OXFORD

Human Molecular Genetics, 2017, Vol. 26, No. 2 454–465

doi: 10.1093/hmg/ddw414
Advance Access Publication Date: 26 December 2016
Association Studies Article

ASSOCIATION STUDIES ARTICLE

Association between GWAS-identified lung adenocarcinoma susceptibility loci and *EGFR* mutations in never-smoking Asian women, and comparison with findings from Western populations

Wei Jie Seow[1,2,*,†], Keitaro Matsuo[3,†], Chao Agnes Hsiung[4,‡], Kouya Shiraishi[5,†], Minsun Song[1,6,†], Hee Nam Kim[7,†], Maria Pik Wong[8,†], Yun-Chul Hong[9,†], H. Dean Hosgood III[10,†], Zhaoming Wang[1,11,†], I-Shou Chang[12], Jiu-Cun Wang[13,14], Nilanjan Chatterjee[1], Margaret Tucker[1], Hu Wei[1], Tetsuya Mitsudomi[15], Wei Zheng[16], Jin Hee Kim[17], Baosen Zhou[18], Neil E. Caporaso[1], Demetrius Albanes[1], Min-Ho Shin[7], Lap Ping Chung[8], She-Juan An[19], Ping Wang[20], Hong Zheng[21], Yasushi Yatabe[22], Xu-Chao Zhang[19], Young Tae Kim[23], Xiao-Ou Shu[16], Young-Chul Kim[24,25], Bryan A. Bassig[1], Jiang Chang[26], James Chung Man Ho[27], Bu-Tian Ji[1], Michiaki Kubo[28], Yataro Daigo[29,30], Hidemi Ito[31], Yukihide Momozawa[28], Kyota Ashikawa[28], Yoichiro Kamatani[32], Takayuki Honda[5], Hiromi Sakamoto[33], Hideo Kunitoh[34], Koji Tsuta[35], Shun-Ichi Watanabe[36], Hiroshi Nokihara[37], Yohei Miyagi[38], Haruhiko Nakayama[39], Shingo Matsumoto[40], Masahiro Tsuboi[41], Koichi Goto[42], Zhihua Yin[18], Jianxin Shi[1], Atsushi Takahashi[32], Akiteru Goto[43], Yoshihiro Minamiya[44], Kimihiro Shimizu[45], Kazumi Tanaka[45], Tangchun Wu[46], Fusheng Wei[47], Jason Y.Y. Wong[1], Fumihiko Matsuda[48], Jian Su[19], Yeul Hong Kim[49], In-Jae Oh[24,25], Fengju Song[21], Victor Ho Fun Lee[50], Wu-Chou Su[51], Yuh-Min Chen[52,53], Gee-Chen Chang[54,55], Kuan-Yu Chen[56], Ming-Shyan Huang[57], Pan-Chyr Yang[58], Hsien-Chih Lin[4], Yong-Bing Xiang[59], Adeline Seow[2], Jae Yong Park[60], Sun-Seog Kweon[7,80], Chien-Jen Chen[61],

[†]These authors contributed equally.
[‡]These authors co-supervised the work.

Received: June 23, 2016. **Revised:** November 29, 2016. **Accepted:** December 5, 2016

Haixin Li[21], Yu-Tang Gao[59], Chen Wu[62], Biyun Qian[21], Daru Lu[13,14], Jianjun Liu[2,63,64], Hyo-Sung Jeon[65], Chin-Fu Hsiao[4], Jae Sook Sung[49], Ying-Huang Tsai[66], Yoo Jin Jung[23], Huan Guo[46], Zhibin Hu[67], Wen-Chang Wang[68], Charles C. Chung[1,69], Charles Lawrence[70], Laurie Burdett[1,69], Meredith Yeager[1,69], Kevin B. Jacobs[1,69], Amy Hutchinson[1,69], Sonja I. Berndt[1], Xingzhou He[71], Wei Wu[18], Junwen Wang[72,73], Yuqing Li[99], Jin Eun Choi[65], Kyong Hwa Park[49], Sook Whan Sung[74], Li Liu[75], Chang Hyun Kang[23], Lingmin Hu[76,77], Chung-Hsing Chen[12], Tsung-Ying Yang[55], Jun Xu[78], Peng Guan[18,79], Wen Tan[62], Chih-Liang Wang[81], Alan Dart Loon Sihoe[82], Ying Chen[2], Yi Young Choi[65], Jen-Yu Hung[57], Jun Suk Kim[83], Ho-Il Yoon[84], Qiuyin Cai[16], Chien-Chung Lin[51], In Kyu Park[23], Ping Xu[85], Jing Dong[76,77], Christopher Kim[1], Qincheng He[18], Reury-Perng Perng[86], Chih-Yi Chen[87,88], Roel Vermeulen[89], Junjie Wu[13,14], Wei-Yen Lim[90], Kun-Chieh Chen[55], John K.C. Chan[91], Minjie Chu[76,77], Yao-Jen Li[61], Jihua Li[92], Hongyan Chen[13,14], Chong-Jen Yu[58], Li Jin[13,14], Yen-Li Lo[4], Ying-Hsiang Chen[4], Joseph F. Fraumeni, Jr[1], Jie Liu[93], Taiki Yamaji[94], Yang Yang[95], Belynda Hicks[1,69], Kathleen Wyatt[1,69], Shengchao A. Li[1,69], Juncheng Dai[67], Hongxia Ma[67], Guangfu Jin[67], Bao Song[93], Zhehai Wang[93], Sensen Cheng[93], Xuelian Li[18,79], Yangwu Ren[18,79], Ping Cui[21], Motoki Iwasaki[94], Taichi Shimazu[94], Shoichiro Tsugane[94], Junjie Zhu[95], Gening Jiang[95], Ke Fei[95], Guoping Wu[47], Li-Hsin Chien[4], Hui-Ling Chen[4], Yu-Chun Su[4], Fang-Yu Tsai[12], Yi-Song Chen[4], Jinming Yu[93], Victoria L. Stevens[96], Ite A. Laird-Offringa[97], Crystal N. Marconett[97], Dongxin Lin[62,‡], Kexin Chen[21,‡], Yi-Long Wu[19,‡], Maria Teresa Landi[1,‡], Hongbing Shen[67,98,‡], Nathaniel Rothman[1,‡], Takashi Kohno[5,‡], Stephen J. Chanock[1,‡] and Qing Lan[1,‡]

[1]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA, [2]Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore, [3]Division of Molecular Medicine, Aichi Cancer Center Research Institute, Nagoya, Japan, [4]Institute of Population Health Sciences, National Health Research Institutes, Zhunan, Taiwan, [5]Division of Genome Biology, National Cancer Center Research Institute, Tokyo, Japan, [6]Department of Statistics, Sookmyung Women's University, Seoul, Republic of Korea, [7]Department of Preventive Medicine, Chonnam National University Medical School, Gwangju, Republic of Korea, [8]Department of Pathology, The University of Hong Kong, Queen Mary Hospital, Hong Kong, [9]Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea, [10]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, USA, [11]Department of Computational Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA, [12]National Institute of Cancer Research, National Health Research Institutes, Zhunan, Taiwan, [13]Ministry of Education Key Laboratory of Contemporary Anthropology, School of Life Sciences, Fudan University, Shanghai, People's Republic of China, [14]State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai, People's Republic of China,

[15]Division of Thoracic Surgery, Kinki University School of Medicine, Sayama, Japan, [16]Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center and Vanderbilt-Ingram Cancer Center, Nashville, TN, USA, [17]Department of Integrative Bioscience & Biotechnology, Sejong University, Seoul, Republic of Korea, [18]Department of Epidemiology, School of Public Health, China Medical University, Shenyang, People's Republic of China, [19]Guangdong Lung Cancer Institute, Medical Research Center and Cancer Center of Guangdong General Hospital, Guangdong Academy of Medical Sciences, Guangzhou, People's Republic of China, [20]Department of Radiotherapy, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center of Cancer, Key Laboratory of Cancer Prevention and Therapy, Tianjin, People's Republic of China, [21]Department of Epidemiology and Biostatistics, Tianjin Medical University Cancer Institute and Hospital, Tianjin, People's Republic of China, [22]Department of Pathology and Molecular Diagnostics, Aichi Cancer Center Central Hospital, Nagoya, Japan, [23]Department of Thoracic and Cardiovascular Surgery, Cancer Research Institute, Seoul National University College of Medicine, Seoul, Republic of Korea, [24]Lung and Esophageal Cancer Clinic, Chonnam National University Hwasun Hospital, Hwasun-eup, Republic of Korea, [25]Department of Internal Medicine, Chonnam National Univerisity Medical School, Gwangju, Republic of Korea, [26]Department of Etiology & Carcinogenesis, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, People's Republic of China, [27]Department of Medicine, The University of Hong Kong, Queen Mary Hospital, Pokfulam Road, Hong Kong, [28]Laboratory for Genotyping Development, Center for Integrative Medical Sciences, RIKEN, Yokohama, Japan, [29]Center for Antibody and Vaccine Therapy, Research Hospital, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, [30]Department of Medical Oncology and Cancer Center, Shiga University of Medical Science, Otsu, Japan, [31]Division of Epidemiology & Prevention, Aichi Cancer Center Research Institute, Nagoya, Japan, [32]Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan, [33]Division of Genetics, National Cancer Center Research Institute, Tokyo, Japan, [34]Department of Medical Oncology, Japanese Red Cross Medical Center, Tokyo, Japan, [35]Department of Pathology, National Cancer Center Hospital, Tokyo, Japan, [36]Division of Thoracic Surgery, National Cancer Center Hospital, Tokyo, Japan, [37]Department of Thoracic Oncology, National Cancer Center Hospital, Tokyo, Japan, [38]Molecular Pathology and Genetics Division, Kanagawa Cancer Center Research Institute, Kanagawa, Japan, [39]Department of Thoracic Surgery, Kanagawa Cancer Center, Kanagawa, Japan, [40]Division of Translational Research, Exploratory Oncology Research and Clinical Trial Center (EPOC), National Cancer Center, Chiba, Japan, [41]Department of Thoracic Surgery, National Cancer Center Hospital East, Chiba, Japan, [42]Department of Thoracic Oncology, National Cancer Center Hospital East, Japan, [43]Department of Cellular and Organ Pathology and [44]Department of Thoracic Surgery, Graduate School of Medicine, Akita University, Akita City, Japan, [45]Department of Integrative Center of General Surgery, Gunma University Hospital, Gunma, Japan, [46]Department of Occupational and Environmental Health and Ministry of Education Key Lab for Environment and Health, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, People's Republic of China, [47]China National Environmental Monitoring Center, Beijing, People's Republic of China, [48]Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, [49]Department of Internal Medicine, Division of Oncology/Hematology, College of Medicine, Korea University Anam Hospital, Seoul, Republic of Korea, [50]Department of Clinical Oncology, The University of Hong Kong, Queen Mary Hospital, Hong Kong, [51]Department of Internal Medicine, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan, [52]Department of Chest Medicine, Taipei Veterans General Hospital, Taipei, Taiwan, [53]College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan, [54]School of Medicine, Faculty of Medicine, National Yang-Ming University, Taipei, Taiwan, [55]Division of Chest Medicine, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan, [56]Division of Pulmonary Medicine, Department of Internal Medicine, National Taiwan University Hospital and College of Medicine, National Taiwan University, Taipei, Taiwan, [57]Department of Internal Medicine, Kaohsiung Medical University Hospital, School of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan, [58]Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan, [59]Department of Epidemiology, Shanghai Cancer Institute, Shanghai, People's Republic of China, [60]Lung Cancer Center, Kyungpook National University Medical Center, Daegu, Republic of Korea, [61]Genomic Research Center, Academia Sinica, Taipei, Taiwan, [62]Department 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, People's Republic of China, [63]Department of Human Genetics, Genome Institute of Singapore, Singapore, Singapore, [64]School of Life Sciences, Anhui Medical University, Hefei, People's Republic of China, [65]Cancer Research Center, Kyungpook National University Medical Center, Daegu, Republic of Korea, [66]Division of Pulmonary and Critical Care Medicine, Chiayi Chang Gung Memorial Hospital, Chiayi, Taiwan, [67]Department of Epidemiology and Biostatistics, Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Personalized Medicine, School of Public Health, Nanjing Medical University, Nanjing, People's Republic of China, [68]The Ph.D. Program for Translational Medicine, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan, [69]Cancer Genomics Research Laboratory, Leidos Biomedical Research Inc, Gaithersburg, MD, USA, [70]Westat, Rockville, MD, USA, [71]Chinese Center for Disease Control and Prevention, Beijing, People's Republic of China, [72]Department of Health Sciences Research, [73]Center for Individualized Medicine, Mayo Clinic, Scottsdale, AZ, USA, [74]Department of Thoracic and Cardiovascular Surgery, Seoul St Mary's Hospital, The Catholic University of Korea, Republic of Korea, [75]Department of Oncology, Cancer Center, Union Hospital, Huazhong University of Science and Technology, Wuhan, People's Republic of China, [76]Ministry of Education Key Laboratory of Modern Toxicology and [77]Jiangsu Key Laboratory of Cancer Biomarkers, Prevention and Treatment, Nanjing Medical University, Nanjing, People's Republic of China, [78]School of Public Health, Li Ka Shing (LKS) Faculty of Medicine, The University of Hong Kong, Hong Kong, People's Republic of China, [79]Key Laboratory of Cancer Etiology and Intervention, University of Liaoning Province, Shenyang, People's Republic of China, [80]Jeonnam Regional Cancer Center, Chonnam National University Hwasun, Hwasun Hospital, Republic of Korea, [81]Department of Pulmonary and Critical Care, Chang Gung Memorial Hospital, Taoyuan, Taiwan, [82]Department of Surgery, Li Ka Shing (LKS) Faculty of Medicine, The University of Hong Kong, Hong Kong, People's Republic of China, [83]Division of Medical Oncology, Department of Internal Medicine, College of Medicine, Korea University Guro Hospital, Seoul, Republic of Korea, [84]Department of Internal Medicine, Seoul National University Bundang Hospital, Seongnam, Republic of Korea, [85]Department of Oncology, Wuhan Iron and Steel (Group) Corporation Staff-Worker Hospital, Wuhan, People's Republic of China, [86]Chest Department, Taipei Veterans General Hospital, Taipei, Taiwan, [87]Institute of Medicine, Chung Shan Medical University, Taichung, Taiwan, [88]Division of Thoracic Surgery, Department of Surgery, Chung Shan Medical University Hospital, Taichung, Taiwan, [89]Division of Environmental Epidemiology, Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, The Netherlands, [90]Agency for Integrated Care, Singapore, [91]Department of Pathology, Queen Elizabeth Hospital, Hong Kong, People's Republic of China, [92]Qujing Center for Diseases Control and Prevention, Qujing, People's Republic of China, [93]Department of Oncology, Shandong Cancer Hospital and Institute, Shandong Academy of Medical Sciences, Jinan, People's Republic of China, [94]Epidemiology and Prevention Group, Center for Public Health Sciences, National Cancer Center, Tokyo, Japan, [95]Shanghai Pulmonary Hospital, Shanghai, People's Republic of China, [96]Laboratory Services, American Cancer Society, Atlanta, GA, USA, [97]Department of Surgery, Department of Biochemistry and Molecular Medicine, Norris Comprehensive Cancer Center, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA, [98]Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing, People's Republic of China and [99]Cancer Prevention Institute of California, Fremont, CA, USA

*To whom correspondence should be addressed at: Tahir Foundation Building, 12 Science Drive 2, #09-01H, Singapore 117549. Tel: +65 66011243; Fax: +65 67791489; Email: ephswj@nus.edu.sg

Abstract

To evaluate associations by *EGFR* mutation status for lung adenocarcinoma risk among never-smoking Asian women, we conducted a meta-analysis of 11 loci previously identified in genome-wide association studies (GWAS). Genotyping in an additional 10,780 never-smoking cases and 10,938 never-smoking controls from Asia confirmed associations with eight known single nucleotide polymorphisms (SNPs). Two new signals were observed at genome-wide significance ($P < 5 \times 10^{-8}$), namely, rs7216064 (17q24.3, *BPTF*), for overall lung adenocarcinoma risk, and rs3817963 (6p21.3, *BTNL2*) which is specific to cases with *EGFR* mutations. In further sub-analyses by *EGFR* status, rs9387478 (*ROS1/DCBLD1*) and rs2179920 (*HLA-DPB1*) showed stronger estimated associations in *EGFR*-positive compared to *EGFR*-negative cases. Comparison of the overall associations with published results in Western populations revealed that the majority of these findings were distinct, underscoring the importance

of distinct contributing factors for smoking and non-smoking lung cancer. Our results extend the catalogue of regions associated with lung adenocarcinoma in non-smoking Asian women and highlight the importance of how the germline could inform risk for specific tumour mutation patterns, which could have important translational implications.

Introduction

Lung cancer is a prominent global health burden that accounts for approximately 1.5 million annual deaths worldwide (1). It is also the leading cause of cancer mortality among women in China, accounting for 21.3% of all cancer deaths in 2010, even though the majority of Asian women do not smoke (2). Although tobacco smoking is a major risk factor for lung cancer, approximately 25% of all lung cancer cases occur in never-smokers (3). We previously conducted a multi-stage genome-wide association study (GWAS) of lung cancer among never-smoking women in the Female Lung Cancer Consortium in Asia (FLCCA) and identified eight lung cancer susceptibility loci on chromosomes 3q28, 5p15.33, 6p21.1, 6p21.32, 6q22.2, 9p21.3, 10q25.2 and 12q13.13 (4,5). Most of these loci were distinct from those identified in smokers of European-ancestry; suggesting that the genetic susceptibility and the aetiology of lung cancer could differ between groups of distinct ancestral origin but more importantly by smoking status (6–11).

Epidermal growth factor receptor (EGFR) is a transmembrane protein important for the regulation of cellular proliferation and apoptosis (12). Mutations in the *EGFR* gene are a defining hallmark of lung adenocarcinoma, which commonly occur in exons 18-21 (tyrosine kinase encoding region). *EGFR* mutation rates in lung cancer tumours are generally higher in Asian compared to Western populations, non-smokers compared to smokers, women compared to men, and the adenocarcinoma subtype compared to other subtypes (13,14). Lung adenocarcinoma patients harbouring different *EGFR* mutations have differential responses to tyrosine kinase inhibitor treatment, which is increasingly used as a targeted therapy (15,16).

We conducted a meta-analysis with further follow-up genotyping in never-smoking Asian women (11,725 lung adenocarcinoma cases and 14,490 controls) to investigate the possible association between *EGFR* mutation status and common genetic susceptibility alleles as well as the identification of new risk loci; 11 SNPs were selected for the analysis. Eight SNPs, corresponding to known susceptibility alleles were significantly associated with lung adenocarcinoma at a genome-wide significance level ($P < 5 \times 10^{-8}$) among never-smoking Asian women (rs2736100 (*TERT*), rs4488809 (*TP63*), rs7086803 (*VTI1A*), rs11610143 (12q13.13), rs72658409 (9p21.3), rs7741164 (*FOXP4*), rs9387478 (*ROS1/DCBLD1*), and rs2395185 (HLA class II)); however, these associations have not been previously evaluated by *EGFR* mutation status (4,5). The additional SNPs (N = 3) were identified from two previous Japanese GWAS of lung adenocarcinoma (rs3817963 (*BTNL2*), rs7216064 (*BPTF*)) or in tumour subgroups (i.e., carrying *EGFR* mutations) (rs2179920 (*HLA-DPB1*)) (17,7). However, associations with these SNPs have yet to be found in never-smoking women. Because the *EGFR* mutation status affects targeted therapy, and nearly 50% of lung adenocarcinoma in the Asian population have *EGFR* mutations (13,17), we evaluated the associations between the 11 SNPs and lung adenocarcinoma risk among never-smoking Asian women, differentiated by *EGFR* mutation status. Further, we compared the findings with results in Western populations.

Results

We conducted a fixed effects meta-analysis of case-control studies that included a total of 11,725 never-smoking female lung adenocarcinoma cases and 14,490 never-smoking controls from the FLCCA, the Nanjing GWAS study and the Japanese Lung Cancer Collaborative Study (JLCCS) (Supplementary Material, Tables S1–S3). There were 3,576 lung adenocarcinoma cases with *EGFR* data in the FLCCA and JLCCS study.

Case–control comparisons

In the meta-analysis, we observed that the SNP marker, rs7216064 (*BPTF*, odds ratio (OR) = 0.86, $P = 6.19 \times 10^{-9}$) achieved genome-wide significance for risk for lung adenocarcinoma among never-smoking Asian women. In addition, we confirmed eight known genome-wide significant risk loci at the genome-wide significance level in among never-smoking Asian women including: rs2736100 (*TERT*), rs4488809 (*TP63*), rs7086803 (*VTI1A*), rs11610143 (12q13.13), rs72658409 (9p21.3), rs7741164 (*FOXP4*), rs9387478 (*ROS1/DCBLD1*), and rs2395185 (HLA class II) (Table 1, Supplementary Material, Table S4). For the two SNPs (rs3817963 in *BTNL2* and rs2179920 in *HLA-DPB1*) that showed significant heterogeneity, the random effect method yielded similar results (data not shown).

We then restricted the case-control analysis to include only lung adenocarcinoma cases with tumours that have *EGFR* mutations in either exon 19 or 21 (*EGFR*-positive) and identified another genome-wide significant locus at 6p21.3 (rs3817963 (*BTNL2*), OR = 1.30, $P = 4.67 \times 10^{-8}$) (Table 2, Supplementary Material, Table S5). This finding is in contrast to the comparison between lung adenocarcinoma cases with tumours without *EGFR* mutations (*EGFR*-negative), for which the association was weaker (*BTNL2*, OR = 1.18, $P = 1.52 \times 10^{-3}$), thus suggesting the contribution of the germline variants in *EGFR* mutation-positive lung adenocarcinoma. In the *EGFR*-specific analyses, there was a suggestion that the estimated OR may be higher in the subset with *EGFR* mutations in the tyrosine kinase region for six of 11 SNPs; the remaining five did not indicate a difference by *EGFR* status.

Case–case comparisons

In the meta-analysis of case-case comparisons, we evaluated the associations between each of the 11 SNPs and the occurrence of *EGFR* mutation in lung adenocarcinoma tissues. We found a statistically significant association between two SNPs, rs2179920 (*HLA-DPB1*, OR = 1.20, $P = 0.0079$, false discovery rate (FDR) = 0.087) and rs9387478 (*ROS1/DCBLD1*, OR = 0.89, $P = 0.021$, FDR = 0.12), and risk of *EGFR*-positive lung adenocarcinoma, compared to *EGFR*-negative lung adenocarcinoma patients (Table 2, Supplementary Material, Tables S6 and S7). In addition, stratification by the two major racial/ethnic groups (self-reported Chinese and Japanese) yielded similar associations for both rs2179920 (OR (95% CI) = 1.28 (0.82–1.98) in Chinese; OR (95% CI) = 1.19 (1.04–1.38) in Japanese, P for heterogeneity = 0.76) and rs9387478 (OR (95% CI) = 0.77 (0.59–1.02) in Chinese; OR (95% CI) = 0.90 (0.81–1.01) in Japanese, P for heterogeneity = 0.30)

Table 1. Association of GWAS identified SNPs and lung adenocarcinoma risk among never-smoking Asian women[a]

SNP	Nearest gene(s)	Chr	Major/minor allele	MAF	No. of subjects				
				Ca/Co	Ca	Co	OR (95% CI)	P[b]	P_{het}
rs4488809[4]	*TP63*	3q28	T/C	0.47/0.53	7,448	7,007	**0.80 (0.76–0.85)**	$\mathbf{4.30 \times 10^{-17}}$	8.65×10^{-1}
rs2736100[4]	*TERT*	5p15.33	A/C	0.49/0.46	7,505	7,070	**1.43 (1.36–1.50)**	$\mathbf{6.12 \times 10^{-43}}$	8.67×10^{-1}
rs7741164[5]	*FOXP4*	6p21.1	G/A	0.37/0.33	10,531	10,648	**1.17 (1.12–1.22)**	$\mathbf{3.96 \times 10^{-13}}$	8.82×10^{-1}
rs3817963[36]	*BTNL2*	6p21.3	T/C	0.33/0.29	7,255	6,745	1.16 (1.10 – 1.22)	1.63×10^{-7}	3.57×10^{-2}
rs2179920 (rs115859925)[31]	*HLA-DPB1*	6p21.32	C/T	0.14/0.13	7,457	7,020	1.17 (1.09 – 1.26)	1.69×10^{-5}	1.37×10^{-2}
rs2395185 (rs28366298)[4]	HLA class II	6p21.32	G/T	0.41/0.38	7,757	9,637	**1.16 (1.10–1.22)**	$\mathbf{2.04 \times 10^{-9}}$	8.04×10^{-1}
rs9387478[4]	*ROS1/DCBLD1*	6q22.2	C/A	0.47/0.50	8,022	9,970	**0.86 (0.82–0.90)**	$\mathbf{5.25 \times 10^{-11}}$	7.91×10^{-1}
rs72658409[5]		9p21.3	C/T	0.05/0.07	10,780	10,938	**0.76 (0.70–0.83)**	$\mathbf{2.37 \times 10^{-10}}$	8.47×10^{-1}
rs7086803[4]	*VTI1A*	10q25.2	G/A	0.30/0.26	7,964	9,914	**1.25 (1.19–1.32)**	$\mathbf{9.22 \times 10^{-17}}$	7.73×10^{-1}
rs11610143[5]		12q13.13	C/G	0.30/0.33	10,267	10,634	**0.85 (0.81–0.89)**	$\mathbf{3.55 \times 10^{-13}}$	2.48×10^{-1}
rs7216064[36]	*BPTF*	17q24.3	A/G	0.40/0.42	7,720	8,630	**0.86 (0.82–0.90)**	$\mathbf{6.19 \times 10^{-9}}$	1.84×10^{-1}

Chr, chromosome; Ca, cases; Co, controls; MAF, minor allele frequency; OR, odds ratio; CI, confidence interval; P_{het}, P-value for heterogeneity.

[a]We used all previously reported GWAS and Taqman data from Female Lung Cancer Consortium in Asia (FLCCA), one independent GWAS from Nanjing (4,5) and Taqman data from the Japanese Lung Cancer Collaborative Study (JLCCS). We excluded FLCCA GWAS and Taqman data from overlapping subjects with the JLCCS.

[b]Adjusted for age (<40, 40–49, 50–59, 60–69, ≥70), racial/ethnic groups (Chinese, Japanese, Korean) and significant eigenvectors in the FLCCA; adjusted for age (<40, 41–50, 51–60, 61–70, >70) in the JLCCS; adjusted for age (<40, 40–49, 50–59, 60–69, ≥70) in the Nanjing study; $P < 5 \times 10^{-8}$ are considered genome-wide significant.

Table 2. Association of GWAS identified SNPs and risk for lung adenocarcinoma by *EGFR* mutation among never-smoking Asian women

SNP	Nearest gene(s)	Chr	Major/minor allele	*EGFR* status	MAF	No. of subjects			Case-control	Case-Case
					Ca/Co	Ca	Co	OR (95% CI)[a]	P[a]	P[b]
rs4488809	*TP63*	3q28	T/C	+	0.47/0.53	1,949	6,687	0.82 (0.75 – 0.89)	8.46×10^{-6}	
				–	0.47/0.53	1,324	6,687	0.80 (0.72 – 0.88)	3.42×10^{-6}	9.99×10^{-1}
rs2736100	*TERT*	5p15.33	A/C	+	0.47/0.39	1,947	6,750	**1.49 (1.36–1.63)**	$\mathbf{1.64 \times 10^{-17}}$	
				–	0.46/0.39	1,325	6,750	**1.40 (1.27–1.54)**	$\mathbf{4.94 \times 10^{-12}}$	1.87×10^{-1}
rs7741164	*FOXP4*	6p21.1	G/A	+	0.33/0.30	1,908	6,239	1.19 (1.08 – 1.31)	3.19×10^{-4}	
				–	0.33/0.30	1,287	6,239	1.20 (1.08 – 1.33)	6.31×10^{-4}	5.50×10^{-1}
rs3817963	*BTNL2*	6p21.3	T/C	+	0.35/0.29	1,945	6,749	**1.30 (1.18–1.43)**	$\mathbf{4.67 \times 10^{-8}}$	
				–	0.33/0.29	1,326	6,749	1.18 (1.06 – 1.31)	1.52×10^{-3}	1.73×10^{-1}
rs2179920	*HLA-DPB1*	6p21.32	C/T	+	0.18/0.14	1,947	6,700	1.29 (1.15 – 1.45)	1.42×10^{-5}	
				–	0.16/0.14	1,325	6,700	1.08 (0.95 – 1.24)	2.33×10^{-1}	7.93×10^{-3}
rs2395185	HLA class II	6p21.32	G/T	+	0.42/0.37	1,952	6,744	1.23 (1.12 – 1.35)	1.24×10^{-5}	
				–	0.39/0.37	1,326	6,744	1.09 (0.99 – 1.20)	8.25×10^{-2}	8.26×10^{-2}
rs9387478	*ROS1/DCBLD1*	6q22.2	C/A	+	0.44/0.49	1,951	6,745	0.84 (0.77 – 0.92)	9.36×10^{-5}	
				–	0.47/0.49	1,326	6,745	0.92 (0.84 – 1.01)	9.11×10^{-2}	2.14×10^{-2}
rs72658409		9p21.3	C/T	+	0.05/0.06	1,938	6,693	0.77 (0.63 – 0.95)	1.36×10^{-2}	
				–	0.04/0.06	1,320	6,693	0.72 (0.57 – 0.91)	4.94×10^{-3}	7.73×10^{-1}
rs7086803	*VTI1A*	10q25.2	G/A	+	0.29/0.24	1,952	6,693	1.29 (1.16 – 1.42)	8.07×10^{-7}	
				–	0.27/0.24	1,325	6,693	1.22 (1.09 – 1.36)	3.91×10^{-4}	5.49×10^{-1}
rs11610143		12q13.13	C/G	+	0.30/0.34	1,952	6,694	0.78 (0.71 – 0.86)	6.05×10^{-7}	
				–	0.30/0.34	1,324	6,694	0.83 (0.75 – 0.92)	2.86×10^{-4}	1.26×10^{-1}
rs7216064	*BPTF*	17q24.3	A/G	+	0.27/0.33	1,946	6,661	0.79 (0.72 – 0.87)	1.46×10^{-6}	
				–	0.28/0.33	1,322	6,661	0.82 (0.74 – 0.91)	2.76×10^{-4}	7.62×10^{-1}

Chr, chromosome; MAF, minor allele frequency; Ca, cases; Co, controls; OR, odds ratio; CI, confidence interval.

[a]Case–control analyses were adjusted for age (<40, 40–49, 50–59, 60–69, ≥70), racial/ethnic groups (Chinese, Japanese, Korean) and significant eigenvectors in the Female Lung Cancer Consortium in Asia (FLCCA); adjusted for age (<40, 41–50, 51–60, 61–70, >70) in the Japanese Lung Cancer Collaborative Study (JLCCS); $P < 5 \times 10^{-8}$ are considered genome-wide significant.

[b]Case–case analyses were adjusted for age (<40, 40–49, 50–59, 60–69, ≥70), racial/ethnic groups (Chinese, Japanese, Korean) and significant eigenvectors in the FLCCA; adjusted for age (<40, 41–50, 51–60, 61–70, >70) in the JLCCS.

(Supplementary Material, Table S8). We excluded seven study centres which did not provide *EGFR* information and the results were similar (data not shown).

Association of SNPs identified from studies of non-smoking Asian women in Western never-smokers and smokers

Of the 10 SNPs observed to be associated with lung adenocarcinoma risk never-smoking Asian women, three SNP markers, in *TERT*, *TP63* and 9p21.3 were significantly associated with lung adenocarcinoma risk in a small pooled study of Western never smokers ($P < 0.05$) (Supplementary Material, Table S9, 219 cases and 1,379 controls). The remaining seven SNPs were not found to be statistically significant in Western never-smokers. Among Western smokers, only two out of 10 SNPs (i.e., *TERT* and *TP63*) were significantly associated with lung adenocarcinoma risk (Supplementary Material, Table S10, 1,612 cases and 4,336 controls).

Association of SNPs identified from studies in Western populations among non-smoking Asian women

We analysed 15 SNPs identified from an extensive set of lung cancer GWAS among Western populations that showed significant associations for overall lung cancer or specific histologies (i.e., adenocarcinoma or squamous cell carcinoma) conducted to date (6,7,18–24) in our pooled GWAS dataset of never-smoking Asian women population (Supplementary Material, Table S11) and identified only two overlapping loci, marked by SNPs rs2736100 (*TERT*) and rs4488809 (*TP63*), which achieved genome-wide significance in our dataset of up to 5,512 cases and 6,277 controls.

Discussion

In our meta-analysis, two new SNPs (rs7216064, *BPTF* and rs3817963, *BTNL2*) achieved genome wide significance for lung adenocarcinoma risk in never-smoking Asian women. Additionally, we found that two loci at 6q22.2 (rs9387478, *ROS1/DCBLD1*) and 6p21.32 (rs2179920, *HLA-DPB1*) were more strongly associated with risk in *EGFR*-positive cases compared to *EGFR*-negative lung adenocarcinoma. We also confirmed eight previous GWAS signals (rs11610143 (12q13.13), rs2736100 (*TERT*), rs4488809 (*TP63*), rs2395185 (HLA class II), rs7086803 (*VTI1A*), rs72658409 (9p21.3), rs7741164 (*FOXP4*), and rs9387478 (*ROS1/DCBLD1*)) and risk of lung adenocarcinoma among never-smoking Asian women. Overall, we observed genome-wide significant associations for 10 of the 11 SNPs tested; the one remaining SNP showed a borderline association, which might have been due to the reduced sample size for this marker.

One of the new loci at 17q24.3 (rs7216064, *BPTF*) maps to a plausible candidate gene, the bromodomain PHD finger transcription factor (*BPTF*), which is the largest subunit of a nucleosome-remodeling factor that regulates transcription and the development of eukaryotic cells. A recent study showed *BPTF* to be highly expressed in non-small cell lung cancer (NSCLC) cell lines and tumour tissues compared to normal tissues and that its overexpression predicted a poor prognosis in lung adenocarcinoma patients (25). The second new locus associated with *EGFR* mutation status maps to the markers in the vicinity of the butyrophilin-like 2 (*BTNL2*) gene in MHC class II in the HLA region of chromosome 6, which is a negative regulator of T-cell proliferation and has already been shown to be associated with several immune-associated diseases (26). The discovery of a new HLA signal highlights the importance of immune regulation in adenocarcinogenesis.

The SNP marker, rs9387478 resides in the vicinity of several plausible candidate genes, ROS proto-oncogene receptor tyrosine kinase (*ROS1*) and discoidin, CUB and LCCL domain containing 1 (*DCBLD1*) genes; its association with lung adenocarcinoma in never-smoking Asian women suggests regulation of one or more of these genes could be critical in primary lung adenocarcinogenesis (4). Notably, this is the first study to report differential associations by *EGFR* mutation status for a genetic variant in the *ROS1/DCBLD1* gene. Similar to *EGFR*, the *ROS1* gene codes for a protein with tyrosine kinase activity. *ROS1* rearrangements are defined as a unique genetic subtype of lung adenocarcinoma that are more commonly found in never-smokers and younger individuals (27). Further, *in vitro* evidence in NSCLC tumour cells indicates that the *EGFR* pathway is activated as a mechanism of resistance to *ROS1* inhibition (28), suggesting a potential interaction between the genes.

A genetic variant in *HLA-DPB1* was also significantly associated with *EGFR*-positive lung adenocarcinoma. Class II major histocompatibility complex, DP beta 1 (*HLA-DPB1*) encodes a human MHC Class II lymphocyte antigen β chain and plays an important role in regulating the immune system. Interestingly, a previous study identified *HLA-DPB1* as part of a 39-gene signature that was differentially expressed among the lung adenocarcinoma patients and showed lower expression of *HLA-DPB1* to be associated with poor survival (29). Our study demonstrates the association of this region in never-smoking Asian women, consistent with a concurrent report (30). Our findings suggest a genetic predisposition to the acquisition of *EGFR* mutations among lung adenocarcinoma cases but not among lung cancer cases compared to controls, which could inform our understanding of how specific germline variants influence the acquisition of specific mutational patterns in lung adenocarcinoma. Given the high prevalence of *EGFR* mutations in lung adenocarcinomas of Asian populations, and that the efficacy of targeted therapies is influenced by *EGFR* status, findings from this study may help guide clinical decision making.

Except for the *TERT* and *TP63* gene regions, which were drawn from the ten genome-wide significant signals in our population, there was no evidence that other signals were identified in Western studies for all or specific lung cancer histologies to date (6,7,18–24); notably we did not see a signal at the 15q25 locus, which is associated with smoking and nicotine dependence, and would not necessarily be expected to show the association in our non-smoking population[4]. Overall, this suggests striking differences between Asian and Western populations that could be partially due to differences in the underlying genetic architecture of lung adenocarcinoma by ancestral history, but more likely is due to smoking status, or other distinct environmental exposures such as stove ventilation (31), household coal use (32), diet (33) and cooking fumes (34). Shiraishi et al. (35) compared a number of these associations between men and women, as well as between smokers and non-smokers in the Japanese population, and found no differences in the association for some SNPs and equivocal results for others. Further studies are needed to clarify the differences and provide new opportunities to understand the interaction of environmental risk factors with underlying genetic susceptibility to a common cancer, such as lung adenocarcinoma, which is a major global health problem.

In summary, we identified two new regions, namely SNPs, rs7216064 (*BPTF*) and rs3817963 (*BTNL2*), that were associated with risk for lung adenocarcinoma in never-smoking Asian women. Furthermore, we confirmed the strength of eight known loci. Together, we now have observed associations for ten loci in relation to lung cancer risk in non-smoking Asian women. In our analyses, we also observed two loci, namely, rs9387478 (*ROS1/DCBLD1*) and rs2179920 (*HLA-DPB1*), that were associated with lung adenocarcinoma cases with *EGFR*-positive mutations. The identification of such genetic variants establishes a foundation for investigating how and in what way the germline variants can inform our understanding of somatic patterns, particularly here, a mutational spectrum that is amenable to new targeted therapies. In turn, this could have important public health implications with respect to risk stratification, screening and treatment for lung cancer among never-smoking women in Asia. Future studies are warranted to expand on these findings and evaluate the contribution of gene-environment interactions on the acquisition of *EGFR* mutations (36).

Materials and Methods

Study population

Female lung cancer consortium in Asia (FLCCA) and Nanjing GWAS

We utilized all GWAS and Taqman data from the FLCCA and one other independent GWAS from Nanjing for this study. FLCCA consists of epidemiological studies of lung cancer, which are restricted to never-smoking female lung cancer cases and never-smoking female controls (Supplementary Material, Table S1). FLCCA included studies from Mainland China, Hong Kong, Taiwan, Singapore, Japan, and South Korea, and its efforts have been previously described (4,5). Details on the Nanjing GWAS were described in a previous publication (37).

In addition to using all previously reported data (4,5) in this pooling effort, 11 study centres in the FLCCA contributed *EGFR* mutation status data (Supplementary Material, Table S2). These 11 study centres include the Chinese Academy of Medical Sciences Cancer Hospital Study (CAMSCH), the Guangdong Study (GDS), the Genetic Epidemiological Study of Lung Adenocarcinoma (GELAC), the Hong Kong Study (HKS), the Hospital-based Epidemiologic Research Program at Aichi Cancer Center (HERPACC), the Japan National Cancer Center (NCC), the South Korea Multi-Center Lung Cancer Study (SKLCS- Korea University, Kyungpook University, Seoul National University Department of Preventive Medicine (SNUPM)), the Chonnam National University Lung Cancer Study (CNUCLS), and the Tianjin Lung Cancer Study (TLCS), which have all been previously described (4,10,38–44). All studies were case–control by design. All lung cancer cases were histologically confirmed. Each study was approved by their local institutional review board and all study participants provided informed consent prior to participation.

Japanese lung cancer collaborative study (JLCCS)

Details of the JLCCS study population have been described elsewhere (30) and their characteristics are shown in the Supplementary Material, Table S3. Briefly, JLCCS consists of lung adenocarcinoma cases and cancer-free, healthy controls from the BioBank Japan project (45), National Cancer Center Hospital (NCCH), National Cancer Center Hospital East (NCCHE), Kanagawa Cancer Center, Akita University Hospital and Gunma University Hospital. We excluded FLCCA GWAS and Taqman data from Japanese study centers consisting of overlapping subjects with the JLCCS. This study was approved by the ethics committees of each participating institution and all the participants provided written informed consent.

Genetic data

FLCCA and Nanjing GWAS

Genotyping and imputation for the FLCCA and Nanjing GWAS studies were previously described (4,5,37,46). Briefly, samples were genotyped using the Illumina 660W SNP microarray, Illumina 370K SNP microarray, Illumina 610Q SNP microarray, AffymetrixGenome-WideSNPArray6.0 and TaqMan custom genotyping assay (only for FLCCA samples) (Applied Biosystems, CA, USA). Genotype imputation was conducted by using IMPUTE2 software version 2.2.2 and the 1,000 Genomes Project Phase 1 version 3 data as the reference panel (5).

JLCCS

The 11 SNPs were genotyped using the TaqMan method or multiplex PCR-based Invader assay (Third Wave Technologies), according to the manufacturer's protocol as previously described (30). Genotyping data on five of the 11 SNPs (*HLA-DPB1*, *TERT*, *BTNL2*, *TP63*, *BPTF*) were presented in a GWAS of *EGFR*-positive lung adenocarcinoma in a larger study population in Japan (30). Data on the other six SNPs (12q13.13, HLA class II, *VTI1A*, 9p21.3, *FOXP4*, *ROS1/DCBLD1*) are presented for the first time in this study.

EGFR tumour mutation

Both FLCCA and JLCCS provided *EGFR* information. Details on DNA extraction methods used by each study center have been previously described (4,5,30). Genomic DNA was extracted from fresh, frozen or formalin-fixed, paraffin-embedded (FFPE) samples, following the respective manufacturer's protocols. Each participating study independently genotyped the tumour tissues from lung cancer cases for *EGFR* mutations on exons 19 and 21 using polymerase chain reaction (PCR) based sequencing (16,47,48), Peptide Nucleic Acid (PNC) Clamping (49), a cycleave PCR technique (50), high-resolution melting (HRM) analysis (51), invader assay (52) and SCORPION-ARMS (53). Cases with *EGFR* mutation on either exon 19 or 21 were defined as being *EGFR*-positive. Cases for whom genotyping for *EGFR* events yielded null findings were deemed as not having an *EGFR* mutation (*EGFR*-negative).

Statistical analyses

We restricted our analyses to 11 SNPs that achieved genome-wide significance (i.e., $P \leq 5 \times 10^{-8}$) in previous GWAS studies of various populations and subgroups in Asia: rs11610143 (12q13.13), rs2179920 (*HLA-DPB1*), rs2395185 (HLA class II), rs2736100 (*TERT*), rs3817963 (*BTNL2*), rs4488809 (*TP63*), rs7086803 (*VTI1A*), rs7216064 (*BPTF*), rs72658409 (9p21.3), rs7741164 (*FOXP4*) and rs9387478 (*ROS1/DCBLD1*). Each SNP was coded as the number of minor alleles (in controls) a subject carried (additive model). Subjects from FLCCA were consisted of three major racial/ethnic groups (self-reported) which included Chinese, Korean, and Japanese (Supplementary Material, Table S1). All analyses were restricted to lung adenocarcinoma patients only.

To assess associations between each SNP and *EGFR* mutations in the tumours (dichotomous, presence or absence of

EGFR mutation), logistic regression models adjusted for age (10-year categories: <40, 40–49, 50–59, 60–69 and ≥70), racial/ethnic groups (Chinese, Korean and Japanese) and significant eigenvectors (EV1, EV2, EV4, EV6, EV7 and EV8 for case-control analysis; EV1, EV2, EV6 and EV7 for *EGFR*-positive (or *EGFR*-negative) case-control analysis) were used to estimate odds ratios (ORs) and 95% confidence intervals (CIs). Three different comparisons were carried out – 1) case-control (FLCCA, Nanjing and JLCCS), 2) *EGFR*-positive (or *EGFR*-negative) case-control (included only cases with *EGFR* mutation from FLCCA and JLCCS), and 3) all cases only (compared cases with and without *EGFR* mutation from FLCCA and JLCCS). A fixed effect meta-analysis was used to estimate the overall effect in each of the analyses and the Cochran's Q statistic was used to test for heterogeneity across studies.

Sensitivity analyses for the 11 SNPs in the subgroup of subjects who did not contribute any *EGFR* mutation data yielded very similar results for overall case-control associations and had similar age distribution as those who contributed *EGFR* mutation data (data not shown). We also stratified the case-case analysis by the major racial/ethnic groups (Chinese and Japanese).

All statistical analyses were conducted using R, version 3.2.2 (54). Multiple comparisons were adjusted for by calculating the false discovery rate (FDR) using the Benjamini-Hochberg method (55). All statistical tests were conducted as two-sided, and a P-value of $< 5 \times 10^{-8}$ was considered as genome-wide significant for the case-control analyses and an FDR of 0.15 was considered significant for the case-case analyses.

Comparison of SNP associations between Asian and Western populations

We examined the associations of our 10 statistically significant SNPs and lung adenocarcinoma risk in a separate population of never-smokers (219 lung adenocarcinoma cases and 1,379 controls) and smokers (1,612 lung adenocarcinoma cases and 4,336 controls) of European ancestry in the NCI GWAS consisting of The Environment and Genetics in Lung Cancer Etiology (EAGLE) study; Prostate, Lung, Colon, Ovary (PLCO) Screening Trial; the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) Trial, and the Cancer Prevention Study-II (CPS-II) Nutrition Cohort (6). Further, to determine if there are shared regions in GWAS reports for lung adenocarcinoma between never-smoking women in Asia and predominately smoking studies in Western populations, we analysed known signals from Western populations that showed significant associations for all or specific lung cancer histologies to date (6,7,18–24) for those SNPs that we had genotyping data, in our never-smoking Asian women population.

Supplementary Material

Supplementary Material is available at *HMG* online.

Conflict of Interest statement. None declared.

Funding

This study was supported by a Grant-in-Aid for Scientific Research on Priority Areas from the Ministry of Education, Science, Sports, Culture and Technology of Japan, a Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry Health, Labor and Welfare of Japan, by Health and Labor Sciences Research Grants for Research on Applying Health Technology from the Ministry of Health, Labor and Welfare of Japan, by the National Cancer Center Research and Development Fund, the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (grant No. 2011-0016106), a grant of the National Project for Personalized Genomic Medicine, Ministry for Health & Welfare, Republic of Korea (A111218-11-GM04), the Program for Changjiang Scholars and Innovative Research Team in University in China (IRT_14R40 to K.C.), the National Science & Technology Pillar Program (2011BAI09B00), MOE 111 Project (B13016), the National Natural Science Foundation of China (No. 30772531, and 81272618), Guangdong Provincial Key Laboratory of Lung Cancer Translational Medicine (No. 2012A061400006), Special Fund for Research in the Public Interest from the National Health and Family Planning Commission of PRC (No. 201402031), and the Ministry of Science and Technology, Taiwan (MOST 103-2325-B-400-023 & 104-2325-B-400-012). The Japan Lung Cancer Study (JLCS) was supported in part by the Practical Research for Innovative Cancer Control from Japan Agency for Medical Research and Development (15ck0106096h0002) and the Management Expenses Grants from the Government to the National Cancer Center (26-A-1) for Biobank. BioBank Japan was supported by the Ministry of Education, Culture, Sports, Sciences and Technology of the Japanese government. The Japan Public Health Center-based prospective Study (the JPHC Study) was supported by the National Cancer Center Research and Development Fund (23-A-31[toku] and 26-A-2) (since 2011) and a Grant-in-Aid for Cancer Research from the Ministry of Health, Labour and Welfare of Japan (from 1989 to 2010). The Taiwan GELAC Study (Genetic Epidemiological Study for Lung AdenoCarcinoma) was supported by grants from the National Research Program on Genomic Medicine in Taiwan (DOH99-TD-G-111-028), the National Research Program for Biopharmaceuticals in Taiwan (MOHW 103-TDUPB-211-144003, MOST 103-2325-B-400-023) and the Bioinformatics Core Facility for Translational Medicine and Biotechnology Development (MOST 104-2319-B-400-002). This work was also supported by the Jinan Science Research Project Foundation (201102051), the National Key Scientific and Technological Project (2011ZX09307-001-04), the National Natural Science Foundation of China (No.81272293), the State Key Program of National Natural Science of China (81230067), the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. NRF-2014R1A2A2A05003665), Sookmyung Women's University Research Grants, Korea (1-1603-2048), Agency for Science, Technology and Research (A*STAR), Singapore and the US National Institute of Health Grant (1U19CA148127-01).

The overall GWAS project was supported by the intramural program of the US National Institutes of Health/National Cancer Institute. The following is a list of grants by study center: SKLCS (Y.T.K.)—National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (2011-0016106). (J.C.) – This work was supported by a grant from the National R&D Program for Cancer Control, Ministry of Health &Welfare, Republic of Korea (grant no. 0720550-2). (J.S.S) – grant number is A010250. WLCS (T.W.)—National Key Basic Research and Development Program (2011CB503800). SLCS (B.Z.)—National Nature Science Foundation of China (81102194). Liaoning Provincial Department of Education (LS2010168). China Medical Board (00726). GDS (Y.L.W.)—Foundation of Guangdong Science and Technology Department (2006B60101010, 2007A032000002, 2011A030400010). Guangzhou Science and Information

Technology Bureau (2011Y2-00014). Chinese Lung Cancer Research Foundation, National Natural Science Foundation of China (81101549). Natural Science Foundation of Guangdong Province (S2011010000792). TLCS (K.C., B.Q)—Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT), China (IRT1076). Tianjin Cancer Institute and Hospital. National Foundation for Cancer Research (US). FLCS (J.C.W., D.R., L.J.)—Ministry of Health (201002007). Ministry of Science and Technology (2011BAI09B00). National S&T Major Special Project (2011ZX09102-010-01). China National High-Tech Research and Development Program (2012AA02A517, 2012AA02A518). National Science Foundation of China (30890034). National Basic Research Program (2012CB944600). Scientific and Technological Support Plans from Jiangsu Province (BE2010715). NLCS (H.S.)—China National High-Tech Research and Development Program Grant (2009AA022705). Priority Academic Program Development of Jiangsu Higher Education Institution. National Key Basic Research Program Grant (2011CB503805). GEL-S (A.S.)—National Medical Research Council Singapore grant (NMRC/0897/2004, NMRC/1075/2006). (J.Liu)—Agency for Science, Technology and Research (A*STAR) of Singapore. GELAC (C.A.H.)—National Research Program on Genomic Medicine in Taiwan (DOH98-TDG-111-015). National Research Program for Biopharmaceuticals in Taiwan (DOH 100-TD-PB-111-TM013). National Science Council, Taiwan (NSC 100-2319-B-400-001). YLCS (Q.L.)—Supported by the intramural program of U.S. National Institutes of Health, National Cancer Institute. SWHS (W.Z., W-H.C., N.R.)—The work was supported by a grant from the National Institutes of Health (R37 CA70867) and the National Cancer Institute intramural research program, including NCI Intramural Research Program contract (N02 CP1101066). JLCS (K.M., T.K.)—Grants-in-Aid from the Ministry of Health, Labor, and Welfare for Research on Applying Health Technology and for the 3rd-term Comprehensive 10-year Strategy for Cancer Control; by the National Cancer Center Research and Development Fund; by Grant-in-Aid for Scientific Research on Priority Areas and on Innovative Area from the Ministry of Education, Science, Sports, Culture and —Technology of Japan. (W.P.)—NCI R01-CA121210. HKS (J.W.)—General Research Fund of Research Grant Council, Hong Kong (781511M).

The Environment and Genetics in Lung Cancer Etiology (EAGLE), Prostate, Lung, Colon, Ovary Screening Trial (PLCO), and Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) studies were supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute (NCI), Division of Cancer Epidemiology and Genetics. ATBC was also supported by U.S. Public Health Service contracts (N01-CN-45165, N01-RC-45035, and N01-RC-37004) from the NCI. PLCO was also supported by individual contracts from the NCI to the University of Colorado Denver (NO1-CN-25514), Georgetown University (NO1-CN-25522), the Pacific Health Research Institute (NO1-CN-25515), the Henry Ford Health System (NO1-CN-25512), the University of Minnesota, (NO1-CN-25513), Washington University (NO1-CN-25516), the University of Pittsburgh (NO1-CN-25511), the University of Utah (NO1-CN-25524), the Marshfield Clinic Research Foundation (NO1-CN-25518), the University of Alabama at Birmingham (NO1-CN-75022), Westat, Inc. (NO1-CN-25476), and the University of California, Los Angeles (NO1-CN-25404). The Cancer Prevention Study-II (CPS-II) Nutrition Cohort was supported by the American Cancer Society. The NIH Genes, Environment and Health Initiative (GEI) partly funded DNA extraction and statistical analyses (HG-06-033-NCI-01 and RO1HL091172-01), genotyping at the Johns Hopkins University Center for Inherited Disease Research.

References

1. Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D., Forman, D. and Bray, F. (2013). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer*, doi:10.1002/ijc.29210 PMID:25220842
2. Chen, W., Zheng, R., Zeng, H. and Zhang, S. (2015) Epidemiology of lung cancer in China. *Thorac. Cancer*, **6**, 209–215.
3. Sun, S., Schiller, J.H. and Gazdar, A.F. (2007) Lung cancer in never smokers–a different disease. *Nat. Rev. Cancer*, **7**, 778–790.
4. Lan, Q., Hsiung, C.A., Matsuo, K., Hong, Y.C., Seow, A., Wang, Z., Hosgood, H.D., 3rd, Chen, K., Wang, J.C., Chatterjee, N., *et al.* (2012) Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. *Nat. Genet.*, **44**, 1330–1335.
5. Wang, Z., Seow, W.J., Shiraishi, K., Hsiung, C.A., Matsuo, K., Liu, J., Chen, K., Yamji, T., Yang, Y., Chang, I.S., *et al.* (2016) Meta-analysis of genome-wide association studies identifies multiple lung cancer susceptibility loci in never-smoking Asian women. *Hum. Mol. Genet.*, **25**, 620–629.
6. 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.
7. Amos, C.I., Wu, X.F., Broderick, P., Gorlov, I.P., Gu, J., Eisen, T., Dong, Q., Zhang, Q., Gu, X.J., 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.
8. 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.
9. 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.
10. Wu, C., Hu, Z., Yu, D., Huang, L., Jin, G., Liang, J., Guo, H., Tan, W., Zhang, M., Qian, J., *et al.* (2009) Genetic variants on chromosome 15q25 associated with lung cancer risk in Chinese populations. *Cancer Res.*, **69**, 5065–5072.
11. Truong, T., Hung, R.J., Amos, C.I., Wu, X., Bickeboller, H., Rosenberger, A., Sauter, W., Illig, T., Wichmann, H.E., Risch, A., *et al.* (2010) 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.*, **102**, 959–971.
12. da Cunha Santos, G., Shepherd, F.A. and Tsao, M.S. (2011) EGFR mutations and lung cancer. *Annu. Rev. Pathol.*, **6**, 49–69.
13. Dearden, S., Stevens, J., Wu, Y.L. and Blowers, D. (2013) Mutation incidence and coincidence in non small-cell lung cancer: meta-analyses by ethnicity and histology (mutMap). *Ann. Oncol.*, **24**, 2371–2376.
14. Rosell, R., Moran, T., Queralt, C., Porta, R., Cardenal, F., Camps, C., Majem, M., Lopez-Vivanco, G., Isla, D., Provencio,

M., et al. (2009) Screening for epidermal growth factor receptor mutations in lung cancer. N. Engl. J. Med., 361, 958–967.
15. Mitsudomi, T., Kosaka, T., Endoh, H., Horio, Y., Hida, T., Mori, S., Hatooka, S., Shinoda, M., Takahashi, T. and Yatabe, Y. (2005) Mutations of the epidermal growth factor receptor gene predict prolonged survival after gefitinib treatment in patients with non–small-cell lung cancer with postoperative recurrence. J. Clin. Oncol., 23, 2513–2520.
16. Lynch, T.J., Bell, D.W., Sordella, R., Gurubhagavatula, S., Okimoto, R.A., Brannigan, B.W., Harris, P.L., Haserlat, S.M., Supko, J.G., Haluska, F.G., et al. (2004) Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib. N. Engl. J. Med., 350, 2129–2139.
17. Shigematsu, H., Lin, L., Takahashi, T., Nomura, M., Suzuki, M., Wistuba, I.I., Fong, K.M., Lee, H., Toyooka, S., Shimizu, N., et al. (2005) Clinical and Biological Features Associated With Epidermal Growth Factor Receptor Gene Mutations in Lung Cancers. J. Natl. Cancer Inst., 97, 339–346.
18. Brenner, D.R., Amos, C.I., Brhane, Y., Timofeeva, M.N., Caporaso, N., Wang, Y., Christiani, D.C., Bickeboller, H., Yang, P., Albanes, D., et al. (2015) Identification of lung cancer histology-specific variants applying Bayesian framework variant prioritization approaches within the TRICL and ILCCO consortia. Carcinogenesis, 36, 1314–1326.
19. Pande, M., Spitz, M.R., Wu, X., Gorlov, I.P., Chen, W.V. and Amos, C.I. (2011) Novel genetic variants in the chromosome 5p15.33 region associate with lung cancer risk. Carcinogenesis, 32, 1493–1499.
20. Park, S.L., Fesinmeyer, M.D., Timofeeva, M., Caberto, C.P., Kocarnik, J.M., Han, Y., Love, S.A., Young, A., Dumitrescu, L., Lin, Y., et al. (2014) Pleiotropic associations of risk variants identified for other cancers with lung cancer risk: the PAGE and TRICL consortia. J. Natl. Cancer Inst., 106, dju061.
21. Timofeeva, M.N., Hung, R.J., Rafnar, T., Christiani, D.C., Field, J.K., Bickeboller, H., Risch, A., McKay, J.D., Wang, Y., Dai, J., et al. (2012) Influence of common genetic variation on lung cancer risk: meta-analysis of 14 900 cases and 29 485 controls. Hum. Mol. Genet., 21, 4980–4995.
22. Wang, Y., McKay, J.D., Rafnar, T., Wang, Z., Timofeeva, M.N., Broderick, P., Zong, X., Laplana, M., Wei, Y., Han, Y., et al. (2014) Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. Nat. Genet., 46, 736–741.
23. Kachuri, L., Amos, C.I., McKay, J.D., Johansson, M., Vineis, P., Bueno-de-Mesquita, H.B., Boutron-Ruault, M.C., Johansson, M., Quiros, J.R., Sieri, S., et al. (2016) Fine mapping of chromosome 5p15.33 based on a targeted deep sequencing and high density genotyping identifies novel lung cancer susceptibility loci. Carcinogenesis, 37, 96–105.
24. Shi, J., Chatterjee, N., Rotunno, M., Wang, Y., Pesatori, A.C., Consonni, D., Li, P., Wheeler, W., Broderick, P., Henrion, M., et al. (2012) Inherited variation at chromosome 12p13.33, including RAD52, influences the risk of squamous cell lung carcinoma. Cancer Discov., 2, 131–139.
25. Dai, M., Lu, J.J., Guo, W., Yu, W., Wang, Q., Tang, R., Tang, Z., Xiao, Y., Li, Z., Sun, W., et al. (2015) BPTF promotes tumor growth and predicts poor prognosis in lung adenocarcinomas. Oncotarget, 6, 33878–33892.
26. Valentonyte, R., Hampe, J., Huse, K., Rosenstiel, P., Albrecht, M., Stenzel, A., Nagy, M., Gaede, K.I., Franke, A., Haesler, R., et al. (2005) Sarcoidosis is associated with a truncating splice site mutation in BTNL2. Nat. Genet., 37, 357–364.
27. Bergethon, K., Shaw, A.T., Ou, S.H., Katayama, R., Lovly, C.M., McDonald, N.T., Massion, P.P., Siwak-Tapp, C., Gonzalez, A., Fang, R., et al. (2012) ROS1 rearrangements define a unique molecular class of lung cancers. J. Clin. Oncol., 30, 863–870.
28. Davies, K.D., Mahale, S., Astling, D.P., Aisner, D.L., Le, A.T., Hinz, T.K., Vaishnavi, A., Bunn, P.A., Jr., Heasley, L.E., Tan, A.C., et al. (2013) Resistance to ROS1 inhibition mediated by EGFR pathway activation in non-small cell lung cancer. PLoS One, 8, e82236.
29. Borczuk, A.C., Shah, L., Pearson, G.D.N., Walter, K.L., Wang, L.Q., Austin, J.H.M., Friedman, R.A. and Powell, C.A. (2004) Molecular signatures in biopsy specimens of lung cancer. Am. J. Respir. Crit. Care Med., 170, 167–174.
30. Shiraishi, K., Okada, Y., Takahashi, A., Kamatani, Y., Momozawa, Y., Ashikawa, K. and K, H. (2016) Association of variations in HLA-class II and other loci with susceptibility to lung adenocarcinoma with EGFR mutation. Nat. Commun., 7, 12451.
31. Lan, Q., Chapman, R.S., Schreinemachers, D.M., Tian, L. and He, X. (2002) Household stove improvement and risk of lung cancer in Xuanwei, China. J. Natl. Cancer Inst., 94, 826–835.
32. Hosgood, H.D., 3rd, Wei, H., Sapkota, A., Choudhury, I., Bruce, N., Smith, K.R., Rothman, N. and Lan, Q. (2011) Household coal use and lung cancer: systematic review and meta-analysis of case-control studies, with an emphasis on geographic variation. Int. J. Epidemiol., 40, 719–728.
33. Matsuo, K., Hiraki, A., Ito, H., Kosaka, T., Suzuki, T., Hirose, K., Wakai, K., Yatabe, Y., Mitsudomi, T. and Tajima, K. (2008) Soy consumption reduces the risk of non-small-cell lung cancers with epidermal growth factor receptor mutations among Japanese. Cancer Sci., 99, 1202–1208.
34. Gao, Y.T., Blot, W.J., Zheng, W., Ershow, A.G., Hsu, C.W., Levin, L.I., Zhang, R. and Fraumeni, J.F.J. (1987) Lung cancer among Chinese women. Int. J. Cancer, 40, 604–609.
35. Shiraishi, K., Kunitoh, H., Daigo, Y., Takahashi, A., Goto, K., Sakamoto, H., Ohnami, S., Shimada, Y., Ashikawa, K., Saito, A., et al. (2012) A genome-wide association study identifies two new susceptibility loci for lung adenocarcinoma in the Japanese population. Nat. Genet., 44, 900–903.
36. Hosgood, H.D., Pao, W., Rothman, N., Hu, W., Pan, Y.H., Kuchinsky, K., Jones, K.D., Xu, J., Vermeulen, R., Simko, J., et al. (2013) Driver mutations among never smoking female lung cancer tissues in China identify unique EGFR and KRAS mutation pattern associated with household coal burning. Resp. Med., 107, 1755–1762.
37. 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.
38. Piao, J.M., Kim, H.N., Song, H.R., Kweon, S.S., Choi, J.S., Yun, W.J., Kim, Y.C., Oh, I.J., Kim, K.S. and Shin, M.H. (2011) p53 codon 72 polymorphism and the risk of lung cancer in a Korean population. Lung Cancer, 73, 264–267.
39. Qian, B., Zhang, H., Zhang, L., Zhou, X., Yu, H. and Chen, K. (2010) Association of genetic polymorphisms in DNA repair pathway genes with non-small cell lung cancer risk. Lung Cancer, 73, 138–146.
40. Ito, H., McKay, J.D., Hosono, S., Hida, T., Yatabe, Y., Mitsudomi, T., Brennan, P., Tanaka, H. and Matsuo, K. (2012) Association between a genome-wide association study-identified locus and the risk of lung cancer in Japanese population. J. Thorac. Oncol., 7, 790–798.
41. Jou, Y.S., Lo, Y.L., Hsiao, C.F., Chang, G.C., Tsai, Y.H., Su, W.C., Chen, Y.M., Huang, M.S., Chen, H.L., Chen, C.J., et al. (2009) Association of an EGFR intron 1 SNP with never-

smoking female lung adenocarcinoma patients. *Lung Cancer*, **64**, 251–256.
42. Jung, H.Y., Whang, Y.M., Sung, J.S., Shin, H.D., Park, B.L., Kim, J.S., Shin, S.W., Seo, H.Y., Seo, J.H. and Kim, Y.H. (2008) Association study of TP53 polymorphisms with lung cancer in a Korean population. *J. Hum. Genet.*, **53**, 508–514.
43. Kim, J.H., Kim, H., Lee, K.Y., Choe, K.H., Ryu, J.S., Yoon, H.I., Sung, S.W., Yoo, K.Y. and Hong, Y.C. (2006) Genetic polymorphisms of ataxia telangiectasia mutated affect lung cancer risk. *Hum. Mol. Genet.*, **15**, 1181–1186.
44. Park, J.Y., Park, S.H., Choi, J.E., Lee, S.Y., Jeon, H.S., Cha, S.I., Kim, C.H., Park, J.H., Kam, S., Park, R.W., *et al.* (2002) Polymorphisms of the DNA repair gene xeroderma pigmentosum group A and risk of primary lung cancer. *Cancer Epidemiol. Biomarkers Prev.*, **11**, 993–997.
45. Nakamura, Y. (2007) The BioBank Japan Project. *Clin. Adv. Hematol. Oncol.*, **5**, 696–697.
46. Hsiung, C.A., Lan, Q., Hong, Y.C., Chen, C.J., Hosgood, H.D., Chang, I.S., Chatterjee, N., Brennan, P., Wu, C., Zheng, W., *et al.* (2010) The 5p15.33 locus is associated with risk of lung adenocarcinoma in never-smoking females in Asia. *PLoS Genet.*, **6**, pii: e1001051.
47. An, S.J., Chen, Z.H., Su, J., Zhang, X.C., Zhong, W.Z., Yang, J.J., Zhou, Q., Yang, X.N., Huang, L., Guan, J.L., *et al.* (2012) Identification of enriched driver gene alterations in subgroups of non-small cell lung cancer patients based on histology and smoking status. *PLoS One*, **7**, e40109.
48. Tam, I.Y., Chung, L.P., Suen, W.S., Wang, E., Wong, M.C., Ho, K.K., Lam, W.K., Chiu, S.W., Girard, L., Minna, J.D., *et al.* (2006) Distinct epidermal growth factor receptor and KRAS mutation patterns in non-small cell lung cancer patients with different tobacco exposure and clinicopathologic features. *Clin. Cancer. Res.*, **12**, 1647–1653.
49. Yoon, S.H., Choi, Y.D., Oh, I.J., Kim, K.S., Choi, H., Chang, J., Shin, H.J., Park, C.K. and Kim, Y.C. (2015) Peptide Nucleic Acid Clamping Versus Direct Sequencing for the Detection of EGFR Gene Mutation in Patients with Non-small Cell Lung Cancer. *Cancer Res. Treat.*, **47**, 661–669.
50. Yatabe, Y., Hida, T., Horio, Y., Kosaka, T., Takahashi, T. and Mitsudomi, T. (2006) A rapid, sensitive assay to detect EGFR mutation in small biopsy specimens from lung cancer. *J. Mol. Diagn.*, **8**, 335–341.
51. Nomoto, K., Tsuta, K., Takano, T., Fukui, T., Fukui, T., Yokozawa, K., Sakamoto, H., Yoshida, T., Maeshima, A.M., Shibata, T., *et al.* (2006) Detection of EGFR mutations in archived cytologic specimens of non-small cell lung cancer using high-resolution melting analysis. *Am. J. Clin. Pathol.*, **126**, 608–615.
52. Naoki, K., Soejima, K., Okamoto, H., Hamamoto, J., Hida, N., Nakachi, I., Yasuda, H., Nakayama, S., Yoda, S., Satomi, R., *et al.* (2011) The PCR-invader method (structure-specific 5' nuclease-based method), a sensitive method for detecting EGFR gene mutations in lung cancer specimens; comparison with direct sequencing. *Int. J. Clin. Oncol.*, **16**, 335–344.
53. Horiike, A., Kimura, H., Nishio, K., Ohyanagi, F., Satoh, Y., Okumura, S., Ishikawa, Y., Nakagawa, K., Horai, T. and Nishio, M. (2007) Detection of epidermal growth factor receptor mutation in transbronchial needle aspirates of non-small cell lung cancer. *Chest*, **131**, 1628–1634.
54. R Core Team. (2014). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. https://www.R-project.org/; date last accessed May 20, 2016.
55. Benjamini, Y. and Hochberg, Y. (1995) Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Series B. Stat. Methodol*, **57**, 289–300.