Defining the genetic architecture of lung cancer etiology.

Christopher I. Amos, Rayjean Hung, Younghun Han2, Yohan Bosse, Xiangjun Xiao2, Yafang Li2, John Field3, Xuchen Zong4, Heike Bickeböller5, David C. Christiani6, Paul Brennan1, Maria-Teresa Landi7, James Dowling Mckay1, on behalf of the OncoArray Lung Cancer Group. 1International Agency for Res. on Cancer, Lyons, France; 2Dartmouth College, NH; 3University of Liverpool, United Kingdom; 4Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, ON, Canada; 5University Medical Centre Göttingen, Göttingen, Germany; 6Harvard School of Public Health, Boston, MA; 7Division of Cancer Epidemiology & Genetics, National Cancer Institute, Bethesda, MD

Lung cancer is considered an archetypical environmentally induced disease because of the high risk from exposure to tobacco smoke. However, family studies clearly identified familial aggregation, beyond effects from familial correlations in smoking behavior. Family studies and genome wide association studies have identified selected variants influencing lung cancer risk but have been underpowered to provide a more comprehensive assessment of genetic architecture. Here, we present the largest genome-wide scan of lung cancer susceptibility, in European-descent populations, comprising data derived from 29,863 patients and 55,586 controls queried for SNPs imputed from the 1000 Genomes Project. Analyses were adjusted for age, sex and the first three principal components for overall lung cancer and stratified by histology (adenocarcinoma, squamous carcinoma, and small cell carcinoma) and ever or never smoking. Results identified 24 loci influencing lung cancer risk for loci with minor allele frequencies of 0.5% or higher, of which 14 had not previously reached genome-wide significance levels. Effect sizes ranged from 0.37 for the functional variants rs17879961 in *CHEK2* to rs11571833 in *BRCA2*, indicating the role of selected uncommon variants in strongly influencing lung cancer risk. Among the novel variants identified, 6 influenced overall lung cancer risk, 6 were specific to adenocarcinoma and 1 each influenced never and ever smokers respectively. Aside from previously described variants in the *CHRNA5* and *CYP2A6* regions that influence cigarette consumption, all other loci showed significant heterogeneity among histologies. Array based heritability analysis also indicates very no significant shared heritability between adenocarcinoma and squamous carcinoma, again pointing to striking etiological heterogeneity other than that due to known smoking behavior related loci between these forms of lung cancer. Evaluation of eQTL results derived from normal lung tissue identified consistent cis-acting effects of the variant rs77468143 influencing expression of a little know gene SECISBP2L, rs6920364 influencing the extracellular ribonuclease *RNASET2*, and rs146729428 in the *EPHX2,*  involved in inflammation. Several loci that influenced adenocarcinoma alone influenced telomere maintenance or cell cycle, while several of the squamous cancer-specific loci are involved in recombination repair. These results indicate striking variation among histological subtypes of lung cancer. Further studies are ongoing to identify the impact that smoking has on lung cancer risk for these loci. These analyses provide a comprehensive assessment of genetic effects on lung cancer risk and will elucidate how these interact with smoking behavior.