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 Seow1,2,*,†, Keitaro Matsuo3,†, Chao Agnes Hsiung4,‡, Kouya Shiraishi5,†, Minsun Song1,6,†, Hee Nam Kim7,†, Maria Pik Wong8,†, Yun-Chul Hong9,†, H. Dean Hosgood III10,†, Zhaoming Wang1,11,†, I-Shou Chang12, Jiu-Cun Wang13,14, Nilanjan Chatterjee1, Margaret Tucker1, Hu Wei1, Tetsuya Mitsudomi15, Wei Zheng16, Jin Hee Kim17, Baosen Zhou18, Neil E. Caporaso1, Demetrius Albanes1, Min-Ho Shin7, Lap Ping Chung8, She-Juan An19, Ping Wang20, Hong Zheng21, Yasushi Yatabe22, Xu-Chao Zhang19, Young Tae Kim23, Xiao-Ou Shu16, Young-Chul Kim24,25, Bryan A. Bassig1, Jiang Chang26, James Chung Man Ho27, Bu-Tian Ji1, Michiaki Kubo28, Yataro Daigo29,30, Hideki Ito31, Yukihide Momozawa28, Kyota Ashikawa28, Yoichiro Kamatani32, Takayuki Honda5, Hiromi Sakamoto33, Hideo Kunitoh34, Koji Tsuta35, Shun-Ichi Watanabe36, Hiroshi Nokihara37, Yohei Miyagi38, Haruhiko Nakayama39, Shingo Matsumoto40, Masahiro Tsuboi41, Koichi Goto42, Zhihua Yin18, Jianxin Shi1, Atsushi Takahashi32, Akiteru Goto43, Yoshihiro Minamiya44, Kimihiro Shimizu45, Kazumi Tanaka45, Tangchun Wu46, Fusheng Wei47, Jason Y.Y. Wong1, Fumihiko Matsuda48, Jian Su19, Yeul Hong Kim49, In-Jae Oh24,25, Fengiu Song21, Victor Ho Fun Lee50, Wu-Chou Su51, Yuh-Min Chen52,53, Gee-Chen Chang54,55, Kuan-Yu Chen56, Ming-Shyan Huang57, Pan-Chyr Yang58, Hsien-Chih Lin4, Yong-Bing Xiang59, Adeline Seow2, Jae Yong Park60, Sun-Seog Kweon7,80, Chien-Jen Chen61,
Haixin Li21, Yu-Tang Gao59, Chen Wu62, Biyun Qian21, Daru Lu13,14, Jianjun Liu2,63,64, Hyo-Sung Jeon65, Chin-Fu Hsiao4, Jae Sook Sung49, Ying-Huang Tsai66, Yoo Jin Jung23, Huan Guo46, Zhibin Hu67, Wen-Chang Wang68, Charles C. Chung1,69, Charles Lawrence70, Laurie Burdett1,69, Meredith Yeager1,69, Kevin B. Jacobs1,69, Amy Hutchinson1,69, Sonja I. Berndt1, Xingzhou He71, Wei Wu18, Junwen Wang72,73, Yuqing Li99, Jin Eun Choi65, Kyong Hwa Park49, Sook Whan Sung74, Li Liu75, Chang Hyun Kang23, Lingmin Hu76,77, Chung-Hsing Chen12, Tsung-Ying Yang55, Jun Xu78, Peng Guan18,79, Wen Tan62, Chih-Liang Wang81, Alan Dart Loon Sihoe82, Ying Chen2, Yi Young Choi65, Jen-Yu Hung57, Jun Suk Kim83, Ho-Il Yoon84, Quyin Cai16, Chien-Chung Lin51, In Kyu Park23, Ping Xu85, Jing Dong76,77, Christopher Kim1, Qincheng He18, Reury-Peng Perg6, Chih-Yi Chen87,88, Roel Vermeulen89, Junjie Wu13,14, Wei-Yen Lim90, Kun-Chieh Chen55, John K.C. Chan91, Minjie Chu76,77, Yao-Jen Li61, Jihua Li92, Hongyan Chen13,14, Chong-Jen Yu58, Li Jin13,14, Yen-Li Lo4, Ying-Hsiang Chen4, Joseph F. Fraumeni, Jr1, Jie Liu93, Taiki Yamaji94, Yang Yang95, Belynda Hicks1,69, Kathleen Wyatt1,69, Shengchao A. Li1,69, Juncheng Dai67, Hongxia Ma57, Guangfu Jin67, Bao Song93, Zhehai Wang93, Sensen Cheng93, Xuelian Li18,79, Yangwu Ren18,79, Ping Cui21, Motoki Iwashaki94, Taichi Shimazu94, Shoichiro Tsugane94, Junjie Zhu95, Gening Jiang95, Ke Fei95, Guoping Wu47, Li-Hsin Chien4, Hui-Ling Chen4, Yu-Chun Su4, Fang-Yu Tsai12, Yi-Song Chen4, Jinming Yu93, Victoria L. Stevens96, Ite A. Laird-Offringa97, Crystal N. Marconett97, Dongxin Lin62,‡, Kexin Chen21,‡, Yi-Long Wu19,‡, Maria Teresa Landi1,‡, Hongbing Shen67,98,‡, Nathaniel Rothman1,‡, Takashi Kohno5,‡, Stephen J. Chanock1,‡ and Qing Lan1,‡
Division of Thoracic Surgery, Kinki University School of Medicine, Sayama, Japan, Division of Epidemiology, Department of Medicine, Vanderbilt University Medical Center and Vanderbilt-Ingram Cancer Center, Nashville, TN, USA, Department of Integrative Bioscience & Biotechnology, Sejong University, Seoul, Republic of Korea, Department of Epidemiology, School of Public Health, China Medical University, Shenyang, People’s Republic of China, 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, 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, Department of Epidemiology and Biostatistics, Tianjin Medical University Cancer Institute and Hospital, Tianjin, People’s Republic of China, Department of Pathology and Molecular Diagnostics, Aichi Cancer Center Central Hospital, Nagoya, Japan, Department of Thoracic and Cardiovascular Surgery, Cancer Research Institute, Seoul National University College of Medicine, Seoul, Republic of Korea, Lung and Esophageal Cancer Clinic, Chonnam National University Hwasun Hospital, Hwasun-eup, Republic of Korea, Department of Internal Medicine, Chonnam National University Medical School, Gwangju, Republic of Korea, Department of Etiology & Carcinogenesis, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, People’s Republic of China, Department of Medicine, The University of Hong Kong, Queen Mary Hospital, Pokfulam Road, Hong Kong, Laboratory for Genotyping Development, Center for Integrative Medical Sciences, RIKEN, Yokohama, Japan, Center for Antibody and Vaccine Therapy, Research Hospital, Institute of Medical Science, The University of Tokyo, Tokyo, Japan, Department of Medical Oncology and Cancer Center, Shiga University of Medical Science, Otsu, Japan, Division of Epidemiology & Prevention, Aichi Cancer Center Research Institute, Nagoya, Japan, Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan, Division of Genetics, National Cancer Center Research Institute, Tokyo, Japan, Department of Medical Oncology, Japanese Red Cross Medical Center, Tokyo, Japan, Department of Pathology, National Cancer Center Hospital, Tokyo, Japan, Division of Thoracic Surgery, National Cancer Center Hospital, Tokyo, Japan, Department of Thoracic Oncology, National Cancer Center Hospital, Tokyo, Japan, Molecular Pathology and Genetics Division, Kanagawa Cancer Center Research Institute, Kanagawa, Japan, Department of Thoracic Surgery, Kanagawa Cancer Center, Kanagawa, Japan, Division of Translational Research, Exploratory Oncology Research and Clinical Trial Center (EPOC), National Cancer Center, Chiba, Japan, Department of Thoracic Surgery, National Cancer Center Hospital East, Chiba, Japan, Department of Thoracic Oncology, National Cancer Center Hospital East, Japan, Department of Cellular and Organ Pathology and Department of Thoracic Surgery, Graduate School of Medicine, Akita University, Akita City, Japan, Department of Integrative Center of General Surgery, Gunma University Hospital, Gunma, Japan, 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, China National Environmental Monitoring Center, Beijing, People’s Republic of China, Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan, Department of Internal Medicine, Division of Oncology/Hematology, College of Medicine, Korea University Anam Hospital, Seoul, Republic of Korea, Department of Clinical Oncology, The University of Hong Kong, Queen Mary Hospital, Hong Kong, Department of Internal Medicine, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan, Department of Chest Medicine, Taipei Veterans General Hospital, Taipei, Taiwan, College of Medical Science and Technology, Taipei Medical University, Taipei, Taiwan, School of Medicine, Faculty of Medicine, National Yang-Ming University, Taipei, Taiwan, Division of Chest Medicine, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan, Division of Pulmonary Medicine, Department of Internal Medicine, National Taiwan University Hospital and College of Medicine, National Taiwan University, Taipei, Taiwan, Department of Internal Medicine, Kaohsiung Medical University Hospital, School of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan, Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan, Department of Epidemiology, Shanghai Cancer Institute, Shanghai, People’s Republic of China, Lung Cancer Center, Kyungpook National University Medical Center, Daegu, Republic of Korea, Genomic Research Center, Academia Sinica, Taipei, Taiwan, Department of Etiology & Carcinogenesis and State Key Laboratory of Molecular Oncology, Cancer
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-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 genetic heterogeneity in lung adenocarcinoma.
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 (BPTF, odds ratio (OR) $= 0.86$, $P = 6.19 \times 10^{-8}$) 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 (ROSI/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, we observed that the SNP marker, rs7216064 (BPTF, odds ratio (OR) $= 0.86$, $P = 6.19 \times 10^{-8}$) 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 (ROSI/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.
Taqman data from the Japanese Lung Cancer Collaborative Study (JLCCS). We excluded FLCCA GWAS and Taqman data from overlapping subjects with the JLCCS. Case–control analyses were adjusted for age (Chr, chromosome; MAF, minor allele frequency; Ca, cases; Co, controls; OR, odds ratio; CI, confidence interval. (rs7216064) BPTF

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

<table>
<thead>
<tr>
<th>SNP</th>
<th>Nearest gene(s)</th>
<th>Chr</th>
<th>Major/minor allele</th>
<th>MAF</th>
<th>No. of subjects</th>
<th>OR (95% CI)</th>
<th>P</th>
<th>Phet</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4488809</td>
<td>TP63</td>
<td>3q28</td>
<td>T/C</td>
<td>0.47/0.53</td>
<td>7,448/7,007</td>
<td>0.80 (0.76–0.85)</td>
<td>4.30 × 10^{-17}</td>
<td>8.65 × 10^{-1}</td>
</tr>
<tr>
<td>rs2736100</td>
<td>TERT</td>
<td>5p15.33</td>
<td>A/C</td>
<td>0.49/0.46</td>
<td>7,505/7,070</td>
<td>1.43 (1.36–1.50)</td>
<td>6.12 × 10^{-13}</td>
<td>8.67 × 10^{-1}</td>
</tr>
<tr>
<td>rs7741164</td>
<td>FOXP4</td>
<td>6p21.1</td>
<td>G/A</td>
<td>0.37/0.33</td>
<td>10,531/10,648</td>
<td>1.17 (1.12–1.22)</td>
<td>3.96 × 10^{-11}</td>
<td>8.82 × 10^{-1}</td>
</tr>
<tr>
<td>rs3817963</td>
<td>BTN2L2</td>
<td>6p21.3</td>
<td>T/C</td>
<td>0.33/0.29</td>
<td>7,255/6,745</td>
<td>1.16 (1.10–1.22)</td>
<td>1.63 × 10^{-7}</td>
<td>5.57 × 10^{-2}</td>
</tr>
<tr>
<td>rs2179920</td>
<td>HLA-DPB1</td>
<td>6p21.32</td>
<td>C/T</td>
<td>0.14/0.13</td>
<td>7,457/7,020</td>
<td>1.17 (1.09–1.26)</td>
<td>1.69 × 10^{-5}</td>
<td>1.37 × 10^{-2}</td>
</tr>
</tbody>
</table>

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.

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, 40–49, 50–59, 60–69, ≥70) in the JLCCS; adjusted for age (<40, 40–49, 50–59, 60–69, ≥70) in the Nanjing study; adjusted for age (<40, 40–49, 50–59, 60–69, ≥70) in the JLCCS; adjusted for age (<40, 40–49, 50–59, 60–69, ≥70) in the Nanjing study; P < 5 × 10^{-8} are considered genome-wide significant.

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

Table 2. Association of GWAS identified SNPs and lung adenocarcinoma risk among never-smoking Asian women

<table>
<thead>
<tr>
<th>SNP</th>
<th>Nearest gene(s)</th>
<th>Chr</th>
<th>Major/minor allele</th>
<th>MAF</th>
<th>No. of subjects</th>
<th>Case-control</th>
<th>Case-Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4488809</td>
<td>TP63</td>
<td>3q28</td>
<td>T/C</td>
<td>0.47/0.53</td>
<td>1,949/6,687</td>
<td>0.82 (0.75–0.89)</td>
<td>8.46 × 10^{-6}</td>
</tr>
<tr>
<td>rs2736100</td>
<td>TERT</td>
<td>5p15.33</td>
<td>A/C</td>
<td>0.47/0.39</td>
<td>1,947/6,750</td>
<td>1.49 (1.36–1.63)</td>
<td>1.64 × 10^{-17}</td>
</tr>
<tr>
<td>rs7741164</td>
<td>FOXP4</td>
<td>6p21.1</td>
<td>G/A</td>
<td>0.33/0.33</td>
<td>1,908/6,239</td>
<td>1.19 (1.08–1.31)</td>
<td>3.19 × 10^{-4}</td>
</tr>
<tr>
<td>rs3817963</td>
<td>BTN2L2</td>
<td>6p21.3</td>
<td>T/C</td>
<td>0.33/0.29</td>
<td>1,326/6,749</td>
<td>1.18 (1.06–1.33)</td>
<td>1.52 × 10^{-3}</td>
</tr>
<tr>
<td>rs2179920</td>
<td>HLA-DPB1</td>
<td>6p21.32</td>
<td>C/T</td>
<td>0.18/0.17</td>
<td>1,947/6,700</td>
<td>1.29 (1.15–1.45)</td>
<td>1.42 × 10^{-5}</td>
</tr>
<tr>
<td>rs2395185</td>
<td>ROS1/DCBLD1</td>
<td>6q22.2</td>
<td>C/A</td>
<td>0.42/0.37</td>
<td>1,952/6,744</td>
<td>1.23 (1.12–1.35)</td>
<td>1.24 × 10^{-5}</td>
</tr>
<tr>
<td>rs9387478</td>
<td>BPTF</td>
<td>17q24.3</td>
<td>A/G</td>
<td>0.44/0.49</td>
<td>1,951/6,745</td>
<td>0.84 (0.77–0.92)</td>
<td>9.36 × 10^{-5}</td>
</tr>
<tr>
<td>rs2658409</td>
<td>TP63</td>
<td>3q28</td>
<td>T/C</td>
<td>0.47/0.53</td>
<td>1,324/6,687</td>
<td>0.80 (0.72–0.88)</td>
<td>3.42 × 10^{-6}</td>
</tr>
<tr>
<td>rs7086803</td>
<td>VTI1A</td>
<td>10q25.2</td>
<td>G/A</td>
<td>0.29/0.24</td>
<td>1,952/6,693</td>
<td>1.29 (1.16–1.42)</td>
<td>8.07 × 10^{-7}</td>
</tr>
<tr>
<td>rs1161043</td>
<td>DCBLD1</td>
<td>6q22.2</td>
<td>C/A</td>
<td>0.30/0.34</td>
<td>1,952/6,694</td>
<td>0.78 (0.71–0.86)</td>
<td>6.05 × 10^{-7}</td>
</tr>
<tr>
<td>rs7216064</td>
<td>BPTF</td>
<td>17q24.3</td>
<td>A/G</td>
<td>0.27/0.33</td>
<td>1,946/6,661</td>
<td>0.79 (0.72–0.87)</td>
<td>1.46 × 10^{-6}</td>
</tr>
</tbody>
</table>
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 (VIT1A), 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. 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. 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); 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. 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. Further, in vitro evidence in NSCLC tumour cells indicates that the EGFR pathway is activated as a mechanism of resistance to ROS1 inhibition, 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. Our study demonstrates the association of this region in never-smoking Asian women, consistent with a concurrent report. 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. 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, household coal use, diet and cooking fumes. 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 (T LCS), 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 660k SNP microarray, Illumina 370K SNP microarray, Illumina 610Q SNP microarray, Affymetrix-Genome-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). J LCCS

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 (PNA) 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 < 5 x 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 × 10^{-4} 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 Epidemiology (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 (6,18–24) 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 (2011BA09B000), 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 (2011ZX09007-001-04), 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-2014R1A2A6A05003665), 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 (U19CA148127-01).

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.W.)—Foundation of Guangdong Science and Technology Department (2006B60101010, 2007A03200002, 2011A030400010). Guangzhou Science and Information

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 Intramural Research Program contract (N01-CN-45165, N01-RC-45035, and N01-RC-37004). HKS (J.W.)—National Science Council, Taiwan (NSC 100-2319-B-009-001). YCLS (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 CP101066). 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).

genotyping at the Johns Hopkins University Center for Inherited Disease Research.

References


14. Rosell, R., Moran, T., Queralt, C., Porta, R., Cardenal, F., Camps, C., Majem, M., Lopez-Vivanco, G., Isla, D., Provencio,


smoking female lung adenocarcinoma patients. Lung Cancer, 64, 251–256.