ORIGINAL ARTICLE

Rare Variants in Known Susceptibility Loci and Their Contribution to Risk of Lung Cancer

Yanhong Liu, PhD,[a,*] Christine M. Lusk, PhD,[b] Michael H. Cho, PhD,[c,†] Edwin K. Silverman, PhD,[c,†] Dandi Qiao, PhD,[c,†] Ruyang Zhang, PhD,[d] Michael E. Scheurer, PhD,[e] Farrah Kheradmand, PhD,[a,f] David A. Wheeler, PhD,[g] Spiridon Tsavachidis, MS,[a] Georgina Armstrong, MPH,[a] Dakai Zhu, MS,[a,h] Ignacio I. Wistuba, MD,[i] Chi-Wan B. Chow, MD,[i] Carmen Behrens, MD,[j] Claudio W. Pikielny, PhD,[k] Christine Neslund-Dudas, PhD,[l] Susan M. Pinney, PhD,[m] Marshall Anderson, PhD,[m] Elena Kupert, MS,[m] Joan Bailey-Wilson, PhD,[n] Colette Gaba, MPH,[o] Diptasri Mandal, PhD,[p] Ming You, MD,[q] Mariza de Andrade, PhD,[r] Ping Yang, PhD, MD,[r] John K. Field, PhD,[s] Triantafillos Liloglou, PhD,[s] Michael Davies, PhD,[s] Jolanta Lissowska, PhD,[t] Beata Swiatkowska, PhD,[u] David Zaridze, PhD,[v] Anush Mukeriya, PhD,[v] Vladimir Janout, PhD,[w] Ivana Holcatova, PhD, MD,[x] Dana Mates, PhD, MD,[y] Sasa Milosavljevic, PhD,[z] Ghislaine Scelo, PhD,[aa] Paul Brennan, PhD,[aa] James McKay, PhD,[aa] Geoffrey Liu, PhD,[bb] Rayjean J. Hung, PhD,[cc] David C. Christiani, MD,[d] Ann G. Schwartz, PhD,[b] Christopher I. Amos, PhD,[a,h] Margaret R. Spitz, MD[a]

[a]*Dan L. Duncan Comprehensive Cancer Center, Department of Medicine, Baylor College of Medicine, Houston, Texas*
[b]*Karmanos Cancer Institute, Wayne State University, Detroit, Michigan*
[c]*Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts*
[d]*Harvard University School of Public Health, Boston, Massachusetts*
[e]*Department of Pediatrics, Baylor College of Medicine, Houston, Texas*
[f]*Michael E. DeBakey Veterans Affairs Medical Center; Houston, Texas*
[g]*Department of Molecular and Human Genetics, Human Genome Sequence Center, Baylor College of Medicine, Houston, Texas*
[h]*Institute for Clinical and Translational Research, Baylor College of Medicine, Houston, Texas*
[i]*Department of Translational Molecular Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas*
[j]*Department of Thoracic/Head and Neck Medical Oncology, The University of Texas M. D. Anderson Cancer Center, Houston, Texas*
[k]*Department of Biomedical Data Science, Geisel School of Medicine, Dartmouth College, Lebanon, New Hampshire*
[l]*Department of Public Health Sciences, Henry Ford health System, Detroit, Michigan*
[m]*University of Cincinnati College of Medicine, Cincinnati, Ohio*
[n]*National Human Genome Research Institute, Bethesda, Maryland*
[o]*The University of Toledo College of Medicine, Toledo, Ohio*
[p]*Louisiana State University Health Sciences Center, New Orleans, Louisiana*
[q]*Medical College of Wisconsin, Milwaukee, Wisconsin*
[r]*Mayo Clinic College of Medicine, Rochester, Minnesota*
[s]*Roy Castle Lung Cancer Research Programme, The University of Liverpool, Department of Molecular and Clinical Cancer Medicine, Liverpool, United Kingdom*
[t]*Maria Sklodowska-Curie Institute of Oncology Center, Warsaw, Poland*
[u]*Nofer Institute of Occupational Medicine, Department of Environmental Epidemiology, Lodz, Poland*
[v]*Russian N. N. Blokhin Cancer Research Centre, Moscow, Russian Federation*

***Corresponding author.**

[†]Writing on behalf of The COPDGene Investigators.

Disclosures: The authors declare no conflict of interest.

Address for correspondence: Yanhong Liu, PhD, Baylor College of Medicine, 1 Baylor Plaza, TX 77030. E-mail: yl10@bcm.edu

ISSN: 1556-0864

https://doi.org/10.1016/j.jtho.2018.06.016

[w]Faculty of Health Sciences, Palacky University, Olomouc, Czech Republic
[x]Institute of Public Health and Preventive Medicine, Charles University, Second Faculty of Medicine, Prague, Czech Republic
[y]National Institute of Public Health, Bucharest, Romania
[z]International Organization for Cancer Prevention and Research, Belgrade, Serbia
[aa]International Agency for Research on Cancer, Lyon, France
[bb]Princess Margaret Cancer Center, Toronto, Ontario, Canada
[cc]Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, Ontario, Canada

Received 12 April 2018; revised 6 June 2018; accepted 17 June 2018
Available online - 4 July 2018

ABSTRACT

Background: Genome-wide association studies are widely used to map genomic regions contributing to lung cancer (LC) susceptibility, but they typically do not identify the precise disease-causing genes/variants. To unveil the inherited genetic variants that cause LC, we performed focused exome-sequencing analyses on genes located in 121 genome-wide association study–identified loci previously implicated in the risk of LC, chronic obstructive pulmonary disease, pulmonary function level, and smoking behavior.

Methods: Germline DNA from 260 case patients with LC and 318 controls were sequenced by utilizing VCRome 2.1 exome capture. Filtering was based on enrichment of rare and potential deleterious variants in cases (risk alleles) or controls (protective alleles). Allelic association analyses of single-variant and gene-based burden tests of multiple variants were performed. Promising candidates were tested in two independent validation studies with a total of 1773 case patients and 1123 controls.

Results: We identified 48 rare variants with deleterious effects in the discovery analysis and validated 12 of the 43 candidates that were covered in the validation platforms. The top validated candidates included one well-established truncating variant, namely, BRCA2, DNA repair associated gene (*BRCA2*) K3326X (OR = 2.36, 95% confidence interval [CI]: 1.38–3.99), and three newly identified variations, namely, lymphotoxin beta gene (*LTB*) p.Leu87Phe (OR = 7.52, 95% CI: 1.01–16.56), prolyl 3-hydroxylase 2 gene (*P3H2*) p.Gln185His (OR = 5.39, 95% CI: 0.75–15.43), and dishevelled associated activator of morphogenesis 2 gene (*DAAM2*) p.Asp762Gly (OR = 0.25, 95% CI: 0.10–0.79). Burden tests revealed strong associations between zinc finger protein 93 gene (*ZNF93*), *DAAM2*, bromodomain containing 9 gene (*BRD9*), and the gene *LTB* and LC susceptibility.

Conclusion: Our results extend the catalogue of regions associated with LC and highlight the importance of germline rare coding variants in LC susceptibility.

Keywords: Exome sequencing; Rare variants; Lung cancer; Susceptibility loci

Introduction

Although almost 80% of lung cancers (LCs) are attributed to smoking, LC develops in only about 15% of smokers. Therefore, it remains of great importance to understand the genetic factors that contribute to LC risk. It is well recognized that tobacco-induced chronic obstructive pulmonary disease (COPD) is an important predictor of LC risk. Genome-wide association studies (GWASs) have identified 45 genome-wide significant loci for LC, 22 loci for COPD, 32 loci for smoking behavior (SM), and 63 loci for pulmonary function (PF) levels, totaling 121 unique susceptibility loci (Supplementary Table 1). Interestingly, there is considerable overlap among these susceptibility loci and genes for these phenotypes (LC, COPD, SM, and PF levels). For example, the 6p21-22, 15q24-25.1, and 19q13.2 regions are shared by all four phenotypes; 5p15.33, 6p21.32, 10q23.31, and 10q25 are shared by three phenotypes; and 15 loci shared by two phenotypes (see Supplementary Table 1).

Although GWASs have been successful in identifying common (minor allele frequency [MAF] >5%) variants of small effect, the overall amount of LC heritability explained by these known common variants remains small. Further, because the tag single-nucleotide polymorphisms (SNPs) utilized in GWASs are used to identify genomic regions of interest rather than being selected for causality, identification of the functional variant at a specific locus generally poses a significant challenge. For example, of the 93 common LC GWAS top hits from the 45 reported susceptibility loci,[1] only two are protein coding (cholinergic receptor nicotinic alpha 3 subunit gene (*CHRNA3*) p.Tyr215 and cholinergic receptor nicotinic alpha 5 subunit gene (*CHRNA5*) p.Asp398Asn), with the remaining 91 variants falling in noncoding regions (four in untranslated regions, seven in flanking regions, 70 in intron regions, and 10 in intergenic regions). Alleles that are functionally deleterious will tend to be underrepresented at high frequencies, which is an assertion that is supported by the observation of a relationship between putative functionality and MAF. Recent studies suggest that multiple low-frequency (1% < MAF < 5%) or rare (MAF <1%) variants