Gene expression classifier for prognosis of early-stage squamous cell carcinoma of the lung

Rintaro Noro,1 Naomi Walsh,2 Teruhide Ishigame,1 Kouya Shiraishi,3 Ana I. Robles,1 Brid M. Ryan,1 Elise Bowman,1 Judith A. Welsh,1 Aaron J. Schetter,4 Vidar Skaug,5 Steen Mollerup,5 Aage Haugen,5 Jun Yokota,6 Takashi Kohno,3 Curtis C. Harris1

1NIH-National Cancer Institute, Bethesda, Maryland, USA
2NIH-National Institute for Cellular Biotechnology, Baltimore, Maryland, USA
3The University of Tokyo, Tokyo, Japan
4National Cancer Center Research Institute, Tokyo, Japan
5National Institute of Occupational Health, Oslo, Norway
6Institute of Predictive and Personalized Medicine of Cancer (IMPPC), Barcelona, Spain

Patients with resected stage II and IIIA non-small cell lung carcinoma (NSCLC) are eligible for adjuvant chemotherapy, however its efficacy for stage I patients is still ambiguous. Importantly, adjuvant chemotherapy trials in early stage lung cancer patients have been conducted in unselected patient populations, which may, in part, explain the lack of conclusive evidence of benefit. There are currently no validated methods that prospectively identify NSCLC patients at a high risk of recurrence after surgery. Many gene expression-based gene signatures that identify patients with lung adenocarcinoma (ADC) at high-risk of recurrence have been described. However, data are still lacking on molecular signatures that stratify risk of recurrence in patients with Squamous Cell Carcinoma (SCC), the second most common histological subtype of NSCLC. SCC is commonly detected in heavy smokers, where the risk of lung cancer is closely correlated with tobacco consumption. The disease is heterogeneous, and patient outcome is complicated by comorbidities such as Chronic Obstructive Pulmonary Disease and Cardiovascular Disease. The National Lung Screening Trial (NLST), which compared low-dose computed tomography (LDCT) versus chest radiography, demonstrated a statistically significant mortality benefit of LDCT screening in patients with ADC. However, there was no benefit for patients with lung SCC despite an increase in the detection of early stage tumors. Surgical resection is the recommended treatment for Stage I NSCLC. However, approximately 30% patients will recur and die within 5 years of surgery. In view of the high rate of relapse and the lack of predictive biomarkers, it is critical to develop biomarkers that can identify high-risk early stage lung cancer patients who may benefit from adjuvant chemotherapy or immunotherapy. We present a prognostic classifier based on gene expression that is broadly applicable to diverse patient cohorts as a clinical tool for guiding postoperative management and therapeutic decisions in patients with early stage lung SCC. In order to maximize the potential for discovering functionally relevant biomarkers we focused our study on genes with known mechanistic roles in lung SCC development and prognosis. Specifically, the expression of 256 genes selected by a literature search was evaluated in microarrays from 107 tumors resected from patients with early stage lung SCC. In order to maximize the potential for discovering functionally relevant biomarkers we focused our study on genes with known mechanistic roles in lung SCC development and prognosis. Specifically, the expression of 256 genes selected by a literature search was evaluated in microarrays from 107 tumors resected from patients with early stage lung SCC. Genes significantly associated with relapse-free survival by univariable Cox regression were technically validated by qRT-PCR in the same sample population and externally validated in an independently collected cohort of 91 SCC. A multivariable regression-based classifier was then established on genes independently associated with survival in these
two cohorts by incorporating the resulting coefficients of multivariable Cox regression model into a score that utilizes linear gene expression values. This gene expression classifier was validated in 6 additional publicly available datasets of stage I/II lung SCC (N = 358). The classifier identified high-risk patients in multiple large-scale and geographically diverse cohorts (N = 570). The results appear to be independent of race and gene expression platform. Canonical pathways associated with this signature encompass proteins involved in signal transduction, tissue remodeling, and cell motility that would broadly lead to cancer cell migration, invasion and proliferation, suggesting that the gene signature identifies molecular subsets of patients with clinical relevance. This gene classifier could be used to guide clinical decisions after surgical resection. Thus, we would advocate that this classifier be incorporated into prospective trials for further evaluation of its clinical effectiveness.

A functional SNP in MRPL43 modulates lung cancer susceptibility and survival through alternative splicing of its isoforms

Andrew Mcclary,1 Kara Calhoun,1 Jessica Roberts,2 Ana Robles,1 Elise Eowman,1 Majda Haznadar,1 Dominic Esposito,2 Cassidy Alexandre,1 Mohammed Khan,1 Rama Modali,3 Hirokazu Okayama,4 Kouya Shiraishi,4 Yun Yokota,4 Takashi Kohno,4 Brid Ryan1,4 Laboratory of Human Carcinogenesis, Center for Cancer Research, National Cancer Institute, Bethesda, MD,2 Frederick National Laboratory, NCI, Frederick, MD,3 Bioserve Ltd, Gaithersburg, MD,4Division of Multistep Carcinogenesis, National Cancer Center Research Institute, Tokyo, Japan

miRNAs, a class of non-coding genes, modulate mRNA translation and stability by primarily binding to the 3’UTR of mRNA transcripts. Genetic variation within a miRNA gene could modulate the thermodynamic interaction between the miRNA and its target mRNA sequence and modulate expression. Many studies have demonstrated the importance of miRNAs in cancer biology, including lung cancer. Indeed, we recently published several studies showing the relationship between functional SNPs in microRNA binding sites and the relationship with lung cancer risk and outcome. However, few studies that systematically examined SNPs in miRNA genes and their relevance to lung cancer have been conducted to date to our knowledge. We conducted a genome-wide analysis of SNPs in microRNA genes and assessed their relevance to human lung cancer. Using samples from the NCI-MD study, we examined 22 SNPs (after selections procedures) in 974 European Americans and found that rs4919510 (Chr10 q24.31) in miR-608 was associated with a 2.55 fold increased risk of lung cancer, after adjustment for age, gender, pack-years of smoking and smoking status (OR 2.55 95% C.I. 1.04-6.28; P=0.041). The G allele of rs4910510 is common; the MAF is 18%, 57% and 53% in European Americans, African Americans and Japanese, respectively. We therefore tested for population conversion of this epidemiological observation in 2 other geographical populations, i.e., an African American population and a Japanese population. In both studies, the G allele of rs4919510 was associated with an increased risk of lung cancer (African American: OR 2.73, 95% C.I. 1.44-5.14; P=0.002; n=566) (Japanese: OR 1.85, 95% C.I. 1.18-2.86; P=0.007; n=768). We also found that rs4910510 was associated with lung cancer survival. The GG genotype of rs4910510 was significantly associated with prolonged survival among European Americans (HR 0.15, 95% C.I. 0.04-0.62; P=0.008) and Japanese patients (HR 0.66, 95% C.I. 0.45-0.97; P=0.036), while a similar trend was observed in African Americans (HR 0.64, 95% C.I. 0.35-1.16; P=0.1). An eQTL analysis using samples from the NCI-MD study found that the SNP did not affect miR-608 processing and functional studies comparing the function of miR-608-C and with miR-608G did not reveal any differences in either mRNA targeting or function. We subsequently mapped the functional significance of the rs4910510 association to a neighboring gene, MRPL43, which is a component of the mitochondrial ribosome. Our results show the miR-608 SNP is actually in linkage with a SNP in MRPL43 that creates a donor splice site at the exon 3/intron 3 boundary. Thus, rs4910510 is a tag allele for the real functional locus in MRPL43. The SNP changes the splicing pattern of MRPL43, whereby expression of the short and predominant isoform is decreased, and the longer isoforms is increased (P<0.0001). We replicated this observation in two different lung cancer tissue sets and also 6 other tissue types using TCGA data (all P<0.000001) and also across populations of different ancestry. We further found that the SNP modulates mitochondrial metabolism by altering the balance between oxidative phosphorylation and glycolysis. Our initial study aimed to evaluate the association between microRNA SNPs and lung cancer. In summary, we did not observe a strong relationship between miRNA SNPs with either lung cancer risk or survival. However, we have identified a SNP in MRPL43, a subunit of the mitochondrial ribosome, which is associated with risk