 |
| rs950776 | 15 | 76713073 | T | C | 7 | 12342 | 24124 | 0.32 | 0.34 | 0.43 | 0 | 0.90 | 7.0E-09 | 0.90 | 7.0E-09 |
| rs4975616 | 5 | 1368660 | A | G | 7 | 12352 | 24144 | 0.39 | 0.41 | 0.15 | 36.9 | 0.90 | 7.1E-09 | 0.90 | 3.2E-05 |
| rs3094127 | 6 | 30805426 | A | G | 9 | 14878 | 29437 | 0.23 | 0.20 | 0.50 | 0 | 1.12 | 8.0E-09 | 1.12 | 8.0E-09 |
| rs7762279 | 6 | 32863268 | T | C | 9 | 14855 | 29429 | 0.11 | 0.10 | 0.66 | 0 | 1.17 | 1.0E-08 | 1.17 | 1.0E-08 |
| rs2517598 | 6 | 30188253 | G | A | 9 | 14893 | 29460 | 0.17 | 0.15 | 0.43 | 1.0 | 1.13 | 1.2E-08 | 1.14 | 1.5E-08 |
| rs1245371 | 6 | 30146331 | A | G | 8 | 14045 | 18195 | 0.30 | 0.28 | 0.38 | 6.5 | 1.11 | 1.2E-08 | 1.11 | 6.0E-08 |
| rs2523989 | 6 | 30186254 | C | T | 9 | 14882 | 29414 | 0.17 | 0.14 | 0.27 | 19.8 | 1.14 | 1.2E-08 | 1.14 | 8.6E-07 |
| rs3129073 | 6 | 29723801 | A | G | 9 | 14880 | 29455 | 0.22 | 0.20 | 0.54 | 0 | 1.12 | 1.9E-08 | 1.12 | 1.9E-08 |
| rs2656069 | 15 | 76532762 | A | G | 3 | 7192 | 7180 | 0.20 | 0.22 | 0.72 | 0 | 0.84 | 2.0E-08 | 0.84 | 2.0E-08 |
| rs1062980 | 15 | 76579582 | T | C | 3 | 7185 | 7174 | 0.36 | 0.38 | 0.22 | 33.9 | 0.87 | 2.2E-08 | 0.85 | 2.5E-04 |
| rs2233956 | 6 | 31189184 | T | C | 8 | 14029 | 18188 | 0.20 | 0.18 | 0.89 | 0 | 1.13 | 2.4E-08 | 1.13 | 2.4E-08 |
| rs2508049 | 6 | 29931862 | A | G | 4 | 4653 | 6123 | 0.19 | 0.17 | 0.62 | 0 | 1.22 | 2.5E-08 | 1.22 | 2.5E-08 |
| rs10519198 | 15 | 76529809 | C | A | 9 | 14850 | 29335 | 0.42 | 0.44 | 0.67 | 0 | 0.91 | 2.6E-08 | 0.91 | 2.6E-08 |
| rs4678 | 6 | 31001920 | G | A | 8 | 13600 | 28045 | 0.22 | 0.20 | 0.12 | 39.2 | 1.12 | 2.8E-08 | 1.13 | 4.5E-05 |
| rs3117143 | 6 | 29139121 | C | A | 9 | 14768 | 29276 | 0.10 | 0.09 | 0.10 | 40.8 | 1.17 | 2.9E-08 | 1.18 | 6.9E-05 |
| rs3130564 | 6 | 31209653 | C | T | 9 | 14882 | 29445 | 0.22 | 0.19 | 0.44 | 0 | 1.12 | 3.0E-08 | 1.12 | 3.0E-08 |
| rs10484399 | 6 | 27642507 | A | G | 9 | 14850 | 29381 | 0.10 | 0.09 | 0.76 | 0 | 1.17 | 3.1E-08 | 1.17 | 3.1E-08 |
| rs2256543 | 6 | 30045812 | C | T | 9 | 14884 | 29442 | 0.45 | 0.43 | 0.05 | 48.7 | 1.09 | 3.4E-08 | 1.10 | 2.7E-04 |
| rs2517861 | 6 | 29929961 | C | T | 9 | 14869 | 29381 | 0.26 | 0.23 | 0.29 | 17.4 | 1.11 | 3.5E-08 | 1.11 | 1.8E-06 |
| rs1150735 | 6 | 30153178 | G | A | 9 | 14870 | 29406 | 0.39 | 0.37 | 0.62 | 0 | 1.09 | 3.8E-08 | 1.09 | 3.8E-08 |
| rs2844573 | 6 | 31443433 | A | C | 8 | 13773 | 16810 | 0.31 | 0.29 | 0.76 | 0 | 1.11 | 3.8E-08 | 1.11 | 3.8E-08 |
| rs2844659 | 6 | 30932511 | C | T | 9 | 14785 | 29339 | 0.22 | 0.19 | 0.35 | 10.0 | 1.12 | 3.8E-08 | 1.12 | 4.4E-07 |
| rs6904596 | 6 | 27599278 | G | A | 9 | 14867 | 29411 | 0.11 | 0.10 | 0.97 | 0 | 1.16 | 4.7E-08 | 1.16 | 4.7E-08 |

MAF - minor allele frequency; P - p value; OR - Odds Ratio; Chr - chromosome;
[1] P value for heterogeneity was assessed using Cochran Q-statistic

Supplementary Tables

**Table S3. Genome-Wide Significant Result (p<5 x 10[-8]) for the Overall Meta-Analysis adjusted for smoking status[1]**

| rs number | Chr | Position (NCBI Build36) | Allele Reference | Allele Effect | Number of studies | N cases | N controls | MAF cases | MAF controls | Heterogeneity P[2] | Heterogeneity I2 (%) | Fixed effect model OR | Fixed effect model P | Random effect model OR | Random effect model P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| rs1051730 | 15 | 76681394 | G | A | 7 | 12243 | 24073 | 0.41 | 0.35 | 0.78 | 0 | 1.32 | 1.4E-45 | 1.32 | 1.4E-45 |
| rs8034191 | 15 | 76593078 | T | C | 7 | 12253 | 24089 | 0.41 | 0.35 | 0.85 | 0 | 1.29 | 2.3E-40 | 1.29 | 2.3E-40 |
| rs12914385 | 15 | 76685778 | C | T | 3 | 7165 | 7145 | 0.46 | 0.40 | 0.10 | 56.5 | 1.31 | 2.5E-24 | 1.25 | 1.2E-04 |
| rs6495309 | 15 | 76702300 | C | T | 7 | 12257 | 24108 | 0.18 | 0.22 | 0.56 | 0 | 0.80 | 9.6E-22 | 0.80 | 9.6E-22 |
| rs2036534 | 15 | 76614003 | T | C | 7 | 12230 | 24053 | 0.19 | 0.22 | 0.59 | 0 | 0.81 | 4.7E-21 | 0.81 | 4.7E-21 |
| rs6495314 | 15 | 76747584 | A | C | 7 | 12249 | 24073 | 0.44 | 0.40 | 0.73 | 0 | 1.19 | 4.7E-21 | 1.19 | 4.7E-21 |
| rs11638372 | 15 | 76770614 | C | T | 7 | 12249 | 24076 | 0.43 | 0.39 | 0.52 | 0 | 1.19 | 8.3E-20 | 1.19 | 8.3E-20 |
| rs1394371 | 15 | 76511524 | C | T | 7 | 12252 | 24084 | 0.33 | 0.29 | 0.67 | 0 | 1.18 | 4.2E-17 | 1.18 | 4.2E-17 |
| rs578776 | 15 | 76675455 | C | T | 3 | 7172 | 7161 | 0.23 | 0.28 | 0.31 | 14.0 | 0.79 | 1.7E-15 | 0.79 | 3.2E-10 |
| rs8042374 | 15 | 76695087 | A | G | 3 | 7165 | 7140 | 0.20 | 0.24 | 0.63 | 0 | 0.78 | 1.7E-15 | 0.78 | 1.7E-15 |
| rs12910984 | 15 | 76678682 | A | G | 3 | 7169 | 7155 | 0.20 | 0.24 | 0.55 | 0 | 0.78 | 1.9E-15 | 0.78 | 1.9E-15 |
| rs938682 | 15 | 76683602 | T | C | 3 | 7176 | 7157 | 0.20 | 0.24 | 0.60 | 0 | 0.78 | 2.2E-15 | 0.78 | 2.2E-15 |
| rs3813565 | 15 | 76806665 | C | A | 3 | 7089 | 7104 | 0.44 | 0.39 | 0.15 | 47.6 | 1.23 | 8.1E-15 | 1.21 | 3.1E-04 |
| rs12441998 | 15 | 76716427 | A | G | 3 | 7173 | 7160 | 0.19 | 0.22 | 0.76 | 0 | 0.79 | 1.5E-13 | 0.79 | 1.5E-13 |
| rs13180 | 15 | 76576543 | T | C | 7 | 12241 | 24053 | 0.35 | 0.39 | 0.61 | 0 | 0.87 | 2.0E-13 | 0.87 | 2.0E-13 |
| rs1316971 | 15 | 76717565 | G | A | 3 | 7165 | 7151 | 0.19 | 0.22 | 0.65 | 0 | 0.80 | 6.7E-13 | 0.80 | 6.7E-13 |
| rs2736100 | 5 | 1339516 | A | C | 7 | 12235 | 24032 | 0.46 | 0.49 | 0.26 | 22.3 | 1.14 | 1.2E-12 | 1.15 | 1.6E-09 |
| rs3130350 | 6 | 30435818 | G | T | 7 | 12226 | 24056 | 0.11 | 0.09 | 0.47 | 0 | 1.24 | 7.8E-11 | 1.24 | 7.8E-11 |
| rs3130380 | 6 | 30387109 | G | A | 7 | 12233 | 24100 | 0.11 | 0.09 | 0.54 | 0 | 1.23 | 1.4E-10 | 1.23 | 1.4E-10 |
| rs12594247 | 15 | 76733688 | C | T | 7 | 12257 | 24096 | 0.14 | 0.16 | 0.98 | 0 | 0.85 | 1.8E-10 | 0.85 | 1.8E-10 |
| rs8321 | 6 | 30140501 | A | C | 7 | 12236 | 24074 | 0.11 | 0.09 | 0.39 | 4.1 | 1.24 | 2.1E-10 | 1.24 | 8.0E-10 |
| rs4887053 | 15 | 76499754 | C | A | 7 | 12257 | 24101 | 0.18 | 0.21 | 0.95 | 0 | 0.87 | 3.8E-10 | 0.87 | 3.8E-10 |
| rs9261290 | 6 | 30146626 | T | C | 7 | 12250 | 24049 | 0.11 | 0.09 | 0.31 | 15.6 | 1.23 | 5.2E-10 | 1.24 | 3.2E-08 |
| rs31489 | 5 | 1395714 | C | A | 7 | 12246 | 24088 | 0.37 | 0.41 | 0.02 | 58.7 | 0.89 | 5.3E-10 | 0.88 | 1.8E-04 |
| rs401681 | 5 | 1375087 | C | T | 7 | 12244 | 24074 | 0.41 | 0.44 | 0.03 | 57.3 | 0.89 | 5.4E-10 | 0.88 | 1.2E-04 |
| rs3094054 | 6 | 30441484 | G | T | 7 | 12212 | 23980 | 0.11 | 0.09 | 0.89 | 0 | 1.22 | 9.7E-10 | 1.22 | 9.7E-10 |
| rs1235162 | 6 | 29645203 | A | G | 7 | 12238 | 24047 | 0.10 | 0.08 | 0.55 | 0 | 1.23 | 1.0E-09 | 1.23 | 1.0E-09 |
| rs3117582 | 6 | 31728499 | T | G | 7 | 12226 | 24029 | 0.11 | 0.09 | 0.42 | 0.2 | 1.21 | 1.3E-09 | 1.21 | 1.4E-09 |
| rs7775397 | 6 | 32369230 | T | G | 7 | 12234 | 24033 | 0.11 | 0.09 | 0.09 | 45.0 | 1.21 | 1.7E-09 | 1.22 | 4.0E-05 |
| rs1794282 | 6 | 32774504 | C | T | 7 | 12252 | 24112 | 0.11 | 0.09 | 0.25 | 23.1 | 1.21 | 2.6E-09 | 1.21 | 1.6E-06 |
| rs3131379 | 6 | 31829012 | G | A | 7 | 12243 | 24067 | 0.11 | 0.09 | 0.35 | 10.8 | 1.20 | 3.0E-09 | 1.21 | 7.3E-08 |

Supplementary Tables

| rs number | Chr | Position (NCBI Build36) | Allele | | Number of studies | N cases | N controls | MAF | | Heterogeneity | | Fixed effect model | | Random effect model | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Reference | Effect | | | | cases | controls | P$^2$ | I2 (%) | OR | P | OR | P |
| rs2734986 | 6 | 29926547 | T | C | 7 | 12252 | 24090 | 0.17 | 0.15 | 0.61 | 0 | 1.17 | 3.4E-09 | 1.17 | 3.4E-09 |
| rs3885951 | 15 | 76612972 | A | G | 7 | 12243 | 24018 | 0.11 | 0.10 | 0.00 | 68.4 | 1.20 | 3.4E-09 | 1.16 | 0.02 |
| rs9262143 | 6 | 30760760 | C | T | 7 | 12245 | 24064 | 0.12 | 0.10 | 0.79 | 0 | 1.20 | 3.8E-09 | 1.20 | 3.8E-09 |
| rs3132610 | 6 | 30652380 | A | G | 7 | 12257 | 24070 | 0.12 | 0.10 | 0.69 | 0 | 1.20 | 3.9E-09 | 1.20 | 3.9E-09 |
| rs680244 | 15 | 76658343 | C | T | 7 | 12245 | 24092 | 0.39 | 0.42 | 0.83 | 0 | 0.89 | 4.6E-09 | 0.89 | 4.6E-09 |
| rs8038920 | 15 | 76761600 | G | A | 7 | 12241 | 24042 | 0.33 | 0.35 | 0.28 | 20.0 | 0.89 | 5.5E-09 | 0.90 | 6.9E-06 |
| rs558702 | 6 | 31978305 | G | A | 7 | 12253 | 24107 | 0.11 | 0.09 | 0.28 | 19.7 | 1.20 | 6.2E-09 | 1.21 | 7.8E-07 |
| rs17487514 | 15 | 76740840 | C | T | 3 | 7165 | 7147 | 0.32 | 0.29 | 0.97 | 0 | 1.18 | 8.0E-09 | 1.18 | 8.0E-09 |
| rs389884 | 6 | 32048876 | A | G | 7 | 12253 | 24045 | 0.11 | 0.09 | 0.33 | 12.5 | 1.20 | 8.9E-09 | 1.20 | 2.5E-07 |
| rs1270942 | 6 | 32026839 | A | G | 7 | 12247 | 24059 | 0.11 | 0.09 | 0.34 | 11.7 | 1.20 | 1.1E-08 | 1.20 | 2.3E-07 |
| rs3132685 | 6 | 30053928 | G | A | 7 | 12212 | 24012 | 0.11 | 0.10 | 0.45 | 0 | 1.20 | 1.4E-08 | 1.20 | 1.4E-08 |
| rs6495306 | 15 | 76652948 | A | G | 7 | 12249 | 24083 | 0.39 | 0.42 | 0.74 | 0 | 0.90 | 1.4E-08 | 0.90 | 1.4E-08 |
| rs1996371 | 15 | 76743861 | T | C | 5 | 5551 | 17411 | 0.43 | 0.39 | 0.89 | 0 | 1.16 | 1.7E-08 | 1.16 | 1.7E-08 |
| rs2746150 | 6 | 29550680 | C | T | 7 | 12256 | 24062 | 0.09 | 0.08 | 0.24 | 24.9 | 1.22 | 1.8E-08 | 1.22 | 5.1E-06 |
| rs3130544 | 6 | 31166319 | C | A | 7 | 12250 | 24096 | 0.12 | 0.10 | 0.43 | 0 | 1.19 | 2.2E-08 | 1.19 | 2.2E-08 |
| rs2187668 | 6 | 32713862 | C | T | 7 | 12228 | 24021 | 0.13 | 0.11 | 0.31 | 15.5 | 1.17 | 4.1E-08 | 1.18 | 1.4E-06 |

MAF - minor allele frequency; P - p value; OR - Odds Ratio; Chr - chromosome;
[1] Neither of the UK studies contributed to the analysis adjusted for smoking because this information was lacking on the controls.
[2] value for heterogeneity was assessed using Cochran Q-statistic

Supplementary Tables

**Table S4.Pairwise Linkage Disequilibrium between rs1051730 and top 15q25 SNPs** [1]

| SNP | Position (NCBI Build36) | Distance from rs1051730 | $R^2$ | D' |
|---|---|---|---|---|
| rs938682 | 76683602 | 2208 | 0.179 | 1 |
| rs12910984 | 76678682 | 2712 | 0.179 | 1 |
| rs12914385 | 76685778 | 4384 | 0.844 | 1 |
| rs578776 | 76675455 | 5939 | 0.228 | 1 |
| rs8042374 | 76695087 | 13693 | 0.179 | 1 |
| rs6495309 | 76702300 | 20906 | 0.152 | 1 |
| rs680244 | 76658343 | 23051 | 0.46 | 1 |
| rs6495306 | 76652948 | 28446 | 0.46 | 1 |
| rs950776 | 76713073 | 31679 | 0.331 | 1 |
| rs12441998 | 76716427 | 35033 | 0.134 | 1 |
| rs1316971 | 76717565 | 36171 | 0.134 | 1 |
| rs12594247 | 76733688 | 52294 | 0.102 | 1 |
| rs17487514 | 76740840 | 59446 | 0.377 | 0.748 |
| rs1996371 | 76743861 | 62467 | 0.562 | 0.803 |
| rs6495314 | 76747584 | 66190 | 0.562 | 0.803 |
| rs2036534 | 76614003 | 67391 | 0.169 | 1 |
| rs3885951 | 76612972 | 68422 | 0.185 | 1 |
| rs8038920 | 76761600 | 80206 | 0.344 | 1 |
| rs4887077 | 76765419 | 84025 | 0.507 | 0.762 |
| rs8034191 | 76593078 | 88316 | 0.934 | 1 |
| rs11638372 | 76770614 | 89220 | 0.507 | 0.762 |
| rs1062980 | 76579582 | 101812 | 0.276 | 0.788 |
| rs13180 | 76576543 | 104851 | 0.263 | 0.783 |
| rs11072793 | 76793497 | 112103 | 0.014 | 0.263 |
| rs3813565 | 76806665 | 125271 | 0.557 | 0.772 |
| rs2656069 | 76532762 | 148632 | 0.109 | 0.78 |
| rs10519198 | 76529809 | 151585 | 0.353 | 0.876 |
| rs1394371 | 76511524 | 169870 | 0.624 | 0.91 |
| rs4887053 | 76499754 | 181640 | 0.045 | 0.504 |
| rs4887053 | 76499754 | 181640 | 0.045 | 0.504 |
| rs4380028 | 76898148 | 216754 | 0.224 | 0.498 |
| rs4539564 | 76915554 | 234160 | 0.097 | 0.381 |

[1] 32 SNPs at the 15q25 locus reached genome-wide significance ($P<5 \times 10^{-8}$). The Linkage disequilibrium between variables was based on the 1000Genome project data, CEU population and retrieved using SNAP software (1897417).

Supplementary Tables

**Table S5. Association of Genetic Variants Identified in Asian Population with the Rsk of Lung Cancer in the Meta-Analysis on Populations of European Descendants**

| Locus | Genes | SNP | Asian population | | | | | proxy in CEU | Meta-analysis | | | | | | |
| | | | Study | effect allele | MAF controls | OR | P | | R2 with the proxy (D'=1) | N Cases | N Controls | Effect Allele | MAF controls | OR | P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3q28 | *TP63* | rs4488809 | Hu et al., 2011; Miki et al., 2010 | C | 0.33 - 0.47 | 1.26 (1.21–1.32) | $7.2 \times 10^{-26}$ | rs1920235 | 0.76 | 1,461 | 1,432 | T | 0.40 | 1.02 (0.92- 1.14) | 0.69 |
| | | rs10937405 | Miki et al., 2010 | T | 0.33 | 1.31 (1.22–1.42) | $7.26 \times 10^{-12}$ | rs10937405 | n/a | 7,194 | 7,179 | T | 0.42 | 0.99 (0.94-1.04) | 0.58 |
| 3q29 | *C3orf21* | rs2131877 | Yoon et al., 2010 | G | 0.199 | 1.30 (1.19–1.42) | $2.4 \times 10^{-8}$ | rs1874104 | 0.83 | 14,881 | 29,385 | T | 0.15 | 1.00 (0.96-1.05) | 0.84 |
| 13q12.12 | *MIPEP-TNFRSF19* | rs753955 | Hu et al., 2011 | G | 0.29 | 1.18 (1.13–1.24) | $1.5 \times 10^{-12}$ | rs745068 | 0.96 | 12,201 | 24,095 | T | 0.37 | 1.01 (0.97-1.04) | 0.70 |
| 22q12.2 | *MTMR3-HORMAD2-LIF* | rs17728461 | Hu et al., 2011 | G | 0.19 | 1.20 (1.14–1.27) | $1.1 \times 10^{-11}$ | rs9614158 | 0.91 | 7,197 | 7,181 | G | 0.31 | 1.07 ( 1.01-1.12) | 0.01 |
| | | rs36600 | Hu et al., 2011 | T | 0.09 | 1.29 (1.20–1.38) | $6.2 \times 10^{-13}$ | rs3788423 | 0.92 | 14,888 | 29,438 | T | 0.26 | 1.02 (0.99-1.06) | 0.25 |

**Table S6. Characteristics of imputation methods and platforms used by the participating studies**

| | MDACC | Liverpool | ICR | SLRI | IARC | deCODE | HGF | Harvard | NCI | Han Chinese |
|---|---|---|---|---|---|---|---|---|---|---|
| **Quality Control exclusion criteria** | **See Study description*** | | | | | **See study description*** | n/a | | | |
| Missing rate per person | | >5% | >5% | >10% | >5% | | | >5% | >5% | >5% |
| X chromosome heterozygosity rate | | >20% for men and < 20% for women | >20% for men and < 20% for women | >10% for men and < 20% for women | >10% for men and < 20% for women | | | >20% for men and < 20% for women | >5% for men and < 20% for women | >10% for men and < 20% for women |
| Unexpected duplicates and first degree relatives | | duplicate ID, genome-wide IBD proportion > 0.185 | duplicate ID, genome-wide IBD proportion > 0.185 | | duplicate ID, genome-wide IBD proportion > 0.20 | | | duplicate ID, genome-wide IBD proportion > 0.185 | duplicate ID | duplicate ID, genome-wide IBD proportion > 0.20 |
| Autosomal heterozygosity rate | | >6 SD of the mean | >6 SD of the mean | >6 SD of the mean | >6 SD of the mean | | | >6 SD of the mean | >34% | >6 SD of the mean |
| Non European ancestry | | Non CEU | Non CEU | <80% CEU | <70% CEU | | | Non CEU | <80% CEU | <80% CEU |
| Test of missingness by case-control status | | | | | $<1\times10^{-6}$ | | | | | $<1\times10^{-6}$ |
| Genotyped call rate | | <0.95 | <0.95 | <0.95 | <0.95 | | | <0.95 | <0.95 | <0.99 |
| MAF | | n/a | n/a | $<1\times10^{-10}$ | n/a | | | <1% | n/a | <1% |
| HWE | | $<1\times10^{-6}$ | $<1\times10^{-6}$ | $<1\times10^{-6}$ | $<1\times10^{-7}$ | | | $<1\times10^{-6}$ | $<1\times10^{-7}$ | $<1\times10^{-6}$ |
| **Imputation Methods** | mach1.0 | Impute II | Impute II | Impute II | minimac | Impute | n/a | mach1.0 | Impute II | minimac |
| **Reference Panel** | 1000Genome | HapMap 3 + 1000Genome | HapMap 3 + 1000Genome | HapMap 3 + 1000Genome | 1000Genome | HapMap 2 (r22) | n/a | 1000Genome (release 2010-06) | HapMap 3 + 1000Genome | 1000Genome(release 2010-06) |
| **Method used for association analysis** | ProbABEL | SNPTEST | SNPTEST | SNPTEST | mach2dat | SNPTEST | n/a | PLINK (--dosage function) | R (glm function) | mach2dat |

SD - standard deviation
IBD - identical by descent
* - No additional quality control before imputation was performed. The quality control before imputation was similar to the standard quality control applied for directly genotyped SNPs.

Supplementary Tables

**Table S7. Effect Allele Frequencies for Top Most Significant SNPs by Study and Case-Control Status (in cases/in controls)**

| | rs401681 | rs2736100 | rs3117582 | rs1333040 | rs10849605 | rs6495309 | rs1051730 |
|---|---|---|---|---|---|---|---|
| | *CLPTM1L* | *TERT* | *BAG6/BAT3* | *CDKN2B-AS* | *RAD52* | *CHRNA3-CHRNB4* | *CHRNA3* |
| Study | T | C | G | C | T | T | A |
| MDACC[a] | 0.41/0.44 | 0.54/0.50 | 0.12/0.11 | 0.43/0.41 | 0.48/0.49 | 0.18/0.21 | 0.40/0.33 |
| Liverpool Lung Project | 0.38/0.44 | 0.52/0.51 | 0.17/0.14 | 0.43/0.41 | 0.43/0.48 | 0.19/0.22 | 0.37/0.32 |
| ICR-GWA study[b] | 0.41/0.44 | 0.53/0.51 | 0.16/0.13 | 0.43/0.42 | 0.46/0.47 | 0.17/0.21 | 0.40/0.34 |
| SLRI[c]/Toronto | 0.43/0.44 | 0.56/0.53 | 0.11/0.09 | 0.40/0.38 | 0.51/0.51 | 0.18/0.20 | 0.40/0.34 |
| IARC[d] GWAS: | 0.39/0.43 | 0.52/0.47 | 0.12/0.09 | 0.45/0.45 | 0.45/0.49 | 0.19/0.22 | 0.41/0.34 |
| DeCODE Genetics | 0.42/0.45 | 0.53/0.50 | 0.09/0.08 | 0.52/0.50 | 0.48/0.50 | 0.18/0.21 | 0.40/0.35 |
| HGF Germany[g] | 0.41/0.42 | 0.56/0.50 | 0.11/0.09 | 0.45/0.42 | 0.48/0.46 | 0.17/0.22 | 0.42/0.38 |
| Harvard | 0.38/0.45 | 0.44/0.50 | 0.11/0.09 | 0.41/0.38 | 0.50/0.49 | 0.19/0.22 | 0.43/0.36 |
| NCI GWAS: | 0.44/0.46 | 0.54/0.52 | 0.10/0.09 | 0.44/0.42 | 0.51/0.49 | 0.19/0.23 | 0.42/0.35 |

[a]    MD Anderson Cancer Center

[b]    Institute of Cancer Research

[c]    Samuel Lunenfeld Research Institute

[d]    International Agency for Research on Cancer

[e]    Carotene and Retinol Efficacy Trial cohort.

[f]    North Trondelag Health Study 2 / Tromsø IV.

[g]    Helmholtz-Gemeinschaft Deutscher Forschungszentren Lung Cancer GWAS.

[h]    Environment And Genetics in Lung cancer Etiology study.

[i]    Alpha-Tocopherol, Beta-Carotene Cancer Prevention study.

[j]    Prostate, Lung, Colon, Ovary screening trial.

[k]    Cancer Prevention Study II nutrition cohort.

[l]    Wellcome Trust Case Control Consortium