Published in final edited form as:

Hum Genet. 2015 March ; 134(3): 333–341. doi:10.1007/s00439-014-1528-z.

Interactions between household air pollution and GWAS-identified lung cancer susceptibility markers in the Female Lung Cancer Consortium in Asia (FLCCA)

A full list of authors and affiliations appears at the end of the article.

Abstract

We previously carried out a multi-stage genome-wide association study (GWAS) on lung cancer among never smokers in the Female Lung Cancer Consortium in Asia (FLCCA) (6,609 cases, 7,457 controls) that identified novel susceptibility loci at 10q25.2, 6q22.2, and 6p21.32, and confirmed two previously identified loci at 5p15.33 and 3q28. Household air pollution (HAP) attributed to solid fuel burning for heating and cooking, is the leading cause of the overall disease burden in Southeast Asia, and is known to contain lung carcinogens. To evaluate the gene-HAP interactions associated with lung cancer in loci independent of smoking, we analyzed data from studies participating in FLCCA with fuel use information available ($n=3$; 1,731 cases; 1,349 controls). Coal use was associated with a 30% increased risk of lung cancer ($OR=1.3$, 95% $CI=1.0$ - 1.6). Among the five *a priori* SNPs identified by our GWAS, two showed a significant interaction with coal use (*HLA Class II* rs2395185, $p=0.02$; *TP63* rs4488809/rs4600802, $p=0.04$). The risk of lung cancer associated with coal exposure varied with the respective alleles for these two SNPs. Our observations provide evidence that genetic variation in *HLA Class II* and *TP63* may modify the association between HAP and lung cancer risk. The roles played in the cell-cycle and inflammation pathways by the proteins encoded by these two genes provide biological plausibility for these interactions; however, additional replication studies are needed in other non-smoking populations.

Keywords

GWAS; Asia; female; lung cancer; solid fuel use; risk factors

Introduction

Genome-wide association studies (GWAS) of lung cancer, consisting primarily of smokers of Caucasian descent, identified susceptibility variants on 5p15 and 15q25[1-6], providing insights into the underlying mechanism(s) of lung cancer susceptibility. It was unclear, however, if these genetic variations were associated with lung cancer and/or tobacco smoking[7]. Interestingly, we did not observe an association at the nicotine receptor coding

*To whom correspondence should be addressed: H. Dean Hosgood, (dean.hosgood@einstein.yu.edu), phone (718-430-2138), fax (718-430-8780), Albert Einstein College of Medicine, Department of Epidemiology and Population Health, Division of Epidemiology, 1300 Morris Park Ave., Belfer 1309, Bronx, New York 10461.

†These authors contributed equally.

‡These authors contributed equally.

region on 15q25[8, 9] in our Female Lung Cancer Consortium in Asia (FLCCA), which consists of epidemiological studies of lung cancer restricted to never smoking female lung cancer cases and never smoking female controls, suggesting that 15q25 is not associated with lung cancer independent of smoking. Further, our multi-stage GWAS of lung cancer among never smokers identified novel lung cancer susceptibility loci[9], which were not associated with lung cancer risk (i.e., $p < 10^{-8}$) in the GWAS consisting primarily of Caucasian smokers[1-6].

Lung cancer GWAS findings highlight the importance of accounting for environmental exposures, via study design or exposure assessment data, that may modify the genetic associations. We set out to evaluate the gene-environment interactions associated with lung cancer loci independent of smoking. Indoor emissions from household combustion of coal have been classified as carcinogenic to humans[10]. We pooled data on household air pollution (HAP) attributed to solid fuel burning for heating and cooking, which is the leading cause of disease in Southeast Asia[11], from three studies included in our GWAS.

Methods

Female Lung Cancer Consortium in Asia (FLCCA)

The Female Lung Cancer Consortium in Asia (FLCCA) consisting of epidemiological studies of lung cancer, which are restricted to never smoking female lung cancer cases and never smoking female controls, was used for this research. To date, FLCCA includes 14 studies from Mainland China, Hong Kong, Taiwan, Singapore, Japan, and South Korea. FLCCA is comprised of over 6,600 cases, 7,400 controls[9].

Three studies in FLCCA contributed solid fuel use data to this pooling effort (Table 1). The studies, which have all been previously described, include the Genetic Epidemiological Study of Lung Adenocarcinoma (GELAC) from Taiwan[12], the Shenyang Lung Cancer Study (SLCS)[13], and the Shanghai Women's Health Cohort Study (SWHS)[14, 15]. Briefly, the GELAC study recruited cases who were 18 years or more of age with incident primary lung cancer from six hospitals in Taiwan. Controls were cancer-free, randomly selected from the health examination clinics of the same hospitals during the same time period of case recruitment and frequency matched by age. The SLCS recruited cases with histologically confirmed lung cancers in Northeast China. Controls were selected from patients who were free of cancer history and symptom, and frequency matched to cases on age. The SWHS is a population-based cohort study of 75,221 women from Shanghai, China, aged 40-70 years. Participants for the current study were selected applying a nested case-control study design. Women with a newly diagnosed malignant neoplasm of the bronchus or lung after study recruitment were included in this study. Controls were selected among the study participants in the cohort who were cancer-free at the time of cancer diagnosis of the matched cases. For the SWHS, one control was randomly selected and matched with each case by age at baseline. After accounting for subjects with missing genotyping data and HAP data in all participating studies, we were left with an analytic data set of 1,731 cases and 1,349 controls.

Environmental Exposure Data

We utilized questionnaire data from each study to determine the type of fuel used for heating and/or cooking for each subject. The GELAC and SLCS studies provided information on the type of fuel used during cooking in their childhood home. SWHS provided information for fuel use for each subjects' most recent three residences lived. Fuel use data from each SWHS subject's oldest home was used to define the respective subject's fuel use. For all studies, subjects using any form of solid fuel, including coal, wood, and other forms of biomass, were classified as ever solid fuel users. Those not using these forms of fuel in their homes were classified as never solid fuel users. Ever solid fuel users were then refined into ever coal users if the specific type of fuel used in their home was coal. The questionnaires and interview methods have been previously reported for each of these studies[12-15]. The main effects of solid fuel use and coal use were assessed by logistic regression, adjusting for age (categorical: less than 40, 40-49, 50-59, 60-69, more than 70) and study (GELAC, SLCS, SWHS).

Genetic Data

The three FLCCA studies with HAP data were genotyped using Illumina 660W arrays at either the NCI Core Genotyping Facility (CGF) (GELAC, SWHS) or Beijing Gene Square (GS) Inc. (SLCS)[9]. The scanned intensity data from GS were collected and the genotypes were clustered and called at CGF using Illumina Genome Studio v2011.1 based on the GenTrain2 calling algorithm. Stringent quality control measures were used when building the final analytic GWAS dataset[9].

Gene-Environmental Interaction Analyses

The analyses for interaction of genotype and HAP exposure were conducted using a wald test under the Empirical Bayes estimation framework[16]. The Empirical Bayes estimator is a shrinkage estimator which corresponds to a weighted average of the standard simple logistic regression estimator and the retrospective likelihood estimator[17] under the assumption of gene-environment independence. This method has additional power relative to standard prospective logistic regression analysis of case-control data and provides superior control of type I error compared with retrospective methods including the case-only approach which are valid under the assumption of gene-environment independence. Our models included the main effects of the SNP and environmental exposure and their interaction term, as well as covariates for age (categorical) and study. To explore potential confounding by environmental tobacco smoke (ETS), we further adjusted our gene-HAP models by ETS (ever, never). We first restricted the analyses to our five *a priori* SNPs that achieved genome-wide significance level (i.e., $p < 10^{-8}$) in the GWAS: rs7086803, rs9387478, rs2395185, rs4488809(rs4600802) and rs2736100. Subsequent exploratory analyses were conducted using the full GWAS dataset.

Results

Genotype and HAP data from three studies in FLCCA, including a study conducted in Taiwan and two in mainland China, were pooled for a total of 1,731 never smoking female lung cancer cases and 1,349 never smoking female controls (Table 1). In all studies, we

observed a 20% increased risk of lung cancer associated with solid fuel use (OR = 1.2; 95% CI = 1.0-1.4) and a 30% increased risk of lung cancer associated with coal use (OR = 1.3; 95% CI = 1.0-1.6). The risk of lung cancer associated with the five *a priori* SNPs in this analytic subset is summarized in Supplemental Table 1.

The gene-environment interactions between the five *a priori* SNPs that achieved genome-wide significance level (i.e., $p = 10^{-8}$) in our GWAS and our two metrics of HAP exposure are summarized in Table 2. For all lung cancers, 1 of the 5 SNPs was associated with a solid fuel use interaction ($p = 0.05$) (Table 2). Two of the 5 SNPs were associated with an interaction between coal use and lung cancer risk (rs2395185, $p = 0.02$; rs4488809/rs4600802, $p = 0.04$). Further adjusting our gene-HAP models by ETS yielded similar results (Supplemental Table 2). The number of SNPs found to have statistically significant multiplicative interactions with coal use significantly exceeded the expected number of SNPs showing interaction [$p = 0.023$ for difference (2 of 5 SNPs vs 0.25 of 5 SNPs)]. The risk of lung cancer associated with coal exposure was found to vary with the respective rs2395185 and rs4488809/rs4600802 alleles (Figure 1). Dose-response relationships were observed between the lung cancer risk associated with coal use when stratifying by allele of these two SNPs. For both SNPs, the effect of the gene was strongest in those not exposed to coal (rs2395185: OR_{per-allele} = 1.29; 95% CI = 1.13-1.48; rs4488809/rs4600802 OR_{per-allele} = 1.34; 95% CI = 1.18-1.53) (Figure 1). Analyses using the full GWAS dataset did not yield any GWAS level significant interactions (i.e., $p = 10^{-8}$) with solid fuel or coal use (Supplemental Figure 1).

The gene-HAP results are presented in Supplemental Table 3 when restricted to only cases with lung adenocarcinoma and controls. Notably, genetic variation on chromosomes 6 (rs2395185) was associated with a coal interaction among adenocarcinomas ($p = 0.002$) (Supplemental Table 3). There was no evidence for interactions ($p > 0.05$) with rs7086803 and rs2736100 for all lung cancers or when restricting to adenocarcinomas.

Discussion

Lung cancer is the most common cancer in the world, with about 25% of cases (53% of those in women, 15% of those in men) attributable to factors other than tobacco use[18, 19]. Lung cancer in never smokers has unique genetic and etiologic risk factors when compared to lung cancer cases attributed to smoking tobacco[20, 21]. Women throughout Asia tend not to smoke, making them an ideal study population for elucidating the risk factors of never smoking lung cancer. Smoke from domestic fuel (i.e., coal, wood, biomass) used for cooking and heating has been associated with lung cancer[22-26], particularly among Asian females using coal[27, 28].

Consistent with the literature, we observed a 20% increased risk of lung cancer associated with solid fuel use and a 30% increased risk of lung cancer associated with coal use[27, 28]. Solid fuel combustion for heating and cooking increases the levels in the home of known carcinogens such as polycyclic aromatic hydrocarbons (PAHs)[29-32]. Genetic variation that influences the metabolism of HAP constituents may identify susceptible populations. Initial explorations into genetic susceptibility to lung cancer attributed to HAP evaluated the

interactions between (PAH-rich) smoky coal use and genetic variation in genes involved in metabolism and detoxification pathways of PAHs, such as glutathione S-transferase M1 (*GSTM1*). In Xuanwei, China where subjects experience some of the highest HAP exposures in the world, the *GSTM1* null genotype was associated with lung cancer risk, which was more pronounced among subjects with high levels of coal use relative to those with low coal use[33]. Beyond the early *GSTM1* observations, for which there is some evidence of replication in additional populations with HAP exposures[34], genetic variation in *AKR1C3*, *OGG1*, and cell cycle genes have been suggested to play a role in lung cancer attributed to HAP[35, 36]. Here, we report additional gene-HAP interactions based on SNPs that were identified by GWAS.

We found that genetic variation in *HLA Class II* and *TP63* may be involved in gene-HAP interactions that are associated with lung cancer risk. Interestingly, we observed a per-allele dose-response relationship between the lung cancer risk associated when stratifying by allele of these two SNPs. These relationships suggest that the major allele is associated with higher risk of lung cancer in our populations. Further, we observed that the effect of the genotype was strongest in those not exposed to coal, suggesting that these SNPs may play a greater role in the disease etiology of unexposed populations. Additional laboratory studies are needed to further determine the functionality of these SNPs and why the presence of the risk allele could lead to a protection from the adverse effects of HAP.

The roles played in human cell-cycle and inflammation pathways by the proteins encoded by these three genes provided biological plausibility for our observed interactions; however, the evidence must be weighed in concert with the strengths and limitations of our study. For example, p63, the product of the *TP63* gene, is involved in the p53 pathway. p53 is critical to proper cell-cycle regulation, and functions as a tumor suppressor in numerous cancers[37]. In addition, *TP63* genomic gains have been identified as potential indicators of pre-invasive lung lesions and early lung cancer diagnosis[38]. Proteins coded by *HLA Class II*, are both involved in the inflammation response. *HLA Class II* is involved in the regulation of lymphocytes necessary for B cell inflammatory responses. *HLA Class II* has both been shown to be involved in rheumatoid arthritis, a chronic inflammatory disease[39]. Interestingly, as a first line of defense against inhalation exposures, such as HAP, the respiratory tract releases cytokines (e.g., TNF- α , IL-1 β) in response to site-specific inflammation[40]. Therefore, our results suggest that HAP-induced lung cancer may be attributed to genetic variation in the cell-cycle and inflammatory pathways. Although we are the first to report these specific gene-HAP interactions, previous studies have suggested interactions between additional cell-cycle and inflammation genes and household coal use. Specifically, genetic variation in *PLA2G6*, *GSK3B*, *AKT1*, *EGF*, *TP53*, *PTEN*, *IL1B*, *IL8RA*, and *IL12A* were associated with lung cancer risk in a rural Chinese population with substantial coal smoke exposures[36, 41].

Our study is the largest study to date to evaluate gene-HAP interactions, and observed three novel gene-HAP interactions between exposures attributed to solid fuel use and genetic variation in *HLA Class II* and *TP63*. A major strength of our analysis is the use of never-smoking Asian female lung cancer cases and Asian female controls. Never smokers are the ideal population to evaluate environmental risk factors associated with lung cancer, since it

minimizes the influence from tobacco smoking, the leading cause of lung cancer. Further, females in Asia experience some of the highest exposures from solid fuel burning. Further research is needed, however, to identify additional populations for confirmatory replication studies and to identify the underlying mechanism(s) of how in-home coal exposures interact with *HLA Class II* and *TP63*.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

H. Dean Hosgood III^{1,2,*†}, Minsun Song^{2,†}, Chao Agnes Hsiung^{3,‡}, Zhihua Yin^{4,†}, Xiao-Ou Shu^{5,†}, Zhaoming Wang^{6,†}, Nilanjan Chatterjee², Wei Zheng⁵, Neil Caporaso², Laurie Burdette⁶, Meredith Yeager⁶, Sonja I Berndt², Maria Teresa Landi², Chien-Jen Chen⁷, Gee-Chen Chang^{8,9}, Chin-Fu Hsiao³, Ying-Huang Tsai¹⁰, Li-Hsin Chien³, Kuan-Yu Chen¹¹, Ming-Shyan Huang¹², Wu Chou Su¹³, Yuh-Min Chen¹⁴, Chung-Hsing Chen¹⁵, Tsung-Ying Yang⁹, Chih-Liang Wang¹⁶, Jen-Yu Hung¹², Chien-Chung Lin¹³, Reury-Perng Perng¹⁴, Chih-Yi Chen¹⁷, Kun-Chieh Chen⁹, Yao-Jen Li⁷, Chong-Jen Yu¹¹, Yi-Song Chen³, Ying-Hsiang Chen³, Fang-Yu Tsai¹⁵, Christopher Kim², Wei Jie Seow², Bryan A. Bassig², Wei Wu⁴, Peng Guan⁴, He Qincheng⁴, Yu-Tang Gao¹⁸, Qiuyin Cai⁵, Wong-Ho Chow¹⁹, Yong-Bing Xiang¹⁸, Dongxin Lin²⁰, Chen Wu²⁰, Yi-Long Wu²¹, Min-Ho Shin²², Yun-Chul Hong²³, Keitaro Matsuo²⁴, Kexin Chen²⁵, Maria Pik Wong²⁶, Dara Lu²⁷, Li Jin²⁷, Jiu-Cun Wang²⁷, Adeline Seow²⁸, Tangchun Wu²⁹, Hongbing Shen³⁰, Joseph F Fraumeni Jr², Pan-Chyr Yang^{31,‡}, I-Shou Chang^{15,†}, Baosen Zhou^{4,‡}, Stephen J Chanock^{2,‡}, Nathaniel Rothman^{2,‡}, and Qing Lan^{2,‡}

Affiliations

¹Department of Epidemiology & Population Health, Albert Einstein College of Medicine, Bronx, New York, USA

²Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Rockville, Maryland, USA

³Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan

⁴Department of Epidemiology, School of Public Health, China Medical University, Shenyang, China

⁵Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center and Vanderbilt-Ingram Cancer Center, Nashville, Tennessee, USA

⁶Core Genotype Facility, SAIC-Frederick, National Cancer Institute-Frederick, National Institutes of Health, Maryland, USA

⁷Genomic Research Center, Academia Sinica, Taipei, Taiwan

⁸Department of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan

⁹Division of Chest Medicine, Taichung Veterans General Hospital, Taichung, Taiwan

¹⁰Department of Respiratory Therapy, Chang Gung Memorial Hospital, Chiayi, Taiwan

¹¹Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

¹²Department of Internal Medicine, Kaohsiung Medical University Hospital, School of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan

¹³Department of Internal Medicine, National Cheng Kung University Hospital, College of Medicine, Cheng Kung University, Tainan, Taiwan

¹⁴Chest Department, Taipei Veterans General Hospital, Taipei, Taiwan

¹⁵National Institute of Cancer Research, National Health Research Institutes, Zhunan, Taiwan

¹⁶Department of Pulmonary and Critical Care, Chang Gung Memorial Hospital, Taoyuan, Taiwan

¹⁷Cancer Center, China Medical University and Hospital, Taichung, Taiwan

¹⁸Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China

¹⁹Department of Epidemiology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

²⁰Department of Etiology & Carcinogenesis and State Key Laboratory of Molecular Oncology, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

²¹Guangdong Lung Cancer Institute, Medical Research Center and Cancer Center of Guangdong General Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China

²²Department of Preventive Medicine, Chonnam National University Medical School, Gwangju, Republic of Korea

²³Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

²⁴Department of Preventive Medicine, Kyushu University Faculty of Medical Sciences, Fukuoka, Japan

²⁵Department of Epidemiology and Biostatistics, Tianjin Medical University Cancer Institute and Hospital, Tianjin, China

²⁶Department of Pathology, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong, China

²⁷Ministry of Education Key Laboratory of Contemporary Anthropology and State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai, China

²⁸Saw Swee Hock School of Public Health, National University of Singapore, Singapore

²⁹Institute of Occupational Medicine and Ministry of Education Key Lab for Environment and Health, School of Public Health, Huazhong University of Science and Technology, Wuhan, China

³⁰Ministry of Education Key Laboratory of Modern Toxicology, Jiangsu Key Laboratory of Cancer Biomarkers, Prevention and Treatment, Nanjing Medical University, Nanjing, China

³¹Department of Internal Medicine, National Taiwan University College of Medicine, Taiwan

Acknowledgments

This work was supported by the NCI intramural system. GELAC was supported in part by grants from the National Research Program for Biopharmaceuticals in Taiwan (MOHW103-TDU-PB-211-144003) and by Taiwan Bioinformatics Institute Core Facility (NSC 102-2319-B-400-001).

References

1. Landi MT, Chatterjee N, Yu K, Goldin LR, Goldstein AM, Rotunno M, Mirabello L, Jacobs K, Wheeler W, Yeager M, Bergen AW, Li Q, Consonni D, Pesatori AC, Wacholder S, Thun M, Diver R, Oken M, Virtamo J, Albanes D, Wang Z, Burdette L, Doheny KF, Pugh EW, Laurie C, Brennan P, Hung R, Gaborieau V, McKay JD, Lathrop M, McLaughlin J, Wang Y, Tsao MS, Spitz MR, Wang Y, Krokan H, Vatten L, Skorpen F, Arnesen E, Benhamou S, Bouchard C, Metspalu A, Vooder T, Nelis M, Valk K, Field JK, Chen C, Goodman G, Sulem P, Thorleifsson G, Rafnar T, Eisen T, Sauter W, Rosenberger A, Bickeboller H, Risch A, Chang-Claude J, Wichmann HE, Stefansson K, Houlston R, Amos CI, Fraumeni JF Jr, Savage SA, Bertazzi PA, Tucker MA, Chanock S, Caporaso NE. A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. *Am J Hum Genet.* 2009; 85:679–691. [PubMed: 19836008]
2. Amos CI, Wu XF, Broderick P, Gorlov IP, Gu J, Eisen T, Dong Q, Zhang Q, Gu XJ, Vijayakrishnan J, Sullivan K, Matakidou A, Wang YF, Mills G, Doheny K, Tsai YY, Chen WV, Shete S, Spitz MR, Houlston RS. Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nature Genetics.* 2008; 40:616–622. [PubMed: 18385676]
3. Hung RJ, McKay JD, 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 JK, Liloglou T, Xinarianos G, Cassidy A, McLaughlin J, Liu G, Narod S, Krokan HE, Skorpen F, Elvestad MB, Hveem K, Vatten L, Linseisen J, Clavel-Chapelon F, Vineis P, Bueno-de-Mesquita HB, Lund E, Martinez C, Bingham S, Rasmussen 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 DI, Znaor A, Healy C, Zelenika D, Boland A, Delepine M, Foglio M, Lechner D, Matsuda F, Blanche H, Gut I, Heath S, Lathrop M, Brennan P. A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature.* 2008; 452:633–637. [PubMed: 18385738]
4. Wang Y, Broderick P, Webb E, Wu X, Vijayakrishnan J, Matakidou A, Qureshi M, Dong Q, Gu X, Chen WV, Spitz MR, Eisen T, Amos CI, Houlston RS. Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat Genet.* 2008; 40:1407–1409. [PubMed: 18978787]

5. Wu C, Hu Z, Yu D, Huang L, Jin G, Liang J, Guo H, Tan W, Zhang M, Qian J, Lu D, Wu T, Lin D, Shen H. Genetic variants on chromosome 15q25 associated with lung cancer risk in Chinese populations. *Cancer Res.* 2009; 69:5065–5072. [PubMed: 19491260]
6. Truong T, Hung RJ, Amos CI, Wu X, Bickeboller H, Rosenberger A, Sauter W, Illig T, Wichmann HE, Risch A, Dienemann H, Kaaks R, Yang P, Jiang R, Wiencke JK, Wrensch M, Hansen H, Kelsey KT, Matsuo K, Tajima K, Schwartz AG, Wenzlaff A, Seow A, Ying C, Staratschek-Jox A, Nurnberg P, Stoelben E, Wolf J, Lazarus P, Muscat JE, Gallagher CJ, Zienolddiny S, Haugen A, van der Heijden HF, Kiemeny LA, Isla D, Mayordomo JI, Rafnar T, Stefansson K, Zhang ZF, Chang SC, Kim JH, Hong YC, Duell EJ, Andrew AS, Lejbkiewicz F, Rennert G, Muller H, Brenner H, Le Marchand L, Benhamou S, Bouchardy C, Teare MD, Xue X, McLaughlin J, Liu G, McKay JD, Brennan P, Spitz MR. Replication of lung cancer susceptibility loci at chromosomes 15q25, 5p15, and 6p21: a pooled analysis from the International Lung Cancer Consortium. *J Natl Cancer Inst.* 2010; 102:959–971. [PubMed: 20548021]
7. Chanock SJ, Hunter DJ. Genomics - When the smoke clears *Nature.* 2008; 452:537–538. [PubMed: 18385720]
8. Hsiung CA, Lan Q, Hong YC, Chen CJ, Hosgood HD, Chang IS, Chatterjee N, Brennan P, Wu C, Zheng W, Chang GC, Wu T, Park JY, Hsiao CF, Kim YH, Shen H, Seow A, Yeager M, Tsai YH, Kim YT, Chow WH, Guo H, Wang WC, Sung SW, Hu Z, Chen KY, Kim JH, Chen Y, Huang L, Lee KM, Lo YL, Gao YT, Kim JH, Liu L, Huang MS, Jung TH, Jin G, Caporaso N, Yu D, Kim CH, Su WC, Shu XO, Xu P, Kim IS, Chen YM, Ma H, Shen M, Cha SI, Tan W, Chang CH, Sung JS, Zhang M, Yang TY, Park KH, Yuenger J, Wang CL, Ryu JS, Xiang Y, Deng Q, Hutchinson A, Kim JS, Cai Q, Landi MT, Yu CJ, Park JY, Tucker M, Hung JY, Lin CC, Perng RP, Boffetta P, Chen CY, Chen KC, Yang SY, Hu CY, Chang CK, Fraumeni JF Jr, Chanock S, Yang PC, Rothman N, Lin D. The 5p15.33 locus is associated with risk of lung adenocarcinoma in never-smoking females in Asia. *PLoS Genet.* 2010; 6
9. Lan Q, Hsiung CA, Matsuo K, Hong YC, Seow A, Wang Z, Hosgood HD 3rd, Chen K, Wang JC, Chatterjee N, Hu W, Wong MP, Zheng W, Caporaso N, Park JY, Chen CJ, Kim YH, Kim YT, Landi MT, Shen H, Lawrence C, Burdett L, Yeager M, Yuenger J, Jacobs KB, Chang IS, Mitsudomi T, Kim HN, Chang GC, Bassig BA, Tucker M, Wei F, Yin Z, Wu C, An SJ, Qian B, Lee VH, Lu D, Liu J, Jeon HS, Hsiao CF, Sung JS, Kim JH, Gao YT, Tsai YH, Jung YJ, Guo H, Hu Z, Hutchinson A, Wang WC, Klein R, Chung CC, Oh IJ, Chen KY, Berndt SI, He X, Wu W, Chang J, Zhang XC, Huang MS, Zheng H, Wang J, Zhao X, Li Y, Choi JE, Su WC, Park KH, Sung SW, Shu XO, Chen YM, Liu L, Kang CH, Hu L, Chen CH, Pao W, Kim YC, Yang TY, Xu J, Guan P, Tan W, Su J, Wang CL, Li H, Siho AD, Zhao Z, Chen Y, Choi YY, Hung JY, Kim JS, Yoon HI, Cai Q, Lin CC, Park IK, Xu P, Dong J, Kim C, He Q, Perng RP, Kohno T, Kweon SS, et al. Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. *Nat Genet.* 2012; 44:1330–1335. [PubMed: 23143601]
10. IARC. Household Use of Solid Fuels and High-temperature Frying. Lyon: World Health Organization; 2010.
11. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, Amann M, Anderson HR, Andrews KG, Aryee M, Atkinson C, Bacchus LJ, Bahalim AN, Balakrishnan K, Balmes J, Barker-Collo S, Baxter A, Bell ML, Blore JD, Blyth F, Bonner C, Borges G, Bourne R, Boussinesq M, Brauer M, Brooks P, Bruce NG, Brunekreef B, Bryan-Hancock C, Bucello C, Buchbinder R, Bull F, Burnett RT, Byers TE, Calabria B, Carapetis J, Carnahan E, Chafe Z, Charlson F, Chen H, Chen JS, Cheng AT, Child JC, Cohen A, Colson KE, Cowie BC, Darby S, Darling S, Davis A, Degenhardt L, Dentener F, Des Jarlais DC, Devries K, Dherani M, Ding EL, Dorsey ER, Driscoll T, Edmond K, Ali SE, Engell RE, Erwin PJ, Fahimi S, Falder G, Farzadfar F, Ferrari A, Finucane MM, Flaxman S, Fowkes FG, Freedman G, Freeman MK, Gakidou E, Ghosh S, Giovannucci E, Gmel G, Graham K, Grainger R, Grant B, Gunnell D, Gutierrez HR, Hall W, Hoek HW, Hogan A, Hosgood HD 3rd, Hoy D, Hu H, Hubbell BJ, Hutchings SJ, Ibeanusi SE, Jacklyn GL, Jasrasaria R, Jonas JB, Kan H, Kanis JA, Kassebaum N, Kawakami N, Khang YH, Khatibzadeh S, Khoo JP, Kok C, Laden F, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet.* 2013; 380:2224–2260.

12. Jou YS, Lo YL, Hsiao CF, Chang GC, Tsai YH, Su WC, Chen YM, Huang MS, Chen HL, Chen CJ, Yang PC, Hsiung CA. Association of an EGFR intron 1 SNP with never-smoking female lung adenocarcinoma patients. *Lung Cancer*. 2009; 64:251–256. [PubMed: 19026460]
13. Yin Z, Su M, Li X, Li M, Ma R, He Q, Zhou B. ERCC2, ERCC1 polymorphisms and haplotypes, cooking oil fume and lung adenocarcinoma risk in Chinese non-smoking females. *J Exp Clin Cancer Res*. 2009; 28:153. [PubMed: 20003391]
14. Zhang Y, Shu XO, Gao YT, Ji BT, Yang G, Li HL, Kilfoy B, Rothman N, Zheng W, Chow WH. Family history of cancer and risk of lung cancer among nonsmoking Chinese women. *Cancer Epidemiol Biomarkers Prev*. 2007; 16:2432–2435. [PubMed: 18006933]
15. Zheng W, Chow WH, Yang G, Jin F, Rothman N, Blair A, Li HL, Wen W, Ji BT, Li Q, Shu XO, Gao YT. The Shanghai Women's Health Study: rationale, study design, and baseline characteristics. *Am J Epidemiol*. 2005; 162:1123–1131. [PubMed: 16236996]
16. Mukherjee B, Chatterjee N. Exploiting gene-environment independence for analysis of case-control studies: an empirical Bayes-type shrinkage estimator to trade-off between bias and efficiency. *Biometrics*. 2008; 64:685–694. [PubMed: 18162111]
17. Chatterjee N, Carroll RJ. Semiparametric maximum likelihood estimation exploiting gene-environment independence in case-control studies. *Biometrika*. 2005; 92:399–418.
18. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin*. 2005; 55:74–108. [PubMed: 15761078]
19. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers--a different disease. *Nat Rev Cancer*. 2007; 7:778–790. [PubMed: 17882278]
20. Subramanian J, Govindan R. Lung cancer in never smokers: a review. *J Clin Oncol*. 2007; 25:561–570. [PubMed: 17290066]
21. Lee YJ, Kim JH, Kim SK, Ha SJ, Mok TS, Mitsudomi T, Cho BC. Lung cancer in never smokers: change of a mindset in the molecular era. *Lung Cancer*. 2011; 72:9–15. [PubMed: 21272954]
22. Mumford JL, He XZ, Chapman RS, Cao SR, Harris DB, Li XM, Xian YL, Jiang WZ, Xu CW, Chuang JC, et al. Lung cancer and indoor air pollution in Xuan Wei, China. *Science*. 1987; 235:217–220. [PubMed: 3798109]
23. Hosgood HD, Chapman R, Shen M, Blair A, Chen E, Zheng T, Lee KM, He X, Lan Q. Portable stove use is associated with lower lung cancer mortality risk in lifetime smoky coal users. *British Journal of Cancer*. 2008; 99:1934–1939. [PubMed: 19034286]
24. Lan Q, Chapman RS, Schreinemachers DM, Tian L, He X. Household stove improvement and risk of lung cancer in Xuanwei, China. *J Natl Cancer Inst*. 2002; 94:826–835. [PubMed: 12048270]
25. Lan Q, He X, Shen M, Tian L, Liu LZ, Lai H, Chen W, Berndt SI, Hosgood HD, Lee KM, Zheng T, Blair A, Chapman RS. Variation in lung cancer risk by smoky coal subtype in Xuanwei, China. *Int J Cancer*. 2008; 123:2164–2169. [PubMed: 18712724]
26. Hernandez-Garduno E, Brauer M, Perez-Neria J, Vedal S. Wood smoke exposure and lung adenocarcinoma in non-smoking Mexican women. *Int J Tuberc Lung Dis*. 2004; 8:377–383. [PubMed: 15139478]
27. Hosgood HD 3rd, Wei H, Sapkota A, Choudhury I, Bruce N, Smith KR, Rothman N, Lan Q. Household coal use and lung cancer: systematic review and meta-analysis of case-control studies, with an emphasis on geographic variation. *Int J Epidemiol*. 2011; 40:719–728. [PubMed: 21278196]
28. Hosgood HD III, Boffetta P, Greenland S, Lee YCA, McLaughlin J, Seow A, Duell EJ, Andrew AS, Zaridze D, Szeszenia-Dabrowska N, Rudnai P, Lissowska J, Fabiánová E, Mates D, Bencko V, Foretova L, Janout V, Morgenstern H, Rothman N, Hung RJ, Brennan P, Lan Q. In-Home Coal and Wood Use and Lung Cancer Risk: A Pooled Analysis of the International Lung Cancer Consortium. *Environ Health Perspect*. 2010; 118:1743–1747. [PubMed: 20846923]
29. Huang S, Yang F, Zeng X, Chen J, Li R, Wen T, Li C, Wei W, Liu J, Chen L, Davis C, Xu J. Preliminary characterization of the oral microbiota of Chinese adults with and without gingivitis. *BMC Oral Health*. 2011; 11:33. [PubMed: 22152152]
30. Zhang J, Smith KR. Indoor air pollution: a global health concern. *Br Med Bull*. 2003; 68:209–225. [PubMed: 14757719]

31. Gustafson P, Barregard L, Strandberg B, Sallsten G. The impact of domestic wood burning on personal, indoor and outdoor levels of 1,3-butadiene, benzene, formaldehyde and acetaldehyde. *J Environ Monit.* 2007; 9:23–32. [PubMed: 17213939]
32. Gustafson P, Ostman C, Sallsten G. Indoor levels of polycyclic aromatic hydrocarbons in homes with or without wood burning for heating. *Environ Sci Technol.* 2008; 42:5074–5080. [PubMed: 18754350]
33. Lan Q, He X, Costa DJ, Tian L, Rothman N, Hu G, Mumford JL. Indoor coal combustion emissions, GSTM1 and GSTT1 genotypes, and lung cancer risk: a case-control study in Xuan Wei, China. *Cancer Epidemiol Biomarkers Prev.* 2000; 9:605–608. [PubMed: 10868696]
34. Hosgood HD III, Berndt SI, Lan Q. GST genotypes and lung cancer susceptibility in Asian populations with indoor air pollution exposures: a meta-analysis. *Mutat Res.* 2007; 636:134–143. [PubMed: 17428724]
35. Lan Q, Mumford JL, Shen M, Demarini DM, Bonner MR, He X, Yeager M, Welch R, Chanock S, Tian L, Chapman RS, Zheng T, Keohavong P, Caporaso N, Rothman N. Oxidative damage-related genes AKR1C3 and OGG1 modulate risks for lung cancer due to exposure to PAH-rich coal combustion emissions. *Carcinogenesis.* 2004; 25:2177–2181. [PubMed: 15284179]
36. Hosgood HD 3rd, Menashe I, Shen M, Yeager M, Yuenger J, Rajaraman P, He X, Chatterjee N, Caporaso NE, Zhu Y, Chanock SJ, Zheng T, Lan Q. Pathway-based evaluation of 380 candidate genes and lung cancer susceptibility suggests the importance of the cell cycle pathway. *Carcinogenesis.* 2008; 29:1938–1943. [PubMed: 18676680]
37. Hernandez-Boussard T, Rodriguez-Tome P, Montesano R, Hainaut P. IARC p53 mutation database: a relational database to compile and analyze p53 mutations in human tumors and cell lines. International Agency for Research on Cancer. *Hum Mutat.* 1999; 14:1–8. [PubMed: 10447253]
38. Massion PP, Zou Y, Uner H, Kiatsimkul P, Wolf HJ, Baron AE, Byers T, Jonsson S, Lam S, Hirsch FR, Miller YE, Franklin WA, Varella-Garcia M. Recurrent Genomic Gains in Preinvasive Lesions as a Biomarker of Risk for Lung Cancer. *PLoS ONE.* 2009; 4:e5611. [PubMed: 19547694]
39. van Gaalen FA, van Aken J, Huizinga TW, Schreuder GM, Breedveld FC, Zanelli E, van Venrooij WJ, Verweij CL, Toes RE, de Vries RR. Association between HLA class II genes and autoantibodies to cyclic citrullinated peptides (CCPs) influences the severity of rheumatoid arthritis. *Arthritis Rheum.* 2004; 50:2113–2121. [PubMed: 15248208]
40. Moldoveanu B, Otmishi P, Jani P, Walker J, Sarmiento X, Guardiola J, Saad M, Yu J. Inflammatory mechanisms in the lung. *J Inflamm Res.* 2009; 2:1–11. [PubMed: 22096348]
41. Lee KM, Shen M, Chapman RS, Yeager M, Welch R, He X, Zheng T, Hosgood HD, Yang D, Berndt SI, Chanock S, Lan Q. Polymorphisms in immunoregulatory genes, smoky coal exposure and lung cancer risk in Xuan Wei, China. *Carcinogenesis.* 2007; 28:1437–1441. [PubMed: 17361014]

a) rs2395185 **b) rs4488809(rs4600802)**

Figure 1. Risk of lung cancer (odds ratio and 95% confidence interval) associated with household coal use exposures stratified by the (a) rs2395185 and (b) rs4488809(rs4600802) genotypes.

Table 1

Studies from the Female Lung Cancer Consortium in Asia (FLCCA) that participated in a gene-environment interaction analysis of GWAS-identified SNPs that confer risk of never smoking lung cancer and household air pollution

	Genetic Epidemiological Study of Lung Adenocarcinoma (GELAC)				Shenyang Lung Cancer Study (SLCS)				Shanghai Women's Health Cohort Study (SWHS)				All Studies			
	Cases		Controls		Cases		Controls		Cases		Controls		Cases		Controls	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Total Subjects	1098	100%	1019	100%	549	100%	260	100%	84	100%	70	100%	1731	100%	1349	100%
Age *																
<59	504	45.9%	479	47.0%	282	51.4%	127	48.8%	40	47.6%	31	44.3%	826	47.7%	637	47.2%
>=59	594	54.1%	540	53.0%	267	48.6%	133	51.2%	44	52.4%	39	55.7%	905	52.3%	712	52.8%
Solid Fuel Use																
Ever	314	28.6%	251	24.6%	311	56.6%	130	50.0%	43	51.2%	42	60.0%	668	38.6%	423	31.4%
Never	784	71.4%	768	75.4%	238	43.4%	130	50.0%	41	48.8%	28	40.0%	1063	61.4%	926	68.6%
Coal Use																
Ever	69	6.3%	46	4.5%	58	10.6%	100	38.5%	43	51.2%	42	60.0%	365	21.1%	188	13.9%
Never	784	71.4%	768	75.4%	253	46.1%	130	50.0%	41	48.8%	28	40.0%	1063	61.4%	926	68.6%
Environmental Tobacco Smoke																
Ever	767	69.9%	615	60.4%	481	87.6%	222	85.4%	59	70.2%	53	75.7%	1307	75.5%	890	66.0%
Never	293	26.7%	382	37.5%	59	10.7%	35	13.5%	16	19.0%	11	15.7%	368	21.3%	428	31.7%
rs7086803																
GG	461	42.0%	488	47.9%	269	49.0%	143	55.0%	33	39.3%	44	62.9%	763	44.1%	675	50.0%
GA	459	41.8%	415	40.7%	239	43.5%	105	40.4%	38	45.2%	21	30.0%	736	42.5%	541	40.1%
AA	128	11.7%	66	6.5%	41	7.5%	12	4.6%	13	15.5%	5	7.1%	182	10.5%	83	6.2%
rs9387478																
AA	240	21.9%	274	26.9%	111	20.2%	58	22.3%	12	14.3%	19	27.1%	363	21.0%	351	26.0%
AC	548	49.9%	518	50.8%	281	51.2%	135	51.9%	39	46.4%	38	54.3%	868	50.1%	691	51.2%
CC	310	28.2%	227	22.3%	157	28.6%	67	25.8%	33	39.3%	13	18.6%	500	28.9%	307	22.8%
rs2395185																
GG	435	39.6%	475	46.6%	226	41.2%	102	39.2%	32	38.1%	32	45.7%	693	40.0%	609	45.1%
GT	515	46.9%	440	43.2%	258	47.0%	126	48.5%	39	46.4%	33	47.1%	812	46.9%	599	44.4%

	Genetic Epidemiological Study of Lung Adenocarcinoma (GELAC)						Shenyang Lung Cancer Study (SLCS)						Shanghai Women's Health Cohort Study (SWHS)						All Studies					
	Cases			Controls			Cases			Controls			Cases			Controls			Cases			Controls		
	N	%		N	%		N	%		N	%		N	%		N	%		N	%		N	%	
TT	148	13.5%		104	10.2%		65	11.8%		32	12.3%		13	15.5%		5	7.1%		226	13.1%		141	10.5%	
rs4488809(rs4600802)																								
TT	266	24.2%		291	28.6%		177	32.2%		84	32.3%		22	26.2%		24	34.3%		465	26.9%		399	29.6%	
TC	506	46.1%		500	49.1%		274	49.9%		130	50.0%		44	52.4%		41	58.6%		824	47.6%		671	49.7%	
CC	275	25.0%		178	17.5%		98	17.9%		46	17.7%		18	21.4%		5	7.1%		391	22.6%		229	17.0%	
rs2736100																								
TT	265	24.1%		394	38.7%		164	29.9%		93	35.8%		18	21.4%		21	30.0%		447	25.8%		508	37.7%	
TG	592	53.9%		481	47.2%		271	49.4%		123	47.3%		46	54.8%		42	60.0%		909	52.5%		646	47.9%	
GG	240	21.9%		144	14.1%		114	20.8%		44	16.9%		20	23.8%		7	10.0%		374	21.6%		195	14.5%	

* based on the median age of controls in all studies

Table 2
Five GWAS-identified SNPs that confer risk of never smoking lung cancer and gene-household air pollution interaction analyses by exposure type

SNP	Chromosome	Gene	All Lung Cancer Cases (N = 1,731)	
			Solid Fuel Interaction	Coal Interaction
			p-value*	p-value*
rs7086803	10	VTI1A	0.49	0.90
rs9387478	6	GOPC	0.05	0.35
rs2395185	6	HLA Class II	0.08	0.02
rs4488809 [†]	3	TP63	0.07	0.04
rs2736100	5	TERT,hTERT	0.90	0.82

[†] rs4488809(rs4600802); bold highlights indicate p 0.05;

* adjusted for age and study